ANNENBERG
BROWN UNIVERSITY

EdWorkingPaper No. 23-791

Can brick phones bridge the digital learning divide? Evidence from SMS-based math practice

Torsten Figueiredo Walter
New York University

Guthrie Gray-Lobe
University of Chicago,

Sarah Kabay
Innovations for Poverty Action

Hardware requirements are a barrier to widespread adoption of digital learning software among low-income populations. We investigate the demand among smallholder-farming households for a simple, adaptive math learning tool that can be accessed by widely available ``brick'' phones, and its effect on educational outcomes. Over a quarter of invited households used the tool, with greater demand among households lacking electricity, radios, or televisions. Usage was highest when schools were out of session. Engagement lapsed without regular reminders to use the service. Using random variation in access to the service, we find evidence that the platform increased test scores, school attendance, and grade attainment. Interpretation of these estimates is complicated by potentially endogenous outcome observation.

VERSION: June 2023

Suggested citation: Figueiredo Walter, Torsten, Guthrie Gray-Lobe, and Sarah Kabay. (2023). Can brick phones bridge the digital learning divide? Evidence from SMS-based math practice. (EdWorkingPaper: 23-791). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/j4kc-kq39

Can brick phones bridge the digital learning divide? Evidence from SMS-based math practice*

Torsten Figueiredo Walter† Guthrie Gray-Lobe‡ Sarah Kabay§

June 14, 2023

Hardware requirements are a barrier to widespread adoption of digital learning software among low-income populations. We investigate the demand among smallholder-farming households for a simple, adaptive math learning tool that can be accessed by widely available "brick" phones, and its effect on educational outcomes. Over a quarter of invited households used the tool, with greater demand among households lacking electricity, radios, or televisions. Usage was highest when schools were out of session. Engagement lapsed without regular reminders to use the service. Using random variation in access to the service, we find evidence that the platform increased test scores, school attendance, and grade attainment. Interpretation of these estimates is complicated by potentially endogenous outcome observation.

*The evaluation received support from the Echidna Foundation, the Douglas B. Marshal, Jr. Family Foundation, Innovations for Poverty Action, and the Development Innovation Lab. We thank Noam Angrist, Amrita Ahuja, Emmanuel Bakirdjian, Jonathan Lehe, Daniel Mokrauer-Madden, and Theresa Solenski for valuable feedback. We thank Nick Ayugi, Jared Otuke, Allison Stone and Asman Suleiman for their support in collecting data. IRB approval was received from Human Subjects Committee for Innovations for Poverty Action IRB-USA (#15719) and from the Maseno University Ethics Review Committee (MUERC) (#00904/20). Data was collected under a National Commission for Science, Technology and Innovation (NACOSTI) permit. The study received IPA country office special approval for face-to-face training of interviewers.

†New York University Abu Dhabi, email: t.f.walter@nyu.edu

‡University of Chicago, email: graylobe@uchicago.edu

§Innovations for Poverty Action, email: skabay@poverty-action.org

1 Introduction

Digital learning is least accessible to children who have the most to gain from it. In sub-Saharan Africa, it is estimated that three-quarters of families are unable to use the internet to access educational materials, and access is even more limited in rural areas (Dreesen et al., 2020; Ngware and Ochieng, 2020). Households that lack electricity, smartphones, or computers – requirements for accessing digital learning – have poor academic outcomes on average (Ngware and Ochieng, 2020). Creating tools that connect to more widely available technology could reduce inequality in access to digital learning and educational outcomes. A number of digital learning tools have emerged in LMICs based on this premise, such as Eneza Education and M-Shule in East Africa, Text2Teach in the Phillippines, and Ustad Mobile in Afghanistan. However, such tools tend to have fewer features than those developed for computers or smartphones. It is unclear if, without these features, these tools will have widespread appeal for families and improve educational outcomes.

The COVID-19 pandemic made the digital learning divide more apparent (UNICEF, 2020). Governments and aid agencies emphasized widely available mass-communication technologies, such as radio or television as a way to reach lower-income households during school closures. However, in many cases, these were not widely used. In Kenya, the government went so far as to cancel the 2020 academic year, citing concerns that many children would be unable to access high-quality home-learning tools (Dahir, 2020).

We report on a large-scale evaluation of ElimuLeo – an adaptive math learning tool that can be accessed by any mobile phone – among smallholder farmers in Kenya. Cellular networks provide the most expansive network for interactive communication with low-income households. In Kenya, nearly eighty percent of the population had access to a mobile phone in 2017 (Pew Research Center, 2018), and the cost of sending a single message is less than half a cent. ElimuLeo – developed by the NGO Precision Development (PxD) in collaboration with the Kenya Institute of Curriculum Development (KICD) – is a free service that provides children with access to adaptive retrieval practice covering primary-school-level mathematics topics. Using a two-way SMS shortcode, the system provides practice problems, real-time feedback on responses, and automatically adjusts the difficulty of problems based on the accuracy of prior responses.

We examine the demand for and impact on educational outcomes of ElimuLeo from two large-

scale randomized evaluations. Study participants were active users of PxD's mobile phone-based agricultural extension services.[1] The first experiment invited households to join the platform in December 2020, while schools were closed due to the pandemic. Among the households that signed up, a random subset was offered early access and received regular reminders to use the service while others were placed on a waiting list until August 2021. The second experiment invited households to sign up in April 2021, when schools were closed for holiday. In this experiment, all households who signed up received access to the service. A random subset was provided with reminders to use the service, while the comparison group did not begin receiving reminders until August 2021. In the first experiment, treated households completed on average 32 problems compared to zero in the control group. In experiment 2, treated households completed 41 problems on average between April and June 2021, compared to 14 problems among control users.

Demand for ElimuLeo is highest among lower socioeconomic-status households. Take up is inversely related to measures of regional development including ward-level night lights, county-level net primary school enrollment rates, and the fraction of children in a county who are enrolled in an age-appropriate grade. Conditional on opting in to use the service, households without electricity, radios, or televisions spent more time on the service and completed more problems. These results suggest that ElimuLeo may be less appealing to wealthier households who likely have a larger menu of home-learning options or are in less need of home-learning tools.

We find evidence that access to ElimuLeo has positive impacts on academic outcomes.[2] Math skills are measured using data generated by usage of the platform. Grade attainment is measured by an independent phone survey. Children who were experimentally induced to use the platform more exhibit more advanced math skills and are less likely to be retained in grade. While random assignment ensures that households assigned to treatment are comparable to those in the control condition, it does not ensure comparability of those households for which outcome data is available. Identification of the effect of access to the platform hinges on an assumption that the treatment and control groups are similar *conditional on having outcome data*. We provide some empirical evidence in support of this assumption, and that, if anything, treated households may be negatively selected, potentially biasing positive academic impacts toward zero. However, follow-up rates on

[1] "Active users" were those that had used the service in the previous 12 months.

[2] For ease of exposition, we use the phrase "access to ElimuLeo" to refer to giving applicants access *and* regular reminders to use the service.

these outcomes are too low to confidently rule out contamination by selection.

Two-way SMS platforms may be a cost-effective tool to expand access to interactive, and adaptive, math learning tools. The marginal cost of providing students with 100 practice problems through the service is approximately 80 cents at market rates. At large scale, governments could lower costs by obtaining bulk discounts or marginal cost pricing from telecommunication providers. The program's cost-effectiveness is further supported by evidence of positive sorting: households with the most to gain from the service appear to have the greatest demand for it. Because costs scale with usage, selective take up diverts resources from those who have the least to gain.

The results in this study may understate the appeal and effectiveness of well-designed, smartphone-accessible tools that incorporate more features without interfering with users' ease of access. ElimuLeo is rudimentary compared to similar tools that have been developed to reach marginalized populations. Eneza Education and M-Shule – two prominent examples in East Africa – both provide a more complex interface including a menu of options to access both instructional content and practice materials.

This paper contributes to a literature that evaluates the impact of computer-assisted and adaptive testing in LMICs (Banerjee et al., 2007; Muralidharan, Singh and Ganimian, 2019; Muralidharan and Singh, 2021; de Barros, Ganimian and Venkatachalam, 2022). Estimated effects of computer-aided adaptive retrieval practice are among the largest of any educational intervention in LMICs (Evans and Yuan, 2019; Angrist et al., 2020*a*). However, these programs typically require expensive technological hardware such as personal computers, tablets, or smartphones. Our study contributes to this literature by evaluating the effect of a simple, low-tech method of providing a similar type of practice via mobile phone.

This study also relates to literature on low-cost tools to reduce learning loss when students are kept out of school for long periods of time. Studies have found that take-up of remote-learning tools like educational radio and television programs, as well as internet-based activities, was low, especially among poorer households and those in rural areas (Wolf et al., 2021; Oza and Cilliers, 2021). Most rigorous evaluations of interventions to reduce learning loss during the pandemic have leveraged various forms of adult tutoring. Carlana and Ferrara (2021) show that volunteer tutors helped disadvantaged students in Italy during the pandemic. Angrist et al. (2020*b*) show that a system that combines weekly SMS practice math problems with teacher calls to parents

was able to improve numeracy in Botswana. Angrist et al. (2020*b*) also evaluate an intervention that provided weekly one-way (non-interactive) SMS messages and find positive, but statistically insignificant effects on test scores. We note two major differences between that study and ours: first, ElimuLeo is designed to provide a larger volume of problems and to automatically adapt to the child's responses without necessarily involving the parent directly; second, the program studied in Angrist et al. (2020*b*) was parent-facing (it provided information for parents to be effective tutors).

The remainder of this paper proceeds as follows. Section 2 discusses relevant background for this study, including access to mobile phones and other communication technologies in Kenya, the history of school closure during the pandemic, and the details of the ElimuLeo service. Section 3 describes the experimental design and the data used in this study. Section 4 presents the results and section 5 discusses mechanisms and external validity and concludes.

2 Background

2.1 Educational inequality in Kenya

Children in rural Kenya tend to lag behind their urban counterparts on measures of academic success. Standardized assessments have found that only 25 percent of rural Grade 3 students can complete Grade 2 work, compared to 41 percent of urban students (Uwezo, 2016). Conditional on age, rural students also complete fewer years of schooling. Figure 1 shows the fraction of children who have reached or surpassed their age-appropriate grade using data from the 2019 Kenya Continuous Household Survey Programme (KCHSP) (KNBS, 2019) separately for rural and urban students.[3][4]At age 5, rural children are about 5 percentage points less likely to have reached their age-appropriate grade. This gap doubles by age 14.

Rural Kenyan children also lack access to resources needed to make use of digital learning. Only 17 percent of rural households report accessing the internet compared to 44 percent of urban households (World Bank Group, 2019; GSMA, 2017). More fundamentally, rural children have less access to electricity. Only 16 percent of children aged 5-14 live in homes lit by electricity.

[3]We define age-appropriate grade as the grade where a student would be found if they entered primary school at age seven and progressed one grade each year.

[4]Although formally prohibited in public schools, Kenyan students repeat grades frequently (Wanzala, 2018, 2019; KNEC, 2020*a*). Common reasons for grade repetition are poor academic performance and chronic absenteeism (KNEC, 2020*b*). In many cases, repetition is requested by parents (KNEC, 2020*b*).

For children in the most marginalized households, these gaps are even more striking. Among households without electricity, around 80 percent start primary school by age seven, but the fraction reaching the next grade at age eight falls to below 70 percent (Figure 2).[5] By age fourteen only around 40 percent of children from households without electricity are in their age-projected grade, suggesting that nearly half of children who started primary school on time will repeat a grade in primary school or drop out entirely.[6]

2.2 School closures

The Kenyan Ministry of Education ordered schools to close on March 15, 2020, to control the spread of COVID-19. Schools remained closed until January 4, 2021. An exception was made for students who were in grades that were preparing for national examinations: children in the fourth, eighth, and twelfth grades. School closures reduced students ability to access teachers for education, with especially large impacts in rural areas (Pape et al., 2020). Schools closed again from March 19 to May 10 for holiday before the start of the 2021 academic year.

Numerous resources for distance learning were available for students in Kenya over this period, although usage of these resources was low. The Kenyan Institute of Curriculum Development (KICD) provided content, free of charge, that was accessible via the internet, including videos, links to free apps, and practice problems (Ngware and Ochieng, 2020). Television and radio programs were available as well, with some of these programs developed by KICD. Use of radio, television, and internet-based educational resources was low in Kenya. Data available from the Kenya COVID-19 Rapid Response Phone Survey 2021 show that the most popular resource while schools were closed was "self-directed study using textbooks and other revision materials" (40 percent of wave 3 respondents). Less than six percent of households reported using radio and television for education, 2.3 percent reported using internet-based tools, including smartphone applications. Zero children reported using SMS-based tools.

[5] Access to electricity is imputed from the households reported primary energy source used for lighting. We treat households who use solar energy for lighting as having no access to electricity, as we think this is more directly comparable to data from the present study.

[6] It is possible that the fourteen-year-old cohort in 2019 started school much later than the six-year-old cohort. However, changes in age-of-entry to primary school over time are unlikely to explain this gap.

2.3 Precision Development

Precision Development (PxD) is a global non-profit that uses mobile technologies to support smallholder farmers, students, and other vulnerable populations in LMICs. In 2022, PxD reached 7.3 million rural households globally with digital agriculture advisory services. In Kenya, PxD reaches over 500,000 households with its mobile advisory services for agriculture (MoA-INFO).

The population of PxD users is predominantly smallholder farmers. A survey conducted in the summer of 2020 with a random sample of users of PxD's agricultural extension services found that 27 percent of their school-aged children were not involved in any educational activities while schools were closed. As noted in Table 1, many households that use the service lack access to electricity, radio, and television – markers of socioeconomic disadvantage and technological assets that are important for accessing distance-learning resources.

2.4 ElimuLeo

ElimuLeo is an automated, two-way SMS platform that enables children to practice math skills and receive feedback on the accuracy of their responses. The service was developed through a collaboration between PxD and the Kenya Institute of Curriculum Development (KICD) in 2020. ElimuLeo was originally conceived as a distance-learning tool that could be deployed, free of charge, to support children during school closures.[7][8] The platform is both responsive and adaptive, informing users whether their answers are correct and adjusting further content based upon users' performance. All prompts, menus, and problems on the service can be displayed on a brick phone.

Topics covered by the service include addition, subtraction, multiplication, division, and algebra. For grades 7 and 8, the system also offers "mixed" and "combined" practice modules.

The difficulty of items is adapted to users' prior performance. Each item was assigned a difficulty rating corresponding to the specific academic grade-term in which the topic is ordinarily covered in primary school. Ratings were based primarily on the Kenyan national primary school curriculum, with minor deviations where the curriculum was non-specific.[9] As an example, Grade 6 Term 1

[7]The project was initially conceived of as an emergency intervention to address Kenya's extended school closures due to the coronavirus pandemic.

[8]Users can choose to interact with the service in either English or Swahili.

[9]The competency-based curriculum is being phased in starting with younger grades. We used the curriculum that was currently prevailing at the time of the intervention. Grades 1 through 4, content was developed based on the Competency-Based Curriculum (KICD, 2017, 2019). Grades 5 through 8 content was based on the 8-4-4 curriculum

math practice focuses on questions that add two numbers with up to six digits, including two digits after the decimal with no regrouping after the decimal. Grade 6 Term 2 practice problems are similar, except that the problems include regrouping after the decimal. A class 7 algebra item would ask students to solve for x where x minus a one-digit number is a two-digit number. The system is designed to automatically produce new problems that satisfy these criteria: it would randomly choose Z, a single-digit number, and Y, a two-digit number, and ask the student to solve for x in $x - Z = Y$.

The service automatically adapts to a student's level of ability. New users of the service enter their grade level. After that, they are provided with a menu of topics available at that grade level. The first time a topic is chosen, the service assesses the student's skills based on a series of five questions. The first question that a student receives corresponds to the topic corresponding to the first term of their initially stated grade level. If they answer the question correctly, they are given a question corresponding to a grade level above. If they respond incorrectly, they move down a grade level. The terminating point results in their assessed level.

After the initial assessment, students can choose to practice the topic further. Students complete five questions at their currently assessed grade-term level at a time. If they complete at least four out of five correctly, they advance to the next grade-term level. If they are unable to complete any correctly, they move down to the preceding grade-term level. Therefore, during practice, the platform is less sensitive to correct and incorrect responses, moving only one term level at a time.

Students may choose to practice other topics at any time after they select their first topic. When they do so, the system uses their most recent assessed grade level to determine the starting point of the assessment algorithm.

3 Experimental design and data

This section describes recruitment of study participants and randomization in two experimental evaluations.

(KICD, 2002).

3.1 Experiment 1

PxD invited 10,618 active users of their agricultural extension service to opt into the service in December 2020.[10] Out of those, 4,551 opted in to use the service.[11] Among those who opted into the service, 2,382 were assigned to receive early access to the service. The remaining 2,150 were placed on a waitlist. Random assignment was stratified at the county level.

At the beginning of the program, treated households received a message informing them that they had been granted access to the service. Subsequently, households received between one and two messages each week reminding them to use the service. Households in the control group were informed that they would be placed on the waitlist. These households received access to the service in August 2021, although this date was not provided in the initial message informing them that they had been waitlisted.

3.2 Experiment 2

PxD invited 19,625 active users of their agricultural extension service to opt-in to ElimuLeo in late March 2021. Out of these, 6,842 opted in. Forty-one phone numbers were selected before randomization as part of a qualitative phone call activity and were removed before assignment, leaving a sample of 6,801. In contrast to Experiment 1, all households that opted in initially received access to the service.[12] On March 30, households were randomly assigned to receive the regular reminder component of the service. Treated households continued to receive reminder messages, whereas control households did not. Randomization was stratified based on the household's county and assessed addition grade level after three days using the service. Households that had not completed an assessment in math (either because they had not completed any assessment, or because they had only been assessed in other topics, such as subtraction, multiplication or division) were treated as being assessed at level zero.

[10] Active users were those that had responded to a text message related to agricultural extension in the past year. Users who were currently participants in other studies related to the extension service were excluded.

[11] Opting in required the user to answer a series of questions about whether a child in the household was in primary school

[12] This design was chosen to try to target the sample at a population that was more likely to use the service intensively. By conditioning on respondents who already demonstrated an interest in using the service, it was hoped that the first stage impact on platform usage would be larger. In the end, this design generated a similar first-stage impact on usage due to the higher level of usage in the counterfactual.

3.3 Data

We combine user data maintained by PxD and a phone-based endline survey conducted by Innovations for Poverty Action (IPA) to conduct the analysis.

3.3.1 Baseline data

All of the data observed before randomization comes from PxD's agricultural extension system. PxD's users are identified by the mobile phone number that had been used to access mobile agricultural extension services in the past. Distinct mobile phone numbers are treated as the unit of analysis and we refer to them as households throughout.[13] In addition to mobile phone numbers, PxD's data contain information on past usage of the agricultural extension system, as well as county and constituency of the household. Geographic data are based on self-reports from the user. PxD data do not contain information on household composition, including whether the household has children.

3.3.2 Application usage data

Usage of the service is observed directly through the ElimuLeo system. These data include a complete report of all exercises completed on the service. These data record, at the topic level, the number of questions answered, the number correct, and the date and time at which the household completed the lesson, measuring take-up on both the extensive margin (any usage of the service) and the intensive margin (number of questions answered).

Math skills are measured using data from performance on the ElimuLeo platform post-intervention. We use responses to the initial assessments and practice problems to determine the grade level corresponding to the student's math skills. Data shared by PxD allow for identification of the assessed level of the household as of October 2021 (two months after comparison households received full access to the service) in each of the topic areas that the household has accessed as well as the last date that the household used the service. We use the maximum assessed grade level of the student across all subjects to construct a single scalar measure of each student's assessed grade level.

Note that all usage data are observed at the household level. For households with more than

[13] We consider identification concerns related to this problem in the discussion section.

one child, usage represents the combined usage of all children. Maximum assessed grade level at different dates may represent different children (or other household members) within the same household. We discuss robustness of our results to restricting to households with a single primary-school-aged child in Section 4.2.1. Survey data (introduced below) indicates that, on average, 1.3 children used ElimuLeo within a household, conditional on there being any children who reported using the system.

3.3.3 Household endline survey

A phone survey collected data on all children aged six to fourteen in the household. The survey was conducted between June and July of 2021. Respondents were, in nearly all cases, parents. For each child, the survey collected information on school enrollment, grade attainment, and whether the child was in school on the day of the interview. Additionally, the survey collected data on learning activities during the school break between the 2021 and 2022 academic school years and children's time use. For home-learning activities, respondents were asked to report activities that each of their children had engaged in without prompts.[14] Two survey items collected information on time use. First, respondents reported on the number of hours in the previous week that each child had spent doing educational activities, caring for young children, doing household chores, and working to earn income. Second, respondents were asked to name activities that children had engaged in when schools were closed.[15]

A callback survey was required after an error was discovered in the survey protocol. An item was mistakenly added to the beginning of the survey that asked households whether they had any recollection of receiving information about the ElimuLeo platform. Households that answered negatively to these questions were then screened out of the activity.[16] All households had been contacted at some point regarding ElimuLeo, however, these items systematically screened out control households, because treatment status in Experiment 1 involved much more intensive messaging regarding the platform. The error also may have resulted in lower follow-up rates overall if some

[14]This question asked "Which learning-related activities has [child] done in an average week in April when schools were closed?".

[15]This question asked specifically "Think of a typical day this past April when schools were closed, what did [child's name] spend their time doing?"

[16]The survey was supposed to skip only those specific questions related to ElimuLeo for households that had no recollection of ElimuLeo.

respondents were not familiar with the service.

The callback survey was completed in October 2021. All households that were successfully reached during the initial survey but screened out as a result of their responses in the survey were called back. All aspects of the interview were the same.

3.3.4 Survey follow-up

The phone survey attempted to reach 4,410 households in Experiment 1, and 4,408 households in Experiment 2. Households with access to ElimuLeo were oversampled due to an error in the survey protocol (see Table A2).[17]

In Experiment 1, the survey successfully interviewed 53 percent of control users who were called, compared to 61 percent of treated households (Table 1). In Experiment 2, 49 percent of control households were interviewed, compared to 52 percent of treated households (Table 1).[18]

Among interviewed households, treated households were more likely to report having children. In Experiment 1, 84 percent of households in the comparison group reported having children. Treated households were 4 percentage points more likely to report having any children. In Experiment 2, 90 percent of households in the comparison group reported having children. Treated households were 3.5 percentage points more likely to report having children. There could be many reasons for this result. First, if survey responses were higher among treated households due to reciprocity, then treated households with children (those households who would have made use of the service) would account for a larger share of households in the treatment group than in the comparison group. It is also possible that access to ElimuLeo had a causal impact on the residential location decisions of children between alternative households.[19] We feel that the latter interpretation is unlikely.

Treatment and control households are similar in terms of observable household characteristics.

[17] the lists provided to interviewers each day ordered treated households first, resulting in oversampling of the treatment group. However, attrition due to non-attempts in the survey could not have been selective because the lists of users provided to the survey firm by PxD were randomized within those experiment and treatment stratus groups and no additional information was shared with the survey firm aside from the user's phone number. Non-representative sampling within the experiment-treatment cells would not, therefore, be possible.

[18] Survey response rates reflect the combined follow-up in the initial survey and the callback activity to address the error in the survey program that had screened out some applicants. We report impacts on this margin in Appendix Table A2).

[19] For example, a child may have been more likely to be sent to a relative's house with greater access to educational resources.

Low follow-up rates and differential attrition raise concerns about the validity of estimated effects using the survey data. Differences between treatment and control could reflect selection in terms of the kinds of households that choose to participate in the survey. We compare treatment and comparison households conditional on being interviewed and having children.[20] Even columns in Panel B of Table 1 test differences in several household-level covariates. Differences between the two groups tend to be small and statistically insignificant. Treated households may be less likely to have electricity, own radios, or televisions. Only radio ownership is statistically significant (and marginally so). The results reported below are robust to including these characteristics as controls.

4 Results

4.1 Demand for the service

Whether households would want to use such a rudimentary system was uncertain prior to the study. Because the ElimuLeo platform was developed in the context of a national effort to control the transmission of COVID-19, the platform was designed specifically so that it would not require users to interact with anyone outside their home. While deliberate, this feature also increased the risk that the platform would be underutilized due to a lack of social interactions which feature in traditional academic environments (Banerjee and Duflo, 2014). There was also uncertainty about whether parents would allow children to use their phones, whether children were interested in doing more educational activities at home, and whether the invitation to use the service would be perceived as a scam.

We examine the demand for ElimuLeo using data on whether households opted in to receive the service and subsequently used the service. Households are linked to measures of ward-level night lights, population density, as well as baseline measures of county-level outcomes including the county net primary school enrollment rate and the fraction of children aged six to fourteen who are found in their age-appropriate grade as defined above. Data on night lights are from Visible and Infrared Imaging Suite (VIIRS) Nighttime lights maps (Evlidge et al., 2013, 2017). Data on population density come from the Global Human Settlement data (Schiavina, Freire and MacManus,

[20] We make the restriction that respondents have children because the main causal estimates using the survey data are at the child level and so implicitly condition on having children.

2022). County-level academic measures are calculated from the 2019 KCHSP. Household-level data come from the survey described above. In the analyses below, we pool Experiments 1 and 2 for conciseness. Results are similar when analysis is conducted separately.

4.1.1 The decision to opt in

The decision to opt in and use the service was likely influenced by many factors, not least whether the household has any children who the phone owner thinks would benefit from the service. The invitation to opt in made clear that PxD expected the service to be most valuable for children in primary school. Ninety-one percent of interviewed households reported having at least one child aged six to fourteen, compared to 57 percent of Kenyan households nationally (Kenya National Bureau of Statistics, 2016).

Demand for the service was high. In Experiment 1, 43 percent of households invited to use the service opted in, compared to 38 percent of households in Experiment 2.[21] Lower demand from households in Experiment 2, which started during a planned academic break after schools had reopened from pandemic closures could be an indication that demand for the service would be lower when schools operating normally.

Measures of regional development are inversely correlated with take-up. Columns (1) through (6) of Table 2 report the bivariate relationship between each regional variable and the fraction of households that opt-in. A doubling of the level of night lights is associated with around a 1.5 percentage point reduction in the fraction who opt in (Column 1). Households in counties with better educational outcomes are also less likely to opt-in. A ten percentage point increase in county-level net primary-school enrollment is associated with a 0.4 percentage point reduction in the fraction who opt in. A ten percentage point increase in the fraction of primary school students who are in their age-appropriate grade is associated with a 1.4 percentage point reduction in the fraction who opt-in.

[21] We treat any household who ever chose to use the service as having opted in. In Experiment 2, 602 households opted in after households were assigned to treatment conditions bringing the total number to 7,444.

4.1.2 The decision to spend time on the platform

Treatment induced greater usage of the platform. In experiment 1, treated households completed on average 32 problems between December 2020 and June 2021. Control households completed zero problems.[22] In experiment 2, treated households completed 41 problems on average between April and June 2021, compared to 14 problems among control users. Figure 5 shows the cumulative distribution of the number of questions completed on the service across households.

Use of ElimuLeo is highest among those households with fewer assets, consistent with the hypothesis that the tool is most valuable for households that lack other study resources. Table 3 reports results from OLS regression of the number of problems answered over the course of the intervention on household characteristics. Households with electricity complete on average 14 fewer problems (Column 4). Households with radios complete 13 fewer problems (Column 5), and those with televisions complete 12 fewer (Column 6). Households with greater assets may have access to more-preferred study options, although these results could also be explained by, for example, differences in tastes.

The impact of giving households access to Elimueleo was similar on average in both experiments. However, there are notable differences in when people used the service. In Experiment 1, usage was concentrated in earlier periods and spanned two periods when schools were closed (Figure 7). Usage in Experiment 2 is more recent and spans only a single period when schools were closed (Figure 8).

4.2 Effect of access on academic outcomes

4.2.1 Effect on math skills

We estimate the effect of treatment on the maximum assessed grade level associated with a user account across mathematics topics. Interactions with the platform generate diagnosed levels of mathematics competencies at a given point in time. Transforming these data into a form that is amenable to conventional impact evaluation requires decisions that are relevant to the research design. We discuss these decisions here.

[22] A single control household used the platform through a separate program that was piloting recruitment processes through teachers.

To begin, we present a graphical description of the progression of the assessed mathematics level for each account that ever used the service over time. For each account, we record the maximum assessed level across subjects on each day. If a household does not use the service on a given day, the daily value is imputed to be the level from the most recent usage. Figures 9 and 10 show the progression of assessed grade level for accounts in treatment and control over time in Experiments 1 and 2 respectively.[23] The two figures lead to strikingly different conclusions about the impact of the study. In Figure 9, no level is observed for the control group until after August 3, when control households received access to ElimuLeo. Treated households advance almost 0.5 grades between January and August, making the most progress during periods when schools are closed. When control households finally receive access, they perform about 0.2 grades *above* the level of treated households. In Experiment 2, treated households advance about 0.4 grades from April to August. Because in Experiment 2 all households had opted into the service and many had completed some exercises, assessed level is observed for control households from the start of the intervention. In Experiment 2, the assessed level is flat until August when control households begin to converge to the treated households. By October 2021, treated households appear to be around 0.25 grades above control households. We discuss below an empirical framework that can reconcile these two results.

Study participants chose whether and when to engage with the service, presenting several identification challenges. First, if a household does not use the service, no outcome is observed. Comparisons between treatment and control may therefore reflect selection, conditional on using the platform. Second, the timing of usage is endogenous to treatment. Households typically use ElimuLeo frequently after receiving access and then gradually slow down. Because control households began receiving full access much later than treated households, their assessed level will tend to be a more recent record of their math skills. If children's math skills advance or decline over time, comparisons of treatment and control households may reflect both the causal effect on learning as well as the effect on when learning was measured in a child's lifetime.[24]

[23]Randomization strata fixed effects are partialed out and the aggregate control mean added in to adjust for differences across strata.

[24]We discuss evidence below that suggests that there is an overall positive trend in math skills. Specifically, in Experiment 1, where treatment had a large positive impact on the date at which households received access and the initial diagnostic assessment, children in the comparison group advanced between one quarter and one-third of a grade level.

We illustrate the problem formally using a simple two-period value-added model of assessed level. Consider the following model of the level of student i in time period $t \in \{1, 2\}$:

$$y_{i1} = \alpha_i + \beta_i D_{i1} + \theta_i + \varepsilon_{i1} \tag{1}$$

and

$$y_{i2} = 2\alpha_i + \beta_i \left(D_{i1} + D_{i2}\right) + \theta_i + \varepsilon_{i2}, \tag{2}$$

where α_i represents the average amount that student i advances between any period t and $t+1$, β_i represents the amount that student i student advances from using ElimuLeo in period t ($D_{it} \in \{0, 1\}$), $\theta_{i,0}$ is the accumulated effect of prior educational inputs, and ε is an idiosyncratic disturbance term specific to period t (e.g. test error). In the first period, $t = 1$, control households, $Z_i = 0$, do not have access to the service, whereas in the second period $t = 2$, all households have access to the service.

We model the decision to use the service in period t using a threshold crossing model. The decision to use the service in period t is modeled as $D_{it} = \mathbb{1}\{\lambda_i \geq \tau + \delta M_{it}\}$, where λ_i is distributed $\mathcal{U}[0, 1]$ and M_{it} represents the number of periods prior to t that a household has had access to ElimuLeo.

Proposition 1 *The estimated impact of ElimuLeo from a comparison of treated and control households yields the following*

$$\begin{aligned}\rho = E[y|Z_i = 1, D_i \geq 1] - E[y|Z_i = 0, D_i \geq 1] \\ = \pi E[\beta_i|\lambda_i \geq \tau + \delta] - (1 - \pi)E[\alpha_i|\tau \leq \lambda_i \leq \tau + \delta]\end{aligned} \tag{3}$$

where $\pi = 1 - \frac{\delta}{1-\tau}$ is the first stage impact on second period usage among the subset of households with $\lambda_i \geq \tau$, and $D_i = \sum_t D_{it}$.

(See Appendix C for derivation.) Importantly, ρ contains the causal (conditional-on-usage) intent-to-treat impact on those households who use the service in the second period, $\pi E[\beta_i|\lambda > \tau + \delta]$, but also reflects the impact on the timing of when tests are observed, $(1 - \pi)E[\alpha_i|\tau \leq \lambda \leq \tau + \delta]$. Observation of the outcome for control students is from older children. If, for example, older

children score higher when they use the service, this will bias the estimated impact toward zero.

One potential solution is to restrict to those scores that are observed in the second period $t = 2$, so that assessments would be more comparable in terms of the age of students. The cost of this approach would be a greater risk of selective attrition: in the second period, it is possible that the subset of treated and control households who use the platform have different α_i or θ_i and that these could be driving the impacts.

We report results restricting the analysis to households for whom the assessment is based on usage past a certain *endline start date*, the earliest date at which usage on the service is treated as measuring a student's endline math level. Modification of the endline start date influences the follow-up differential. Furthermore, restricting the sample to households who used the service in a narrower time window reduces the scope for bias due to differences in the timing of when outcomes are observed.

Overall, we find large positive impacts of earlier ElimuLeo access on assessed grade level regardless of the endline start date (Tables 4 and 5 for Experiments 1 and 2 respectively). Panel A reports the impact on the follow-up differential given the endline start date. Earlier endline start dates are associated with a large positive differential and later dates are associated with a large negative differential (Panel A) because households use the service more intensively shortly after they receive access. We find that when the endline start date is June 15 (Column 3), the rate of follow-up is similar between treatment and control. Changing the endline date also affects the treatment differential in terms of the timing of the last assessment (Panel B). Using an endline start date of August 1, we find no difference between treatment and control in the date of the last time the household used the service. At least in Experiment 2, households are similar in terms of their baseline reported grade level, regardless of the endline start date used (Panel C Table 5). In Experiment 1, ElimuLeo access is negatively associated with initial grade level, which is expected given that the grade level for control students is observed long after treated households began using the service.

Despite the sensitivity of follow-up to the choice of the endline start date, the estimated impacts are uniformly positive and highly statistically significant (Panel E of Tables 4 and 5). The first stage impact on the number of problems solved (Panel D) increases with the use of later endline start dates as does the estimated impact on assessed grade level. In Appendix Table A3 and

A4 we report tests of heterogeneous impacts on households with and without electricity for the subset of households with survey data. The results are mixed with no evidence of heterogeneity in Experiment 2 and larger impacts on households with electricity in Experiment 1. We note that interpretation of heterogeneous impacts is complicated by the fact that estimates derived from application usage data condition on using the platform, as the degree of selection in β_i above could vary across covariate cells.

The fact that many households have multiple children may complicate these results, because the usage data cannot differentiate between different children within the same household. In Appendix Table A5, we report results that are analogous to those in Table 5 for the subset of households that reported having only a single child in the phone survey.

4.2.2 Effect on survey-reported academic outcomes

We examine the effect of treatment on primary school enrollment, school attendance, and grade attainment for children aged six to thirteen, the three academic outcomes included in the survey.[25] Analysis is conducted at the child level. We estimate the following linear model of primary school enrollment and grade reached at endline using OLS:

$$y_{ij} = \alpha + \beta Z_j + X'_{ij}\Gamma + \lambda_j + \varepsilon_i \tag{4}$$

where y_{ij} is the outcome, $Z_j \in \{0, 1\}$ indicates treatment assignment, and λ_j is a common household effect, and ε is an individual child level effect. In all specifications, we control for randomization strata fixed effects, discussed above (Section 3). We report results that pool both experiments as well as separate estimates for each experiment to provide more precise results. Estimates from the pooled specification average two distinct, but closely related experimental programs. The added precision, therefore, comes at a cost of complicating the interpretation of the result, because the counterfactuals are different.

Before discussing the effects on academic outcomes, we report results on the first stage impact

[25] Fourteen-year-olds were included in the survey, but they are excluded from this analysis. The question regarding school enrollment specifically asked whether they attend primary school. We find that fourteen-year-olds are much less likely to be enrolled in primary school. This may reflect dropouts, or it may reflect secondary school transitioning. Due to the wording of the survey item, it is difficult to interpret effects on this group and we exclude fourteen-year-olds from this analysis.

of treatment on measures of engagement with the platform for the sub-samples with survey data. Specifically, we report the impact of access on the number of questions answered, the amount of time that the child's household spent on the service, and an indicator for whether the survey respondent reported that their child had used ElimuLeo in April 2021. In experiment 1, access to ElimuLeo is associated with 49 more questions answered (Column 1 of Table 6), seven more hours (Column 4), and a 16 percentage point increase in the fraction of parents who reported that their child had used ElimuLeo in April 2021 (Column 7). In Experiment 2, reminders increased the number of questions answered by 40 and the number of hours on the platform by 7.5. In Experiment 2, we find no impact on parent-reported use of the platform, presumably because inclusion in that experiment required more active use initially. The pooled estimates (Columns 3 and 6) show lower use among households with electricity. Households with electricity completed 12 fewer problems and spent 1.5 hours less on the platform.

Access to ElimuLeo has no discernible impact on school participation. In the control group, 96.9 percent of primary-school-aged children were attending primary school (Column 3, Table 7). In both Experiment 1 and Experiment 2 we estimate tight zero effects on school enrollment (Columns 1 and 2, Table 7).

Although children are not more likely to be reported as being enrolled in a school, they are more likely to be reported as attending school on the day of the interview. Pooling both experiments, access to use ElimuLeo for households without electricity is associated with a 5.4 percentage point increase in the probability that a child was attending school on the day of the interview from a base of 63.8 percentage points (Column 6 of Table 7). We find no impact on households with electricity. Although the impacts on children in Experiment 1 are not statically significant, the signs of the impacts are similar to those in Experiment 2.

Access to ElimuLeo appears to raise grade attainment as well. Overall, access to ElimuLeo for households without electricity raises grade attainment by 0.11 grades, approximately 57 percent of the gap in grade attainment between households with and without electricity (Column 9 of Table 7). ElimuLeo has no discernible impact on grade attainment of children in households with electricity.

Estimated effects are similar for boys and girls (Appendix Table 11). School closures raised concerns that girls education would be disproportionately harmed (Dreesen et al., 2020), a concern that was based on experience from past public health emergencies in low-income countries (Bandiera

et al., 2020).

4.3 Impact on time use & home-learning activities

This section briefly examines the impact of ElimuLeo on child time use and the mix of home-learning activities. Caregivers were asked two sets of questions about their children's time use. First, they were asked to list the activities of the child in a typical week in April 2021. Emphasis was placed on the fact that this was a period when schools were closed. A second set of questions asked caregivers to report on the number of hours that children had spent, during the same reference period, on a predefined set of activities, including educational activities, child care, chores, tutoring others, and employment. Finally, caregivers were asked to list the educational tools that their child had access to when schools were closed (again in April 2021).

For brevity, we focus discussion on estimates that pool both experiments, although results are also reported separately for each experiment and for households with and without electricity.

Access to ElimuLeo may have increased the fraction of children who were engaged in educational activities when schools were out of session during the April break. Caregivers were 3 percentage points more likely to report that their child had engaged in educational activities, over a base of 78 percent (Column 7, Table 8). There is a corresponding reduction in the fraction of parents who report that their child was engaged in play. Estimates of the effect of ElimuLeo access on the average hours spent on educational activities are positive, but small and statistically insignificant (Column 7, Table 9).

Access to ElimuLeo does not appear to have had any impact on access to other most home-learning tools. We do find an increase of 2.8 percentage points (from a base of 32.1 percent) in reported access to assignments and homework (Table 10). This may reflect practice through ElimuLeo or crowding in of other study resources.

5 Conclusion

This study shows that simple, SMS-based technology can improve access to adaptive math practice and may prevent children from being held back during school closures. Demand for such services is highest among households with indicators of lower socio-economic status, especially households

without access to electricity. These results illustrate the promise of using two-way SMS to build interactive educational tools that reach the growing share of low-income households who possess SMS-enabled mobile phones.

The large estimated impacts of providing access on assessed math competencies and grade attainment suggests that programs like ElimuLeo can be effective at preventing retention in grade. These impacts are especially pronounced for households that lack electricity, a group that is far less likely to be enrolled, attend school regularly, and progress normally through grades.

These results may indicate that children in smallholder farming households would benefit from more opportunities to practice math skills and get feedback. A large literature in cognitive science finds that practice itself is an important pedagogical tool (Karpicke and Roediger, 2007, 2008; Karpicke and Grimaldi, 2012; Nunes and Karpicke, 2015, e.g.). Field evaluations have also found that adaptive and non-adaptive math practice can improve retention of academic skills, especially for lower-achieving students (Muralidharan, Singh and Ganimian, 2019; de Barros, Ganimian and Venkatachalam, 2022). If, as indicated by parent reports of child time use, children who used the platform would have otherwise spent time playing, then the positive impacts on academic outcomes could reflect the higher academic value of practice compared to the displaced leisure activities.

However, the impact could also reflect the influence of the ElimuLeo service on other educational inputs. The impact of any educational intervention is the combined impact of all household and school resources a child receives (Todd and Wolpin, 2007). First, although ElimuLeo did not explicitly aim to affect parental engagement, the intervention involved substantial engagement with parents, who were in most cases the owners of the phones.[26] A large literature has found that SMS-based interventions with parents can themselves result in improvements in child educational outcomes (Bursztyn and Coffman, 2017; Berlinski et al., 2016; Bergman and Chan, 2019; York, Loeb and Doss, 2019; Bettinger et al., 2020). It is possible that frequent SMS messages to households to remind them to use the service led to improvements in education through parental behaviors. Treated parents may have exerted more pressure for their children to re-engage with their schools or may have taken a more active role in the decision for their children to progress in school. Parents may also have taken a more active role in their child's home study. Children may also

[26] Qualitative interviews with users identified a range of different use profiles - including both parents who gave the phone to children to use ElimuLeo independently and parents who copied down the problems onto paper for children to complete.

have adjusted their behavior in response to the ElimuLeo access. We find some evidence that the platform may have crowded in the use of other distance learning resources, such as assignment books and textbooks. It's possible that children used the program for its diagnostic output and were then able to use other educational resources more efficiently. Further research is needed to understand how this tool is used in conjunction with other resources that children have.

These results may be especially informative about the usefulness of two-way SMS messaging to support resilient education systems that can weather periods of unplanned closures. These results were observed in the context of an unprecedented shock to Kenya's educational system.

Extrapolation to other settings like planned closures for vacation (e.g., to address summer learning loss) requires caution. We do find substantial demand for the service during a planned closure in Experiment 2 after schools are closed. However, the aftermath of the COVID-19 school closures may also be an exceptional moment. We find that between the first and second experiments, demand for the service (as measured by opt-ins) falls significantly.

These results do not mean that services like ElimuLeo can eliminate gaps in access to digital learning altogether. Households with mobile phones may still face barriers to accessing ElimuLeo such as reliable electricity access (inside or outside the home), or mobile phones may not be accessible to children. In qualitative interviews, several respondents confirmed that keeping phones charged sometimes prevented them from using ElimuLeo. Finally, our study focuses on existing clients of PxD. As discussed above, mobile phones are widespread, but a substantial share of households still lack phones altogether. Future work would be needed to more accurately quantify the distributional impacts of SMS-based practice.

References

Angrist, Noam, David K. Evans, Deon Filmer, Rachel Glennerster, Halsey F. Rogers, and Shwetlena Sabarwal. 2020*a*. "How to Improve Education Outcomes Most Efficiently? A Comparison of 150 Interventions Using the New Learning-Adjusted Years of Schooling Metric." World Bank Working Paper.

Angrist, Noam, Peter Bergman, Caton Brewster, and Moitshepi Matsheng. 2020*b*. "Stemming Learning Loss During the Pandemic: A Rapid Randomized Trial of Low-Tech Intervention in Botswana." *Working paper*. Link.

Bandiera, Oriana, Niklas Buerhen, Markus Goldstein, Imran Rasul, and Andrea Smurra. 2020. "Do School Closures During an Epidemic have Persistent Effects? Evidence from Sierra Leone in the time of ebola." Working Paper.

Banerjee, Abhijit, Shawn Cole, Esther Duflo, and Leigh Linden. 2007. "Remedying education: Evidence from two randomized experiments in India." *The Quarterly Journal of Economics*, 122.

Banerjee, Abhijit V., and Esther Duflo. 2014. "(Dis)organization and Success in an Economics MOOC." *American Economic Review*, 104(5): 514–18.

Bergman, Peter, and Eric W. Chan. 2019. "Leveraging Parents through Low-Cost Technology: The Impact of High-Frequency Information on Student Achievement." *Working Paper*. Link.

Berlinski, Samuel, Matias Busso, Taryn Dinkelman, and Claudia Martinez. 2016. "Reducing parent-school information gaps and improving education outcomes: Evidence from high frequency text messaging in Chile." *Working Paper*. Link.

Bettinger, Eric, Nina Cunha, Gulherme Lichand, and Ricardo Madeira. 2020. "Are the Effects of Informational Interventions Driven by Salience?" *Working paper*. Link.

Bursztyn, Leonardo, and Lucas C. Coffman. 2017. "The Schooling Decision: Family Preferences, Intergenerational Conflict, and Moral Hazard in the Brazilian *faveleas*." *American Economic Review*, 107(6). Link.

Carlana, Michela, and Eliana La Ferrara. 2021. "Apart but Connected: Online tutoring and student outcomes during the COVID-19 pandemic." IZA Discussion Papers, No. 14094, Institute of Labor Economics (IZA), Bonn.

Dahir, Abdi Latif. 2020. "Kenya's Unusual Solution to the School Problem: Cancel the Year and Start Over." *The New York Times*.

de Barros, Andreas, Alejandro J. Ganimian, and Anuja Venkatachalam. 2022. "Which Students Benefit from Independent Practice? Experimental Evidence from a Math Software in Private Schools in India." *Journal of Research on Educational Effectiveness*, 15(2): 279–301.

Dreesen, Thomas, Spogmai Akseer, Matt Brossard, Pragya Dewan, Juan-Pablo Giraldo, Akito Kamei, Suguru Mizunoya, and Javier Santiago Ortiz. 2020. "Promising practices for equitable remote learning: Emerging lessons from COVID-19 education responses in 127 countries."

Evans, David, and Fei Yuan. 2019. "Equivalent years of schooling: A metric to communicate learning gains in concrete terms." *World Bank Policy Research Working Paper*.

Evlidge, Christopher D., Kimberly E. Baugh, Mikhail Zhizhin, and Feng Chi Hsu. 2013. "Why VIIRS data are superior to DMSP for mapping nighttime lights." *Asia-Pacific Advanced Network*, 35: 62.

Evlidge, Christopher D., Kimberly E. Baugh, Mikhail Zhizhin, and Feng Chi Hsu. 2017. "VIIRS night-time lights." *International Journal of Remote Sensing*, 38(21): 586–=5879.

GSMA. 2017. "The Mobile Economy: Sub-Saharan Africa 2017." GSM Association. Link.

Karpicke, Jeffrey D., and Henry L. Roediger. 2007. "Expanding retrieval practice promotes short-term retention, but equally spaced retrieval enhances long-term retention." *Journal of experimental psychology. Learning, memory, and cognition*, 33 4: 704–19.

Karpicke, Jeffrey D., and Henry L. Roediger. 2008. "The Critical Importance of Retrieval for Learning." *Science*, 319(5865): 966–968.

Karpicke, Jeffrey D., and Phillip J. Grimaldi. 2012. "Retrieval-Based Learning: A Perspective for Enhancing Meaningful Learning." *Educational Psychology Review*, 24: 401–418. Link.

Kenya National Bureau of Statistics. 2016. "2015-2016 Kenya Integrated Household Budget Survey."

KICD. 2002. "Primary Education Syllabus: Volume Two."

KICD. 2017. "Lower Primary Level Curriculum Designs: Volume Two."

KICD. 2019. "Upper Primary Level Designs: Volume Three, Grade 4."

KNBS. 2019. "Kenya Continuous Household Survey Programme (KCHSP) - 2019." *https://statistics.knbs.or.ke/nada/index.php/catalog/1/study-description*.

KNEC. 2020*a*. "Monitoring Learner Achievement at Class 3 Level of Primary School Education in Kenya." Kenya National Examinations Council – National Assessment Centre. Link.

KNEC. 2020*b*. "Monitoring Learner Achievement at Form 2 Level of Secondary Education in Kenya: A Baseline Report." Kenya National Examinations Council – National Assessment Centre. Link.

Muralidharan, Karthik, Abhijeet Singh, and Alejandro J Ganimian. 2019. "Disrupting education? Experimental evidence on technology-aided instruction in India." *American Economic Review*, 109(4): 1426–60.

Muralidharan, Karthik, and Abhijeet Singh. 2021. "Evaluation of a technology-aided instruction platform (Mindspark) in government schools in Rajasthan." https://doi.org/10.1257/rct.2546-1.3.

Ngware, Moses, and Vollan Ochieng. 2020. "EdTech and the COVID-19 response: A case study of Kenya." EdTech Hub Working Papers, November. Link.

Nunes, Ludmila D., and Jeffrey D. Karpicke. 2015. "Retrieval-Based Learning: Research at the Interface between Cognitive Science and Education." *Emerging Trends in the Social and Behavioral Sciences*. Link.

Oza, Shardul, and Jacobus Cilliers. 2021. "What Did Children Do During School Closures." Research on Improving Systems of Education (RISE) Report.

Pape, Utz Johann, Antonia Delius, Ritika Khandelwal, and Rhea Gupta. 2020. "Socioeconomic Impacts of COVID-19 in Kenya." https://microdata.worldbank.org/index.php/catalog/3774/download/56111.

Pew Research Center. 2018. "Majorities in sub-Saharan Africa own mobile phones, but smartphone adoption is modest." Pew Research Center Report.

Schiavina, Marcello, Sergio Freire, and Kytt MacManus. 2022. "GHS-POP R2022A - GHS population grid multitemporal (1975-2030)." *http://data.europa.eu/89h/d6d86a90-4351-4508-99c1-cb074b022c4a*.

Todd, Petra E., and Kenneth I. Wolpin. 2007. "The Production of Cognitive Achievement in CHildren: Home, School, and Racial Test Score Gaps." *Journal of Human Resources*, 1(1).

UNICEF. 2020. "Unequal access to remote schooling amid COVID-19 threatens to deepen global learning crisis."

Uwezo. 2016. "Are Our Children Learning? Uwezo Kenya Sixth Learning Assessment Report." Nairobi: Twaweza East Africa, Nairobi, Kenya. Link.

Wanzala, Ouma. 2018. "Students made to repeat classes in spite of ban." *Nation*. Link.

Wanzala, Ouma. 2019. "Students repeating classes despite ban, dropout rate." *Nation*. Link.

Wolf, Sharon, Elisabetta Aurino, Noelle Suntheimer, Esinam Avornyo, and Edward Tsinigo. 2021. "Learning in the Time of a Pandemic and Implications for Returning to School: Effects of COVID-19 in Ghana." Consortium for Policy Research in Education (CPRE) Working Paper.

World Bank Group. 2019. "Kenya Economic Update: Securing Future Growth: Policies to Support Kenya's Digital Transformation." World Bank Report, October.

World Bank, University of California Berkeley, Kenya National Bureau of Statistics. 2021. "Kenya - COVID-19 Rapid Response Phone Survey with Households 2020-2022, Panel, Waves1-8 (COVIDRS)." *https://microdata.worldbank.org/index.php/catalog/3774/study-description#metadata-data_access*.

York, Benjamin N., Susanna Loeb, and Christopher Doss. 2019. "One Step at a Time: The Effects of an Early Literacy Text Messaging Program for Parents of Preschoolers." *Journal of Human Resources*, 54(3). Link.

Figures

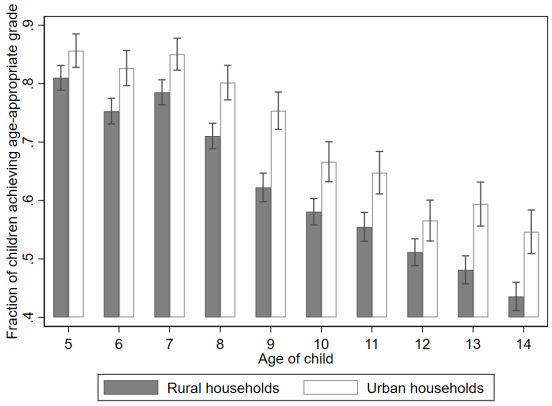

Figure 1: Authors calculations based on the 2019 Kenya Continuous Household Survey. Bars indicate the fraction of children who are in a grade that would be appropriate for their age if a student were to enter primary school at age seven and progress one grade per year. Ranges indicate the 95 percent confidence interval.

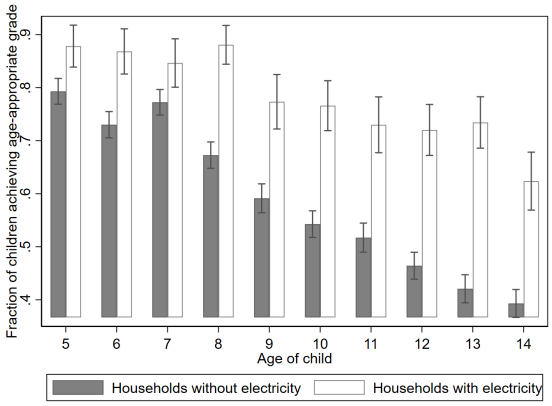

Figure 2: Authors calculations based on the 2019 Kenya Continuous Household Survey. Bars indicate the fraction of children who are in a grade that would be appropriate for their age if a student were to enter primary school at age seven and progress one grade per year. Ranges indicate the 95 percent confidence interval.

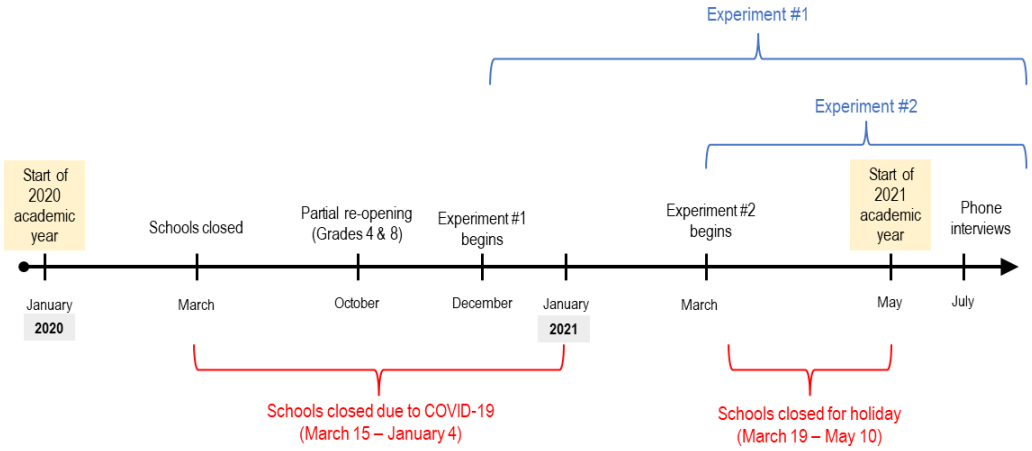

Figure 3: Timeline of events. The 2020 academic year began in January. Schools closed due to the COVID-19 pandemic on March 15, 2020 and reopened on January 4, 2021. Schools were partially reopened in October for students in grades 4 and 8 to help those students prepare for end-of-grade exams. Due to closure, the 2020 academic year extended until March 2021. The 2021 academic year began in May of 2021.

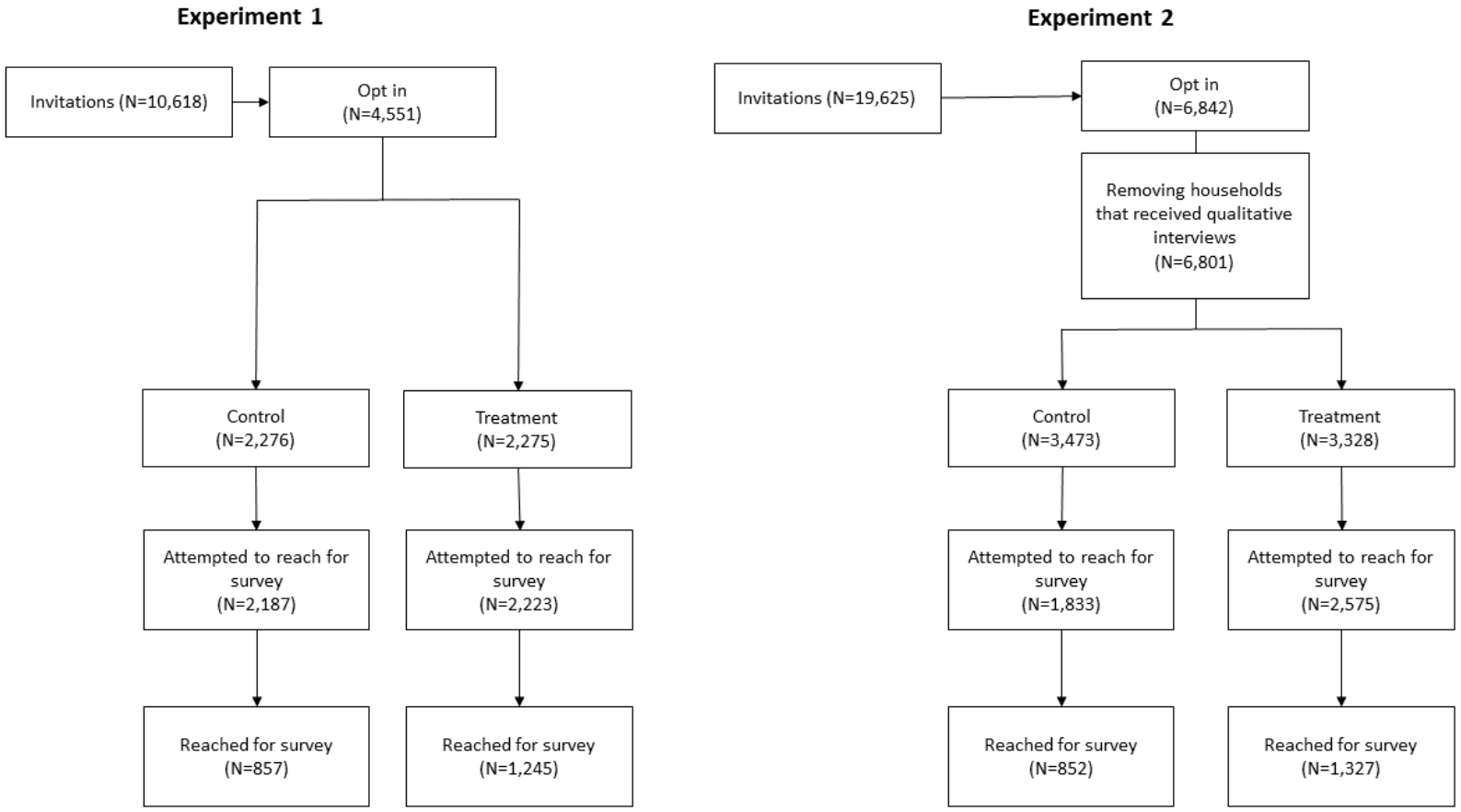

Figure 4: This figure illustrates the treatment assignment and the construction of the final sample.

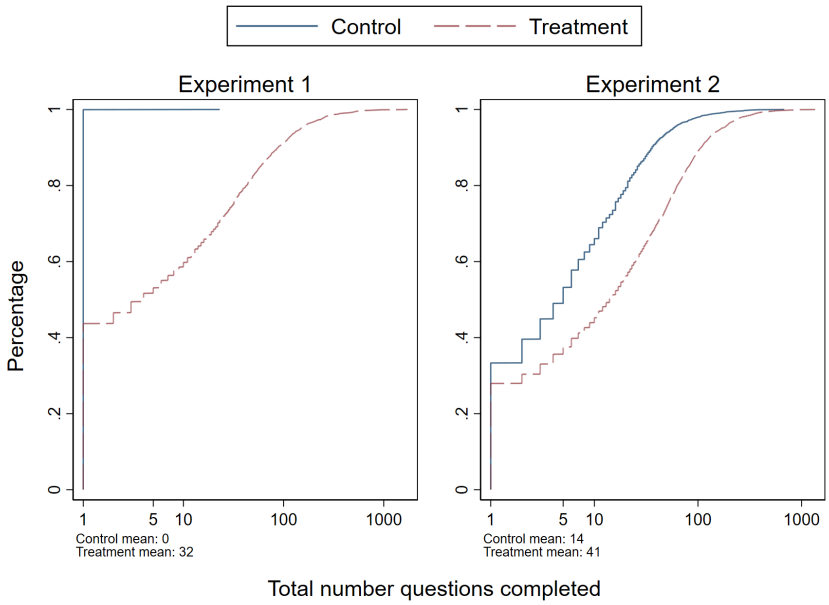

Figure 5: Cumulative distribution of total number of questions completed over the course of the study.

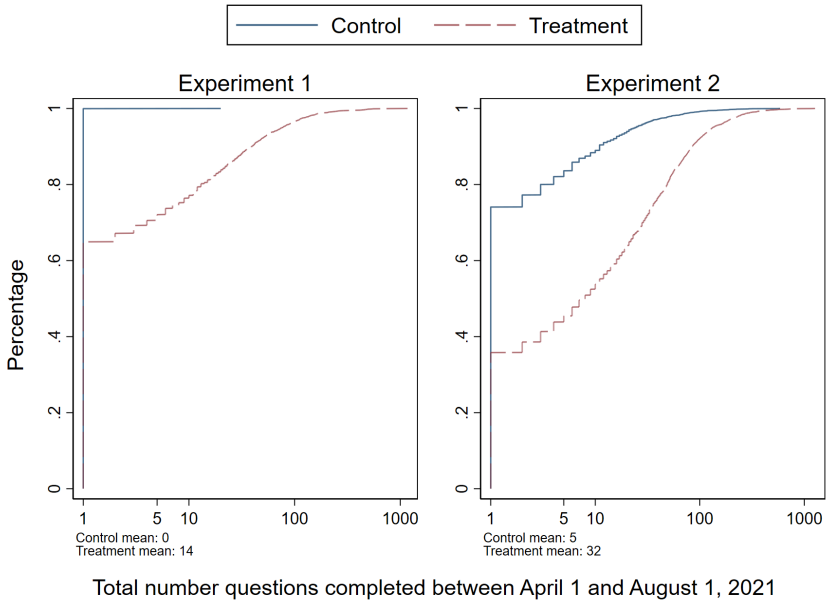

Figure 6: Cumulative distribution of total number of questions completed between April 1 and August 1, 2021.

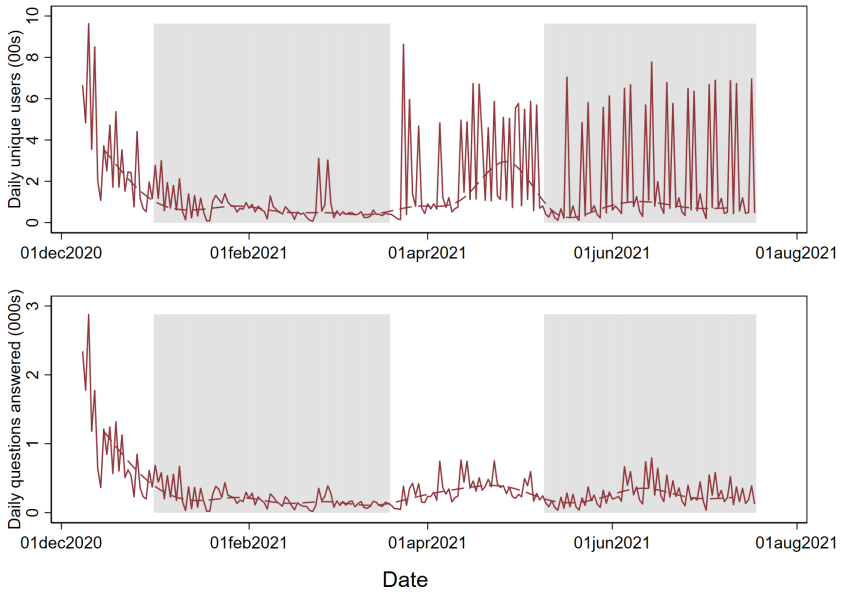

Figure 7: Usage over time for **Experiment 1**. The shaded region indicates periods where school was in session.

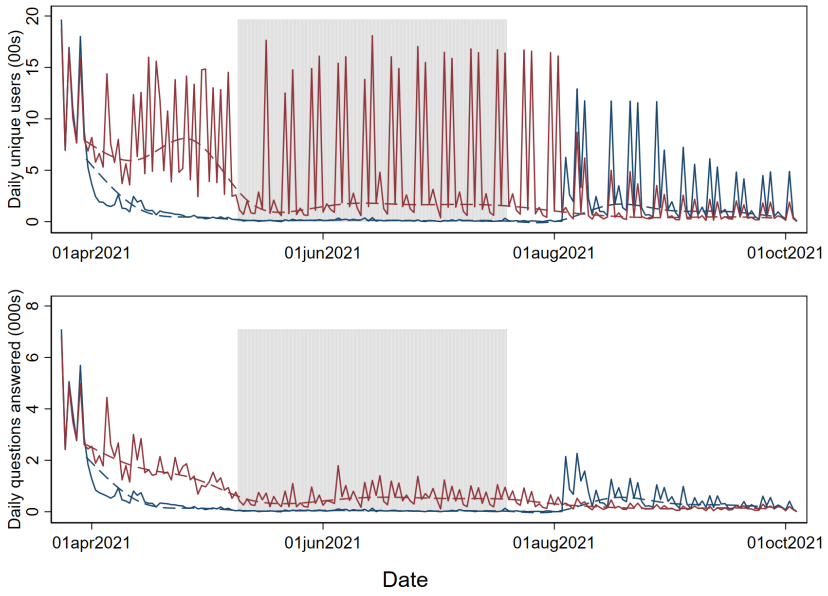

Figure 8: Usage over time for **Experiment 2**. The shaded region indicates periods where school was in session. The control series indicates usage for those households that did not receive reminders to usage the service. The treatment series indicates usage for those households that did receive reminders.

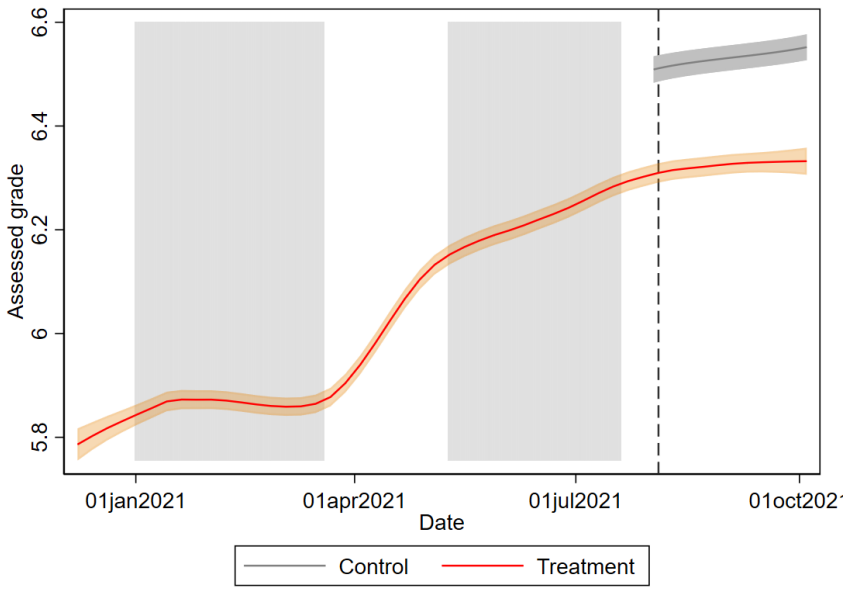

Figure 9: Kernel weighted local polynomial regression of the daily maximum assessed grade level of **Experiment 1** households who used ElimuLeo. The grey vertical bars indicates periods when school was in session. The vertical dashed line indicates when control households received access to ElimuLeo.

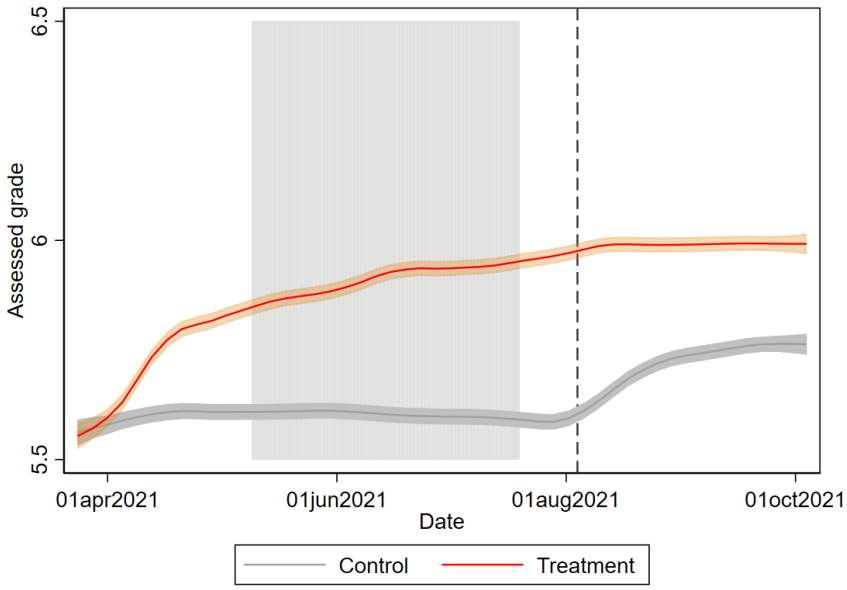

Figure 10: Kernel weighted local polynomial regression of the daily maximum assessed grade level of **Experiment 2** households who used ElimuLeo. The grey vertical bar indicates periods when school was in session. The vertical dashed line indicates when control households received access to ElimuLeo.

Table 1: Follow-up and covariate balance

	Experiment 1		Experiment 2		Pooled	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Follow-up						
Interviewed	0.53	0.075***	0.49	0.032**	0.51	0.054***
		(0.015)		(0.016)		(0.011)
N		4,410		4,408		8,818
Any children (conditional on interview)	0.84	0.040***	0.90	0.035***	0.87	0.038***
		(0.014)		(0.013)		(0.010)
N		2,501		2,246		4,747
Panel B: Covariate balance						
Household characteristics						
Household has electricity	0.37	-0.013	0.38	-0.023	0.37	-0.018
		(0.020)		(0.022)		(0.015)
Household owns radio	0.84	-0.028*	0.82	-0.020	0.83	-0.024**
		(0.016)		(0.018)		(0.012)
Household owns TV	0.47	-0.016	0.49	-0.031	0.48	-0.023
		(0.021)		(0.023)		(0.016)
Number of children	1.82	0.034	1.90	0.040	1.90	0.037
		(0.048)		(0.052)		(0.035)
Interviewed in Swahili	0.84	-0.013	0.85	0.008	0.85	-0.003
		(0.016)		(0.016)		(0.011)
Number of Paddy inquiries	11.44	-0.587				
		(0.392)				
Initial grade level			4.78	-0.053		
				(0.115)		
Child characteristics						
Female child	0.50	0.001	0.51	0.002	0.50	0.002
		(0.015)		(0.016)		(0.011)
Age of child	10.11	-0.075	10.00	0.037	10.14	-0.023
		(0.070)		(0.074)		(0.051)
N		4,707		4,436		9,143

Notes: Panel A shows the effect of treatment on endline follow-up, from the sub-sample of households that were attempted during the endline survey. Panel B reports tests of balance on characteristics of the household conditional on being interviewed and reporting having at least one child. Each test comes from OLS regression of the characteristics of the household on treatment. *Number Paddy inquiries* is a count of the number of times the household's phone number accessed PxD's two-way SMS-based agricultural extension services and proxies for the users degree of engagement on that platform. *Assessed grade level* refers to the household's initially reported grade level on the Elimuleo platform. In Experiment 2, this information was collected prior to randomization. It is not reported for Experiment 1 because the responses were collected many months apart and are likely endogenous to treatment status. All covariates reported in Panel B are come from the survey except the number of Paddy inquiries and the assessed grade level. Robust standard errors are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

Table 2: Predictors of the decision to opt in to use Elimuleo

Estimate	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ln(ward night lights)	-0.015***	-0.024***					-0.018***
	(0.006)	(0.007)					(0.007)
ln(population/km^2)			0.001				0.001
			(0.003)				(0.004)
County net primary-school enrollment				-0.044***			-0.060**
				(0.015)			(0.027)
County fraction children in age-appropriate grade					-0.140***		-0.133***
					(0.027)		(0.029)
Arid county						-0.054***	0.014
						(0.018)	(0.033)
Semi-arid county						-0.006	0.000
						(0.007)	(0.008)
Experiment 2	-0.049***	-0.045***	-0.049***	-0.048***	-0.047***	-0.048***	-0.046***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Observations	30,243	30,243	30,243	30,220	30,220	30,243	30,220
County fixed effects	No	Yes	No	No	No	No	No

Notes: Nightlights are the mean value at the ward level from the Visible and Infrared Imaging Suite (VIIRS) Nighttime lights maps (Evlidge et al., 2013, 2017). Robust standard errors reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

Table 3: Relationship between the total number of questions answered and household characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Number of children (aged 4-14)	0.40**						0.37**
	(0.18)						(0.18)
Number girls (aged 4-14)		0.32					
		(0.21)					
Number boys (aged 4-14)			0.23				
			(0.21)				
Household has electricity				-1.18***			-0.78*
				(0.38)			(0.42)
Household owns radio					-1.30**		-1.03*
					(0.53)		(0.54)
Household owns TV						-1.16***	-0.70*
						(0.37)	(0.40)
Experiment 2	1.59***	1.60***	1.60***	1.57***	1.58***	1.60***	1.55***
	(0.39)	(0.39)	(0.39)	(0.39)	(0.39)	(0.39)	(0.39)
Constant	3.19***	3.71***	3.81***	4.50***	5.13***	4.60***	4.73***
	(0.44)	(0.36)	(0.33)	(0.34)	(0.55)	(0.35)	(0.68)
Observations	2445	2445	2445	2445	2445	2445	2445
R^2	0.026	0.025	0.024	0.027	0.027	0.027	0.032

Notes: Table reports the relationship between the total number of questions answered on the ElimuLeo platform and survey reported household characteristics. All estimates come from OLS regression. The sample is restricted to households in the treatment group who had access to ElimuLeo prior the survey.

Table 4: Effects on maximum assessed math grade level
Experiment 1

Outcome	(01apr) (1)	(01jun) (2)	(15jun) (3)	(01jul) (4)	(01aug) (5)
Panel A: Effect of treatment on follow-up rate					
Coefficient	0.117***	0.022*	-0.018	-0.067***	-0.206***
SE	(0.014)	(0.014)	(0.013)	(0.013)	(0.011)
Control mean	0.388	0.323	0.316	0.313	0.308
Observations	4,472	4,472	4,472	4,472	4,472
Panel B: Effect of treatment on last assessment date					
Coefficient	-64.0***	-44.6***	-38.4***	-31.0***	-1.2
SE	(1.7)	(1.5)	(1.5)	(1.5)	(2.0)
Observations	1,512	1,296	1,205	1,092	777
Panel C: Effect of treatment on highest assessed grade at baseline					
Coefficient	-0.259***	-0.249***	-0.299***	-0.319***	-0.355***
SE	(0.084)	(0.089)	(0.093)	(0.099)	(0.146)
Control mean	5.373	5.370	5.378	5.379	5.396
Observations	1,431	1,228	1,145	1,037	734
Panel D: Number questions answered					
Coefficient	48.6***	61.5***	66.3***	76.7***	145.9***
SE	(5.4)	(6.7)	(7.3)	(8.7)	(21.8)
Control mean	37.9	41.1	41.2	41.5	41.8
Observations	1,511	1,295	1,204	1,091	776
Panel E: Effect of treatment on highest assessed grade level					
Coefficient	0.135**	0.216***	0.235***	0.267***	0.588***
SE	(0.063)	(0.066)	(0.068)	(0.072)	(0.094)
Control mean	6.401	6.460	6.466	6.471	6.495
Observations	1,512	1,296	1,205	1,092	777

Notes: Table shows the effect of receiving access to the Elimuleo platform on the highest assessed grade level for *Experiment 1* households that used the platform. Each observation represents a household. Use of the platform at different points in time varies across treatment and control so that the latest date of assessment is potentially endogenous to treatment. Column 1 restricts the analysis to those applicants who last used the platform after April 1, 2021. Coumns 2-5 test the robustness of the results to changing this cut-off to restrict the analysis to those households that used the platform more recently. Panel A reports the effect of treatment on whether a household has an assessed grade recent enough to satisfy these inclusion criteria. Panel B reports the effect on the date from which the student's assessed level comes from. Panel C reports a comparison of baseline assessed levels in treatment and control for those satisfying the inclusion criteria (balance). Panel D reports the difference in the number of questions answered by October 1, 2021. Panel E reports the estimated effect on the highest assessed grade level, controlling for the baseline level. Robust standard errors reported in parentheses. All specifications include a linear control for the probability of treatment for the household's randomization strata, baseline reported grade level, and subject specific baseline assessed grade. To accommodate missingness in the subject-specific assessed grade level, an indicator is included in the controls for whether the household was missing a particular subject-specific assessment, and missing baseline assessed level is imputed to the number four. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

Table 5: Effects on maximum assessed math grade level
Experiment 2

Outcome	(01apr) (1)	(01jun) (2)	(15jun) (3)	(01jul) (4)	(01aug) (5)
Panel A: Effect of treatment on follow-up rate					
Coefficient	0.194***	0.062***	0.018	-0.042***	-0.226***
SE	(0.012)	(0.012)	(0.012)	(0.012)	(0.010)
Control mean	0.388	0.323	0.316	0.313	0.308
Observations	6,801	6,801	6,801	6,801	6,801
Panel B: Effect of treatment on last assessment date					
Coefficient	-37.7***	-36.5***	-32.8***	-27.1***	-2.3
SE	(1.9)	(1.2)	(1.1)	(1.1)	(1.5)
Observations	4,029	2,799	2,573	2,330	1,650
Panel C: Effect of treatment on highest assessed grade at baseline					
Coefficient	0.049	0.098*	0.069	0.044	0.008
SE	(0.048)	(0.057)	(0.059)	(0.064)	(0.096)
Control mean	5.373	5.370	5.378	5.379	5.396
Observations	3,767	2,652	2,445	2,219	1,576
Panel D: Number questions answered					
Coefficient	22.5***	37.5***	42.5***	51.5***	100.4***
SE	(3.0)	(4.2)	(4.6)	(5.3)	(10.9)
Control mean	37.9	41.1	41.2	41.5	41.8
Observations	4,029	2,799	2,573	2,330	1,650
Panel E: Effect of treatment on highest assessed grade level					
Coefficient	0.360***	0.450***	0.466***	0.500***	0.646***
SE	(0.041)	(0.049)	(0.051)	(0.054)	(0.078)
Control mean	6.401	6.460	6.466	6.471	6.495
Observations	4,029	2,799	2,573	2,330	1,650

Notes: Table shows the effect of receiving access to the Elimuleo platform on the highest assessed grade level for *Experiment 2* households that used the platform. Each observation represents a household. Use of the platform at different points in time varies across treatment and control so that the latest date of assessment is potentially endogenous to treatment. Column 1 restricts the analysis to those applicants who last used the platform after April 1, 2021. Coumns 2-5 test the robustness of the results to changing this cut-off to restrict the analysis to those households that used the platform more recently. Panel A reports the effect of treatment on whether a household has an assessed grade recent enough to satisfy these inclusion criteria. Panel B reports the effect on the date from which the student's assessed level comes from. Panel C reports a comparison of baseline assessed levels in treatment and control for those satisfying the inclusion criteria (balance). Panel D reports the difference in the number of questions answered by October 1, 2021. Panel E reports the estimated effect on the highest assessed grade level, controlling for the baseline level. Robust standard errors reported in parentheses. All specifications include a linear control for the probability of treatment for the household's randomization strata, baseline reported grade level, and subject specific baseline assessed grade. To accommodate missingness in the subject-specific assessed grade level, an indicator is included in the controls for whether the household was missing a particular subject-specific assessment, and missing baseline assessed level is imputed to the number four. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

Table 6: Effect of access to Elimuleo on usage
(Conditional on survey follow-up)

	Questions answered (/10)			Time active on the platform (hours)			Survey-reported use of the platform		
Estimate	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Elimuleo offer	4.871***	4.028***	4.479***	7.243***	7.464***	7.334***	0.176***	0.021	0.108***
	(0.502)	(0.399)	(0.330)	(0.377)	(0.290)	(0.246)	(0.017)	(0.022)	(0.013)
Household has electricity	0.080	0.758	0.385*	-0.001	0.287	0.122	0.013	-0.017	0.004
	(0.133)	(0.520)	(0.232)	(0.114)	(0.228)	(0.116)	(0.018)	(0.028)	(0.016)
Elimuleo × electricity	-1.144*	-1.350	-1.204**	-0.820	-1.438***	-1.101***	-0.033	0.086**	0.020
	(0.695)	(0.925)	(0.559)	(0.605)	(0.476)	(0.388)	(0.029)	(0.038)	(0.023)
Control mean	0.002	1.764	0.777	0.002	1.001	0.441	0.096	0.246	0.162
Observations	3,997	3,763	7,760	3,997	3,763	7,760	3,997	3,763	7,760
Included samples									
Experiment 1	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes
Experiment 2	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes

Notes: Table shows the effect of receiving access to the Elimuleo platform on usage for the sample of households included in the analyses using survey reported outcomes. Each observation represents a child in a household. The analysis restricts to children aged six to thirteen. Of the usage outcomes, onl survey-reported use varies within households. Standard errors are clustered at the household level and are reported in parentheses. All specifications control for randomization strata. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

Table 7: Effects on parent reported academic outcomes

	Enrolled in school			Attending school on day of interview			Grade attainment		
Estimate	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Elimuleo offer	0.005	-0.010	-0.002	0.047*	0.069***	0.057***	0.076	0.160**	0.113**
	(0.008)	(0.008)	(0.006)	(0.025)	(0.029)	(0.019)	(0.065)	(0.078)	(0.050)
Household has electricity	0.015*	0.002	0.010	0.039	0.118***	0.077***	0.201***	0.204**	0.201***
	(0.009)	(0.009)	(0.007)	(0.035)	(0.036)	(0.025)	(0.086)	(0.093)	(0.063)
Elimuleo × electricity	-0.026**	0.011	-0.009	-0.032	-0.126***	-0.077***	-0.134	-0.116	-0.124
	(0.012)	(0.012)	(0.009)	(0.043)	(0.046)	(0.032)	(0.110)	(0.121)	(0.082)
Control mean	0.966	0.974	0.969	0.652	0.625	0.638	5.254	5.143	5.205
Observations	3,997	3,763	7,760	3,436	3,662	7,098	3,853	3,649	7,502
Included samples									
Experiment 1	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes
Experiment 2	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes

Notes: Table shows the effect of receiving access to the Elimuleo platform on caregiver reported outcomes. Each observation represents a child in a household. The analysis restricts to children aged six to thirteen. Standard errors are clustered at the household level and are reported in parentheses. All specifications control for randomization strata. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

Table 8: Effects on child time use (extensive margin)

Outcome	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Educational activites									
Coefficient	0.042***	0.042**	0.038	0.015	-0.014	0.049	0.030***	0.019	0.044**
SE	(0.016)	(0.020)	(0.027)	(0.017)	(0.021)	(0.030)	(0.012)	(0.015)	(0.020)
Control mean	0.765	0.759	0.776	0.802	0.803	0.800	0.781	0.778	0.787
Play									
Coefficient	-0.010	-0.023	0.019	-0.059***	-0.072***	-0.004	-0.031**	-0.043***	0.010
SE	(0.018)	(0.023)	(0.031)	(0.022)	(0.028)	(0.038)	(0.014)	(0.018)	(0.024)
Control mean	0.620	0.619	0.623	0.623	0.624	0.619	0.621	0.621	0.621
Watching television									
Coefficient	-0.005	-0.007	-0.004	-0.014	-0.029	0.011	-0.009	-0.016	0.002
SE	(0.014)	(0.015)	(0.030)	(0.016)	(0.018)	(0.033)	(0.011)	(0.012)	(0.022)
Control mean	0.136	0.087	0.238	0.131	0.098	0.194	0.134	0.092	0.218
Chores									
Coefficient	0.024	0.016	0.042	-0.028	-0.029	-0.042	0.001	-0.004	0.007
SE	(0.017)	(0.021)	(0.031)	(0.020)	(0.026)	(0.035)	(0.013)	(0.016)	(0.023)
Control mean	0.619	0.635	0.584	0.656	0.658	0.652	0.635	0.645	0.615
Visiting friends/relatives									
Coefficient	0.009	0.010	0.009	-0.005	-0.004	-0.008	0.003	0.004	0.001
SE	(0.007)	(0.009)	(0.012)	(0.007)	(0.009)	(0.015)	(0.005)	(0.006)	(0.009)
Control mean	0.034	0.031	0.039	0.039	0.033	0.051	0.036	0.032	0.044
Child care									
Coefficient	-0.003	-0.006	0.005	-0.004	0.007	-0.032*	-0.004	-0.001	-0.011
SE	(0.009)	(0.012)	(0.014)	(0.010)	(0.013)	(0.019)	(0.007)	(0.009)	(0.011)
Control mean	0.060	0.068	0.042	0.069	0.060	0.085	0.064	0.065	0.062
Tutoring others									
Coefficient	0.007	0.002	0.014	-0.003	-0.005	-0.003	0.002	-0.001	0.007
SE	(0.009)	(0.011)	(0.015)	(0.011)	(0.015)	(0.022)	(0.007)	(0.009)	(0.013)
Control mean	0.050	0.055	0.041	0.071	0.066	0.081	0.059	0.060	0.059
Religious activities									
Coefficient	0.021***	0.023***	0.018	0.012	0.003	0.021	0.017***	0.015*	0.019**
SE	(0.007)	(0.010)	(0.011)	(0.009)	(0.012)	(0.016)	(0.006)	(0.008)	(0.010)
Control mean	0.031	0.034	0.024	0.041	0.048	0.028	0.036	0.040	0.026
Earning income									
Coefficient	0.004	0.002	0.011	0.004	0.010	-0.020	0.004	0.005	-0.002
SE	(0.010)	(0.013)	(0.017)	(0.012)	(0.016)	(0.023)	(0.008)	(0.010)	(0.014)
Control mean	0.058	0.061	0.053	0.065	0.058	0.079	0.062	0.060	0.065
Observations	3,999	2,666	1,331	3,763	2,488	1,247	7,762	5,154	2,578
Experiment 1	Yes	Yes	Yes	No	No	No	Yes	Yes	Yes
Experiment 2	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Households w/o electricity	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
Households w/ electricity	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes

Notes: Table shows the effect of receiving access to the Elimuleo platform on caregiver reported outcomes. Each observation represents a child in a household. The analysis restricts to children aged six to thirteen. Standard errors are clustered at the household level and are reported in parentheses. All specificatons control the age of the child and the randomization stratum. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

Table 9: Effects on child time use

Outcome		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Educational activities										
	Coefficient	0.383	-0.076	1.323***	0.009	0.141	0.122	0.225	0.018	0.829**
	SE	(0.312)	(0.390)	(0.524)	(0.361)	(0.452)	(0.658)	(0.237)	(0.295)	(0.412)
	Control mean	12.642	12.594	12.717	12.884	12.489	13.628	12.748	12.549	13.132
Child care										
	Coefficient	-0.042	-0.110	0.240	-0.052	0.580	-1.536***	-0.043	0.179	-0.516
	SE	(0.233)	(0.285)	(0.392)	(0.289)	(0.362)	(0.518)	(0.182)	(0.225)	(0.317)
	Control mean	3.312	3.413	3.104	3.526	3.082	4.361	3.406	3.270	3.677
Chores										
	Coefficient	-0.005	-0.131	0.411	-0.517	-0.397	-1.099*	-0.226	-0.243	-0.225
	SE	(0.320)	(0.401)	(0.526)	(0.356)	(0.468)	(0.595)	(0.238)	(0.305)	(0.395)
	Control mean	9.754	9.954	9.354	9.822	9.909	9.659	9.784	9.935	9.493
Tutoring others										
	Coefficient	0.183	0.023	0.588*	-0.216	0.020	-0.692*	0.010	0.026	0.037
	SE	(0.178)	(0.219)	(0.305)	(0.213)	(0.272)	(0.381)	(0.137)	(0.171)	(0.241)
	Control mean	3.052	3.170	2.812	3.273	3.199	3.412	3.149	3.183	3.086
Employment										
	Coefficient	-0.032	-0.242	0.457	0.245	0.498	-0.227	0.091	0.070	0.170
	SE	(0.220)	(0.278)	(0.360)	(0.255)	(0.330)	(0.453)	(0.166)	(0.213)	(0.284)
	Control mean	2.269	2.486	1.816	2.337	2.376	2.263	2.299	2.439	2.020
Observations		3,999	2,666	1,331	3,763	2,488	1,247	7,762	5,154	2,578
Experiment 1		Yes	Yes	Yes	No	No	No	Yes	Yes	Yes
Experiment 2		No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Households w/o electricity		Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
Households w/ electricity		Yes	No	Yes	Yes	No	Yes	Yes	No	Yes

Notes: Table shows the effect of receiving access to the Elimuleo platform on caregiver reported outcomes. Each observation represents a child in a household. The analysis restricts to children aged six to thirteen. Standard errors are clustered at the household level and are reported in parentheses. All specificatons control for the age of the child and the randomization stratum. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

Table 10: Effects on use of home learning activities when schools were closed

Activity		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Radio										
	Coefficient	-0.013	-0.017	-0.001	-0.000	0.010	-0.025	-0.007	-0.005	-0.012
	SE	(0.009)	(0.011)	(0.017)	(0.011)	(0.014)	(0.020)	(0.007)	(0.009)	(0.013)
	Control mean	0.065	0.066	0.065	0.060	0.052	0.077	0.063	0.060	0.070
Television										
	Coefficient	0.023*	-0.004	0.071***	-0.018	-0.036*	-0.000	0.005	-0.018	0.041*
	SE	(0.014)	(0.015)	(0.029)	(0.018)	(0.019)	(0.038)	(0.011)	(0.012)	(0.023)
	Control mean	0.124	0.090	0.195	0.153	0.104	0.245	0.137	0.096	0.218
Assignments/homework										
	Coefficient	0.020	0.026	0.016	0.037	0.036	0.037	0.028*	0.031	0.024
	SE	(0.020)	(0.025)	(0.033)	(0.023)	(0.029)	(0.039)	(0.015)	(0.019)	(0.025)
	Control mean	0.318	0.319	0.317	0.325	0.322	0.330	0.321	0.320	0.323
Reading textbook										
	Coefficient	0.017	0.028	-0.004	0.019	0.032	-0.011	0.018	0.030	-0.008
	SE	(0.020)	(0.025)	(0.034)	(0.023)	(0.029)	(0.039)	(0.015)	(0.019)	(0.026)
	Control mean	0.453	0.458	0.441	0.459	0.463	0.451	0.456	0.460	0.446
Reading story books										
	Coefficient	-0.003	0.008	-0.022	0.019	0.026	0.011	0.007	0.016	-0.007
	SE	(0.014)	(0.016)	(0.024)	(0.015)	(0.020)	(0.028)	(0.010)	(0.013)	(0.018)
	Control mean	0.141	0.133	0.160	0.137	0.128	0.154	0.139	0.131	0.157
Reading for pleasure										
	Coefficient	0.013	0.016	0.003	-0.004	-0.000	-0.021	0.006	0.009	-0.007
	SE	(0.012)	(0.015)	(0.021)	(0.014)	(0.018)	(0.027)	(0.009)	(0.012)	(0.016)
	Control mean	0.099	0.098	0.100	0.119	0.107	0.144	0.108	0.102	0.120
Tutoring others										
	Coefficient	0.007	0.007	0.006	-0.018**	-0.024**	-0.006	-0.004	-0.006	0.001
	SE	(0.006)	(0.008)	(0.010)	(0.008)	(0.011)	(0.013)	(0.005)	(0.006)	(0.008)
	Control mean	0.028	0.030	0.022	0.043	0.044	0.040	0.034	0.036	0.030
Joint p-value		0.162	0.310	0.365	0.704	0.273	0.455	0.175	0.112	0.888
Observations		3,999	2,666	1,331	3,763	2,488	1,247	7,762	5,154	2,578
Experiment 1		Yes	Yes	Yes	No	No	No	Yes	Yes	Yes
Experiment 2		No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Households w/o electricity		Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
Households w/ electricity		Yes	No	Yes	Yes	No	Yes	Yes	No	Yes

Notes: Table shows the effect of receiving access to the Elimuleo platform on learning activites that students used when schools were closed during the April break between the 2021 and 2022 academic school years. Outcomes are based on caregiver reports. Each observation represents a child in a household. The sample restricts to children aged six to thirteen. Standard errors are clustered at the household level and are reported in parentheses. All specificatons control the age of the child and the randomization stratum. All specificatons control ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

Table 11: Heterogeneous gender effects on parent reported academic outcomes
Households without electricity only

	Enrolled in school			Attending school on day of interview			Grade attainment		
Estimate	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Elimuleo offer	0.011	-0.011	0.001	0.070**	0.074**	0.072***	0.031	0.234**	0.132*
	(0.011)	(0.011)	(0.008)	(0.032)	(0.036)	(0.024)	(0.092)	(0.103)	(0.069)
Female child	-0.005	0.002	-0.002	-0.000	0.024	0.012	0.060	0.247***	0.150**
	(0.013)	(0.011)	(0.009)	(0.031)	(0.034)	(0.023)	(0.086)	(0.096)	(0.064)
Elimuleo × Girl	-0.004	-0.003	-0.003	-0.057	-0.039	-0.047*	0.187*	-0.160	0.018
	(0.016)	(0.015)	(0.011)	(0.039)	(0.042)	(0.029)	(0.112)	(0.121)	(0.082)
Control mean	0.958	0.975	0.966	0.649	0.589	0.619	5.178	5.042	5.110
Observations	2,375	2,468	4,843	2,286	2,439	4,725	2,289	2,390	4,679
Included samples									
Experiment 1	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes
Experiment 2	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes

Notes: Table shows the effect of receiving access to the Elimuleo platform on caregiver reported outcomes. Each observation represents a child in a household. The analysis restricts to children aged six to thirteen. Standard errors are clustered at the household level and are reported in parentheses. All specifications control for randomization strata and the age of the child. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

Appendices

Appendix A Additional results

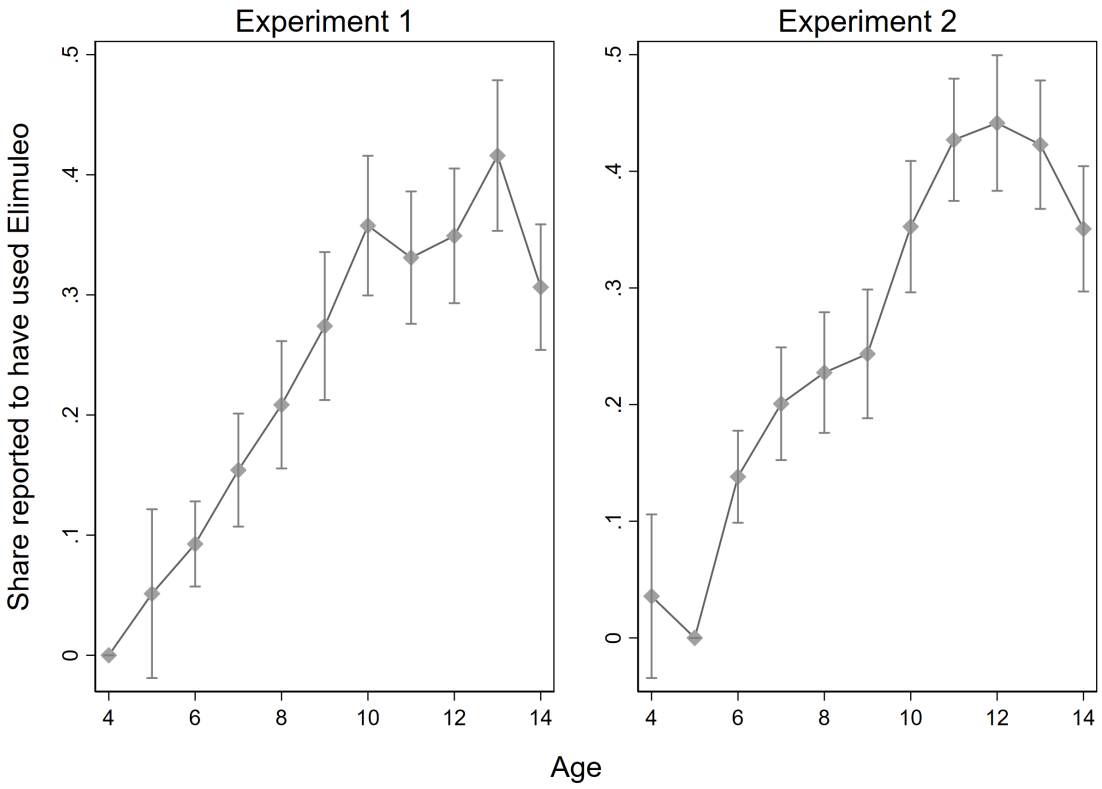

Figure A.1: Figure shows the fraction of children whose parents reported that they ever used the ElimuLeo service by age. Sample is restricted to households who received regular reminders to use the service.

Table A1: Topics available by grade level

Grade	Term	Addition	Subtraction	Multiplication	Division	Mixed	Combined	Algebra
1	1	Y	Y					
	2	Y	Y					
	3	Y	Y					
2	1	Y	Y					
	2	Y	Y					
	3	Y	Y	Y				
3	1	Y	Y	Y	Y			
	2	Y	Y	Y	Y			
	3	Y	Y	Y	Y			
4	1	Y	Y	Y	Y			
	2	Y	Y	Y	Y			
	3	Y	Y	Y	Y			
5	1	Y	Y	Y	Y			
	2	Y	Y	Y	Y			
	3	Y	Y	Y	Y			
6	1	Y	Y	Y	Y			
	2	Y	Y	Y	Y			
	3	Y	Y	Y	Y			
7	1					Y	Y	Y
	2					Y	Y	Y
	3					Y	Y	Y
8	1					Y	Y	Y
	2					Y	Y	Y
	3					Y	Y	Y

Notes: This table shows the topics that are available at each grade level.

Table A2: Follow-up margins

	Experiment 1		Experiment 2		Pooled	
	(1)	(2)	(3)	(4)	(5)	(6)
Any call attempt	0.96	0.016***	0.53	0.239***	0.70	0.148***
		(0.005)		(0.011)		(0.007)
N		4,551		6,801		11,352
Call answered (conditional on attempt)	0.74	0.036***	0.66	-0.006	0.70	0.016
		(0.013)		(0.015)		(0.010)
N		4,410		4,408		8,818
Consented (conditional on call answered	0.53	0.072***	0.48	0.034**	0.51	0.054***
		(0.015)		(0.016)		(0.011)
N		4,410		4,408		8,818
Interviewed successfully (conditional on attempt)	0.53	0.071***	0.48	0.034**	0.51	0.054***
		(0.015)		(0.016)		(0.011)
N		4,410		4,408		8,818
Interviewed (unconditional)	0.51	0.078***	0.26	0.149***	0.36	0.120***
		(0.015)		(0.011)		(0.009)
N		4,551		6,801		11,352
Passed screening questions	0.73	0.182***	0.94	0.025***	0.82	0.114***
		(0.013)		(0.008)		(0.008)
N		3,336		2,930		6,266
Callback interview (conditional on interviewed)	0.27	-0.189***	0.05	-0.012	0.17	-0.112***
		(0.015)		(0.009)		(0.009)
N		2,508		2,249		4,757

Notes: All specifications including the Experiment 1 sample control for randomization strata (the county of the respondent according to PxD records). The pooled specifications also include an indicator for whether the household was in Experiment 2. *Interviewed succesfully* indicates that the household completed at least some survey questions. Twenty households who consented and were available were not surveyed because the respondent was a minor. *Callback interview* indicates that the household was reached as part of the call back survey that resurveyed households that were mistakenly screened out of the first round of surveys. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

Table A3: Heterogeneous impact of access to Elimuleo on math skills
Experiment 1

Estimate	Has outcome (1)	Last assessment date (2)	Assessed grade at baseline (3)	Questions answered (4)	Highest asssessed grade level (5)
Elimuleo offer	-0.028	-37.0***	-0.541***	78.3***	0.021
	(0.024)	(2.2)	(0.136)	(10.4)	(0.124)
Household has electricity	-0.032	-1.0	-0.192	3.7	-0.306*
	(0.031)	(2.9)	(0.176)	(13.4)	(0.160)
Elimuleo × electricity	0.016	-4.8	0.685***	-29.8*	0.501***
	(0.040)	(3.8)	(0.234)	(17.9)	(0.213)
Control mean	0.278		5.576	30.8	6.543
Observations	2,450	853	812	852	853

Notes: Table shows the effect of receiving access to the Elimuleo platform on the highest assessed grade level for *Experiment 1* households that used the platform after June 15, 2021. Each observation represents a household. Use of the platform at different points in time varies across treatment and control so that the latest date of assessment is potentially endogenous to treatment. The outcome in Column 1 is an indicator for whether the household's account used the platform after June 15. The outcome in Column 2 is the date of the last assessment. The outcome in column 3 is the baseline assessment of households that used the the platform after June 15. Column 4 reports the effect of access on the number of questions answered on the platform over the course of the study for those who used the platform after June 15. Column 5 reports the impact on the highest assessed grade level. All specifications control for randomization strata. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

Table A4: Heterogeneous impact of access to Elimuleo on math skills
Experiment 2

Estimate	Has outcome (1)	Last assessment date (2)	Assessed grade at baseline (3)	Questions answered (4)	Highest asssessed grade level (5)
Elimuleo offer	0.050*	-33.8***	-0.017	43.6***	0.286***
	(0.029)	(2.3)	(0.110)	(9.5)	(0.111)
Household has electricity	-0.073**	-3.8	0.066	-1.3	0.019
	(0.036)	(3.2)	(0.152)	(13.1)	(0.153)
Elimuleo × electricity	-0.022	-0.6	-0.108	-16.0	0.002
	(0.047)	(4.0)	(0.193)	(16.6)	(0.193)
Control mean	0.372		5.285	46.2	6.430
Observations	2,113	1,043	1,003	1,043	1,043

Notes: Table shows the effect of receiving access to the Elimuleo platform on the highest assessed grade level for *Experiment 2* households that used the platform after June 15, 2021. Each observation represents a household. Use of the platform at different points in time varies across treatment and control so that the latest date of assessment is potentially endogenous to treatment. The outcome in Column 1 is an indicator for whether the household's account used the platform after June 15. The outcome in Column 2 is the date of the last assessment. The outcome in column 3 is the baseline assessment of households that used the the platform after June 15. Column 4 reports the effect of access on the number of questions answered on the platform over the course of the study for those who used the platform after June 15. Column 5 reports the impact on the highest assessed grade level. All specifications control for randomization strata. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

Table A5: Effects assessed math grade level on single child households
Experiment 2

Outcome	(01apr) (1)	(01jun) (2)	(15jun) (3)	(01jul) (4)	(01aug) (5)
Panel A: Effect of treatment on follow-up rate					
Coefficient	0.192***	0.087**	0.043	-0.026	-0.254***
SE	(0.038)	(0.040)	(0.040)	(0.039)	(0.036)
Control mean	0.392	0.355	0.348	0.345	0.339
Observations	674	674	674	674	674
Panel B: Effect of treatment on last assessment date					
Coefficient	-37.2***	-33.7***	-31.2***	-25.8***	-2.9
SE	(5.3)	(3.4)	(3.0)	(2.9)	(3.5)
Observations	447	348	318	286	187
Panel C: Effect of treatment on highest assessed grade at baseline					
Coefficient	0.365**	0.412**	0.350*	0.239	0.358
SE	(0.173)	(0.190)	(0.199)	(0.204)	(0.227)
Control mean	5.243	5.247	5.250	5.265	5.304
Observations	423	336	309	278	183
Panel D: Number questions answered					
Coefficient	25.7***	35.6***	38.6***	45.8***	78.1***
SE	(7.8)	(9.7)	(10.5)	(11.7)	(18.8)
Control mean	39.7	40.6	41.3	41.4	42.0
Observations	447	348	318	286	187
Panel E: Effect of treatment on highest assessed grade level					
Coefficient	0.312***	0.409***	0.424***	0.480***	0.531***
SE	(0.121)	(0.134)	(0.139)	(0.141)	(0.168)
Control mean	6.405	6.400	6.400	6.405	6.459
Observations	447	348	318	286	187

Notes: Table shows the effect of receiving access to the Elimuleo platform on the highest assessed grade level for *Experiment 1* households that used the platform. Each observation represents a household. Use of the platform at different points in time varies across treatment and control so that the latest date of assessment is potentially endogenous to treatment. Column 1 restricts the analysis to those applicants who last used the platform after April 1, 2021. Coumns 2-5 test the robustness of the results to changing this cut-off to restrict the analysis to those households that used the platform more recently. Panel A reports the effect of treatment on whether a household has an assessed grade recent enough to satisfy these inclusion criteria. Panel B reports the effect on the date from which the student's assessed level comes from. Panel C reports a comparison of baseline assessed levels in treatment and control for those satisfying the inclusion criteria (balance). Panel D reports the difference in the number of questions answered by October 1, 2021. Panel E reports the estimated effect on the highest assessed grade level, controlling for the baseline level. Robust standard errors reported in parentheses. All specifications include a linear control for the probability of treatment for the household's randomization strata, baseline reported grade level, and subject specific baseline assessed grade. To accommodate missingness in the subject-specific assessed grade level, an indicator is included in the controls for whether the household was missing a particular subject-specific assessment, and missing baseline assessed level is imputed to the number four. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

Appendix B Item Difficulty Rating

This appendix provides the item difficulty rating of different types of problems. Items are classified by their difficulty according to the academic term in which the national curriculum would have taught them.

Table B1: Difficulty rating of addition items

	Term 1	Term 2	Term 3
Class 1	Add two one-digit numbers. Solution must be one-digit Ex: 3 + 5	Add three one-digit numbers. Solution must be one-digit Ex: 2 + 1 + 4	Add a two-digit number to a one-digit number. No regrouping. Ex: 52 + 6
Class 2	Add two multiples of 10 between 0 and 90. Solution must be two-digit Ex: 60 + 20	Add three one-digit numbers. Solution must be less than 20. Ex: 4 + 1 + 8	Add two two-digit numbers. Solution must be two-digit. Ex: 43 + 45
Class 3	Add a two-digit number and a three-digit number. Regrouping in 0 or 1 positions. Ex: 275 + 16	Add three one-digit numbers. Ex: 8 + 5 + 9	Add two three-digit numbers. Must regroup in ones and tens but not hundreds. Ex: 657 + 148
Class 4	Add two four-digit numbers. Regroup in 1 position. Ex: 2045 + 4129	Add two four-digit numbers. Regroup in 2 positions. Ex: 1451 + 4656	Round two three-digit numbers then add together. Ex: 164 + 669 = ? (Round each number to the nearest 10 and then add them to estimate the answer).
Class 5	Add two five-digit or six-digit numbers. No regrouping. Ex: 354008 + 114841	Add three six-digit numbers. No regrouping. Ex: 310504 + 374031 + 112034	Add three six-digit numbers. Ex: 121974 + 491267 + 344319
Class 6	Add two three-digit to six-digit numbers (with two digits after the decimal). No regrouping after the decimal Ex: 2582.52 + 7249.13	Add two three-digit to six-digit numbers (with two digits after the decimal). With regrouping in at least 1 position after the decimal Ex: 135.59 + 2989.36	Add two five-digit to eight digit numbers (with four digits after the decimal) (three problems) Ex: 1271.5912 + 3232.0093 Add three five-digit to eight digit numbers (with four digits after the decimal) (two problems) Ex: 3076.8415 + 900.805 + 4387.8262

Notes: This table shows how addition items are rated according to difficulty.

Table B2: Difficulty rating of subtraction items

	Term 1	Term 2	Term 3
Class 1	Subtracting two one-digit numbers Ex: 5 - 3	Subtracting a one-digit number from a two-digit number (between 10 and 19) Ex: 12 - 4	Subtracting a two-digit number from a two-digit number (multiples of 10 only) Ex: 70 - 20
Class 2	Subtracting a one-digit number from a two-digit number (no regrouping) Ex: 27 - 6	Subtracting a two-digit number from a two-digit number (no regrouping) Ex: 55 - 31	Subtracting two one-digit numbers from a two-digit number (no regrouping) Ex: 78 - 3 - 4
Class 3	Subtracting a three-digit number from a three-digit number (no regrouping) Ex: 582 - 260	Subtracting a three-digit number from a three-digit number (with regrouping in one position) Ex: 717 - 276	Subtracting a one-digit number and a two-digit number from a three-digit number (with regrouping) Ex: 455 - 71 - 3
Class 4	Subtracting a three-digit number from a three-digit number (with regrouping in two positions) Ex: 672 - 287	Subtracting a four-digit number from a four-digit number (no regrouping) Ex: 2889-1143	Subtracting a four-digit number from a four-digit number (with regrouping) Ex: 4132 - 3379
Class 5	Using rounding to estimate the result of subtracting a three-digit number from a three-digit number. Ex: Round each number to the nearest ten to estimate 673 - 528	Subtracting a six-digit number from a six-digit number (with regrouping) Ex: 682464 - 163911	Using rounding to estimate the result of subtracting a six-digit number from a six-digit number. Ex: Round each number to the nearest thousand to estimate 304964 - 223178
Class 6	Subtracting a six-digit number from a six-digit number (with two digits after the decimal point) Ex: 6137.55 - 3439.48	Subtracting a five-digit number (with one digit after the decimal point) from a six-digit number (with two digits after the decimal point Ex: 4776.4 - 1129.82	Subtracting a nine-digit number from a nine-digit number (with four digits after the decimal point) Ex: 40485.6633 - 13705.6567

Notes: This table shows how subtraction items are rated according to difficulty.

Table B3: Difficulty rating of multiplication items

	Term 1	Term 2	Term 3
Class 2	N/A	N/A	Each problem has an instruction saying " # + # + # = ##, Then # X ? = ##". The # will be a number between 1 and 9 and the number of plus signs will be between 2 and 5. Ex: 7 + 7 = 14 Then, 7 X ? = 14
Class 3	Multiply 1, 2, 3, 4, 5 or 10 by 1, 2, 3, 4, 5 or 10 Ex: 3 X 5	Multiply a number between 1 and 10 by a number between 1 and 10 Ex: 4 X 7	Multiply a number between 8 and 12 by a number between 1 and 10 Ex: 11 X 6
Class 4	Multiply a number between 1 and 50 by a multiple of 10 between 0 and 90 Ex: 28 X 40	Multiply a one-digit or two-digit number by a one-digit or two-digit number. No regrouping in the ones or tens place. Ex: 21 X 51	Multiply a one-digit or two-digit number by a one-digit or two-digit number. Ex: 88 X 3
Class 5	Multiply a one-digit or two-digit number by a two-digit or three-digit number. No regrouping in the ones, tens, hundreds, or thousands place. Ex: 72 X 110	Multiply a one-digit or two-digit number by a two-digit or three-digit number. Solution must be less than 10000. Ex: 17 X 477	Multiply a one-digit or two-digit number by a two-digit or three-digit number. Ex: 94 X 132
Class 6	Multiply a one-digit or two-digit number by a two-digit or three-digit number (with one digit after the decimal. Ex: 41.6 X 79	Multiply two two-digit or three-digit numbers (each with one digit after the decimal). Ex: 46.1 X 84.3	Multiply two three-digit or four-digit numbers (each with two digit after the decimal). Ex: 12.56 X 73.85

Notes: This table shows how multiplication items are rated according to difficulty.

Table B4: Difficulty rating of division items

	Term 1	Term 2	Term 3
Class 3	Divide a number by 1, 2, 3, 4, 5 or 10 and the solution will be either 1, 2, 3, 4, 5 or 10	Each problem has an instruction saying "## - # - # - # = 0, Then ## / ? = #". The # will be a number between 1 and 10 and the number of minus signs will be between 2 and 5.	Each problem has an instruction saying " If a X b = c Then c / ? = a" where a and b are integers between 1 and 10
	Ex: 20 / 5	Ex: 35 - 7 - 7 - 7 - 7 - 7 = 0 Then, 35 / ? = 7	Ex: If 7 X 6 = 42 Then, 42 / ? = 7
Class 4	Divide a one-digit or two-digit integer by a one-digit or two-digit integer. All answers are one-digit positive integers.	Calculate the remainder when dividing a two-digit number by a one-digit number	Calculate the remainder when dividing a two-digit number by a two-digit number
	Ex: 63 / 21	Ex: What is the remainder of: 37 / 8	Ex: What is the remainder of: 75 / 22
Class 5	Divide a three-digit number by a two-digit number. All answers are two-digit integers.	Divide a four-digit or five-digit number by a two-digit number. All answers are three-digit integers	Calculate the remainder when dividing a four-digit number by a three-digit number
	Ex: 348 / 29	Ex: 36135 / 55	Ex: What is the remainder of 7022 / 529
Class 6	Divide a five-digit number (with one digit after the decimal) by a two-digit number. Answer will have one digit after the decimal	Divide a four-digit number by a two-digit or three-digit number and round to one place after the decimal.	Divide a four-digit number by a four-digit or five-digit number (with one digit after the decimal) and round to two places after the decimal.
	Ex: 1622.4 / 32	Ex: 4707 / 159 (Round your answer to one place after the decimal)	Ex: 4424 / 229.9 (Round your answer to two places after the decimal)

Notes: This table shows how division items are rated according to difficulty.

Table B5: Difficulty rating of algebra items

	Term 1	Term 2	Term 3
Class 7	Solve for x where x minus a one-digit number is a two-digit number (2 problems) Ex: x - 5 = 42 Solve for x where x plus a one-digit number is a two digit number (2 problems) Ex: 7 + x = 21 Solve for x where a two digit number minus x is a one-digit number Ex: 34 - x = 6	Solve for x where x multiplied by a one-digit or two-digit number is a two-digit or three-digit number (3 problems; answer will be a one-digit number or two-digit number) Ex: 6x = 126 Solve for x where x divided by a one-digit number is a two digit number (2 problems) Ex: x / 13 = 3 Solve for x where a two-digit number or three-digit number divided by x is a one-digit number Ex: 189 / x = 7	Solve for x where x multiplied by a one-digit number plus a one-digit number is a two-digit number (3 problems) Ex: 3x + 7 = 46 Solve for x where x multiplied by a one-digit number minus a one-digit number is a two-digit number (2 problems) Ex: 2x - 5 = 25
Class 8	Solve for x where x multiplied by a one-digit number plus a one-digit number is a two-digit number (1 problem) Ex: 6x + 8 = 38 Solve for x where x multiplied by a one-digit number minus a one-digit number is a two-digit number (1 problems) Ex: 4x - 2 = 50 Solve for x where a two-digit number minus x divided by a one-digit is a one-digit or two-digit number (2 problems) Ex: (50 - x) / 2 = 21 Solve for x where a two-digit number plus x divided by a one-digit is a one-digit or two-digit number (1 problems) Ex: (11 + x) / 8 = 2	Solve for x where x divided by a one-digit number plus a one-digit number is a two-digit number (3 problems) Ex: (x / 3) + 2 = 36 Solve for x where x divided by a one-digit number minus a one-digit number is a two-digit number (3 problems) Ex: (x / 7) - 5 = 89	Solve for x where a two-digit number divided by x plus a one-digit number is a two-digit number (3 problems) Ex: (80 / x) + 8 = 13 Solve for x where x divided by a one-digit number minus a one-digit number is a two-digit number (3 problems) Ex: (68 / x) - 6 = 11

Notes: This table shows how algebra items are rated according to difficulty.

Appendix C Proof of Proposition 1

Comparison of treated to control test scores pooling both periods 1 and 2 yields is a consistent estimator of

$$\rho = E[y|Z_i = 1, D_i \geq 1] - E[y|Z_i = 0, D_i \geq 1] \tag{5}$$

In the control group, $Z_i = 0$, if a household uses the service at all, that usage is all in the second period $D_{i2} = 1$ and $D_{i1} = 0$ so that the average asssessed level y for control households is

$$E[y|\lambda_i \geq \tau] = E[2\alpha_i + \beta_i + \theta_{i,0}|\lambda_i \geq \tau] \ . \tag{6}$$

In the treatment group, for a household with $\lambda \geq \tau + \delta$, we observe

$$E[2(\alpha_i + \beta_i) + \theta_{i,0}|\lambda_i \geq \tau + \delta] \tag{7}$$

and for a household with $\tau \leq \lambda_i \leq \tau + \delta$, we observe

$$E[\alpha_i + \beta_i + \theta_{i,0}|\tau \leq \lambda_i \leq \tau + \delta]. \tag{8}$$

Therefore

$$\begin{aligned}\rho = \pi E[2(\alpha_i + \beta_i) + \theta_i|\lambda_i \geq \tau + \delta] + (1 - \pi)E[\alpha_i + \beta + i|\tau \leq \lambda_i \leq \tau + \delta] \\ - E[2\alpha_i + \beta_i + \theta_i|\lambda_i \geq \tau]\end{aligned} \tag{9}$$

$$\begin{aligned}\rho = \pi E[2(\alpha_i + \beta_i) + \theta_i|\lambda_i \geq \tau + \delta] + (1 - \pi)E[\alpha_i + \beta + i|\tau \leq \lambda_i \leq \tau + \delta] \\ - \pi E[2\alpha_i + \beta_i + \theta_i|\lambda_i \geq \tau + \delta] + (1 - \pi)E[2\alpha_i + \beta_i + \theta_i|\tau \leq \lambda_i \leq \tau + \delta]\end{aligned} \tag{10}$$

$$\rho = E[y|Z_i = 1, D_i \geq 1] - E[y|Z_i = 0, D_i \geq 1] = \pi E[\beta_i|\lambda \geq \tau + \delta] - (1 - \pi)E[\alpha_i|\tau \leq \lambda \leq \tau + \delta] \tag{11}$$