Estimating the Treatment Effect of New Device Deployment on Uruguayan Students' Online Learning Activity

Cecilia Aguerrebere
Fundación Ceibal, Uruguay
caguerrebere@ceibal.edu.uy

Cristóbal Cobo
Fundación Ceibal, Uruguay
ccobo@ceibal.edu.uy

Jacob Whitehill
Worcester Polytechnic
Institute, USA
jrwhitehill@wpi.edu

ABSTRACT

When implementing large-scale educational computing initiatives (e.g., One Laptop Per Child) it is vital to allocate resources for training, support, and device deployment judiciously. One question that arises is how learners' engagement with online educational resources is affected by receiving a new computer; do the benefits justify the costs? In this paper, we perform a quasi-experimental analysis to measure the effect of new device deployment on students' online learning activity, operationalized as either the number of interaction events with their LMS, or the number of attempted exercises in their math ITS. The focus is on 6th-grade learners in Uruguay, which to-date has delivered over 750,000 computers to pupils nationwide. Our results suggest that, relative to learners' online learning activity before device deployment, the absolute effects are small but the relative effect are stat. sig. and surprisingly strong: the estimated relative increase on 2016 students' overall LMS activity is 49%. The effects are positive for both 2015 and 2016 and persist several months after device delivery. Moreover, we find that students attempt to solve stat. sig. more (88%) math problems during the month after they receive a new device. We discuss possible reasons and implications for large-scale educational computing programs.

Keywords

One Laptop Per Child; quasi-experimental design

1. INTRODUCTION

During the past 15 years, there have been numerous large-scale educational interventions worldwide – most notably the One Laptop Per Child (OLPC) [16] and One Tablet Per Child (OTPC) [23] programs – that distribute computers to disadvantaged learners to help them bridge the digital divide and achieve better learning outcomes. Early on, such programs were often viewed as a panacea to equalize education worldwide, and indeed some studies have shown that they can boost learners' writing [15] and math [5] skills, verbal fluency [3], basic cognitive processes [3], and self-efficacy [20]. More subtly, they can also help learners to contribute educational content of their own [11] in an educational ecosphere dominated by Western, English-speaking content-makers.

Above all, however, independent evaluations of OLPC and related programs have shown that achieving meaningful learning gains requires more than just giving students laptops and hoping for a positive change [4, 24, 13, 22]. In order for these initiatives to work, it is vital to provide teachers with training on how to make good use of them as part of the curriculum [6]. Computers can break down, and it is important to provide both hardware and software support to ensure these devices remain usable [21]. Finally, even the best maintained device will eventually become obsolete, and thus money for new *device deployment* must be budgeted.

Effectively implementing large-scale educational computing initiatives requires that resources be apportioned judiciously. One question that arises is: **How are learners' interactions with online educational resources affected by receiving a new laptop or tablet computer?** Distributing computers to every student is expensive, and it is important to establish that they are worth the cost. There are several reasons why new devices might impact learners' behavior: (1) *Different affordances*: the new device may offer new features that enable new kinds of interaction. (2) *Novelty*: the mere fact of receiving a shiny new device may incite learners to use it (at least temporarily). (3) *Replacement of broken hardware*: receiving the new device can enable learners simply to *resume* accessing online content.

One way to measure the effect of new device deployment would be to conduct a randomized-controlled trial (RCT), i.e., randomly select a set of students to whom to give a new device at random times throughout the school year, and compare the outcomes of students who received a new device to those who didn't. However, this would be problematic for logistic, political, and ethical reasons, since some people might believe *a priori* that the benefits of receiving a new device could be significant. In this paper, we instead pursue a quasi-experimental approach: One of the potential opportunities offered by educational data-mining is to estimate the causal impact of different interventions from *observational* datasets, i.e., data that were collected containing many covariates/features but *without* random assignment of treatments to participants. Over the past few decades, a variety of techniques have been developed for this

purpose, including propensity score matching [18], principal stratification [9], regression discontinuity analysis [12], and others [19]. Such methods are only applicable in specific contexts, such as in a *natural experiment* in which an exogenous event causes the treatment assignment to be *essentially* random w.r.t. any variable that could conceivably influence the outcome of the treatment itself (i.e., potential confounds). In such situations, random assignment can be imputed *post hoc*, and treatment effects can be estimated by comparing the treated subjects to the untreated ones.

Our paper represents a **case study in quasi-experimental educational data-mining**: We examine how learners, who received computers as part of OLPC, are affected by new device deployment in terms of their interactions with online educational content. Our geographical focus is on Uruguay, which was one of the largest (in terms of number of pupils receiving a laptop) participants in the OLPC program [13]. During 2007-2016, the government of Uruguay together with the Plan Ceibal organization distributed laptops and tablets to over 750,000 pupils nationwide. The nearly universal implementation of this program within Uruguay offers an opportunity to estimate a "new device effect" since there is no selection bias of who receives a new device. We assess the impact of new device deployment on two dependent variables: (1) the *total number of interaction events* with their learning management system (LMS); and (2) the *number of attempted math exercises* within their mathematics intelligent tutoring system (ITS); in prior research, the number of attempted exercises in ITS has been shown to correlate with students' performance on standardized math tests [7, 19, 8].

1.1 Related work
Many studies have examined the educational impact of OLPC programs in general; however, the issue of *new device deployment* within educational computing initiatives and how they are perceived by and affect users, has received much less attention. Oliver & Goerke [17] conducted a survey of engineering and business students in Australia, Ethiopia and Malaysia to assess learners' willingness to adopt a new device (the HP iPAQ) for educational purposes. One notable result was that female students in the participating countries indicated lower willingness to trial the new devices than their male counterparts. In addition, Lai, et al. [14] surveyed students in Hong Kong on their willingness to adopt new educational technology and found that device compatibility with the students' perceived learning styles would affect their likelihood of using it. Neither study examined quantitatively how new devices impact learning behaviors. Hence, these works can be seen as complementary to ours in that they seek to describe the *interactions* between different types of learners and different types of educational technology that might jointly influence their impact on learning.

1.2 OLPC in Uruguay & Plan Ceibal
Since 2007, Plan Ceibal has provided a computer to almost every student in primary and secondary schools in Uruguay, and also ensured Internet access in schools and as well as public access-points. The initial goals were to reduce the digital divide, promote digital inclusion, and ensure the integration of ICT in education. Since 2011, Plan Ceibal has focused on providing the educational community with a wide range of digital tools, such as an LMS, an ITS for mathe-

Figure 1: The CREA2 LMS used by Plan Ceibal. Students can submit homework, send messages to teachers and other students, view content posted by their teachers, etc.
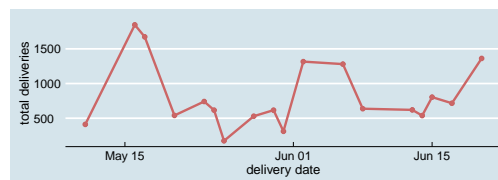

Figure 2: # devices delivered for 6th grade students in 2016

matics, a digital library, a videoconference system to teach English as a second language and facilitate collaboration, etc. The LMS managed by Plan Ceibal is called "CREA2" and is shown in Figure 1. The math ITS is called "PAM".

1.3 Device Delivery Process
Students' devices are upgraded several times during the 9 years of basic education (ages 6-14 years): First graders (6 years old) receive a tablet which they use for two years. In 3rd grade they receive a new tablet which they use for one year only. In 4th grade the tablet is replaced by a laptop, which students use for three years. The laptops are then replaced during either 6th or 7th grade (see Figure 2). Within each school, most (90% of primary and 70% of secondary) students in each classroom receive their new devices at the same time. The schedule is set by Plan Ceibal; larger schools have priority, along with schools located close to the delivery path, etc. While the delivery process is not strictly random, *the delivery dates are independent of many factors* including students' prior LMS and ITS activity, the curriculum the children are learning, dates of examinations, holidays, life-changing events for students, etc. This helps to remove many potential confounds that would impede the inference of treatment effects.

2. EXPERIMENTAL ANALYSIS
We investigate the effect of new device deployment in terms of two dependent variables: (1) Δ **LMS Interaction Events**: The increase in students' activity (total number of interaction events) with the CREA2 LMS *after* receiving their new device compared to their activity *before* receiving it. (2) Δ **ITS Attempted Exercises**: The increase in the number of math exercises that students attempt to solve with the PAM ITS *after* versus *before* receiving their new device.

2.1 Dataset
Δ **LMS Interaction Events**: The dataset includes each student's activity in the CREA2 platform on each day of

Year	Learners	Active L.	Deliv. dates	Sch.	Classr.
2015	13329	7276	21	526	809
2016	25898	16962	18	810	1378

Table 1: Plan Ceibal dataset: total # of considered learners, *active* learners, delivery dates, schools, and classrooms containing students who received devices that year.
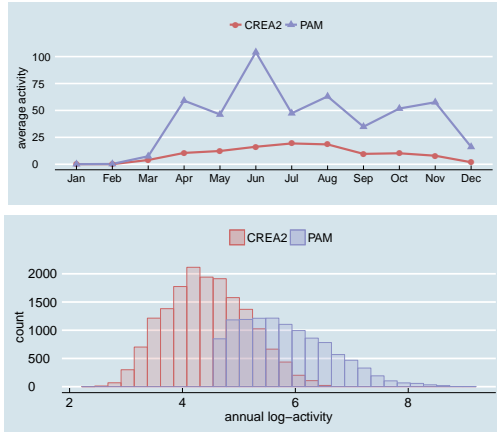

Figure 3: Activity levels in the CREA2 (# actions) and PAM (# attempted exercises) platforms during 2016. **Top:** Average per-student monthly activity. **Bottom:** Histogram of the logarithm of the total annual activity per student.

2015 and 2016, as well as the delivery dates of new devices during that period. A large fraction of the students almost never used the platform. Our focus is on the impact of new devices on *active* students; hence, we limit the universe of study to students who accessed the platform on at least 10 different days in a given year (this is the *active user* definition used at Plan Ceibal). We note that, even with this constraint, the median CREA2 activity level per month is low: only 7 total actions. In addition, we focus exclusively on 6th graders (11 years old), who are the most active CREA2 users. Finally, we only consider delivery dates on which at least 5 new devices were delivered. Table 1 summarizes the sample sizes considered for each dataset for 6th grade.

Δ **ITS Attempted Exercises**: Data were available for 2016 (but not 2015) on 6th-grade students' total math exercises attempted each day. The universe of study is limited to those students who attempted at least 100 exercises in the year (*active user* definition at Plan Ceibal). In addition, during 2017 (but not 2016), the numbers of *correct* and *incorrect* attempted exercises are also available. Figure 3 shows the overall activity levels in CREA2 and PAM. The platforms are offered as a recommended tool for teachers, but their use is not mandatory. Plan Ceibal provides tutorials promoting their use, which are independent of device delivery dates.

Δ **CREA2 activity: 1 month after v. 1 month before**: As a preview of our more detailed analyses below, Figure 4 shows students' *delta* behavior, i.e., their CREA2 activity during the month *after* each delivery date t, minus their activity during the month *before* t. The blue curve shows the deltas for learners who *received* a device on t (along
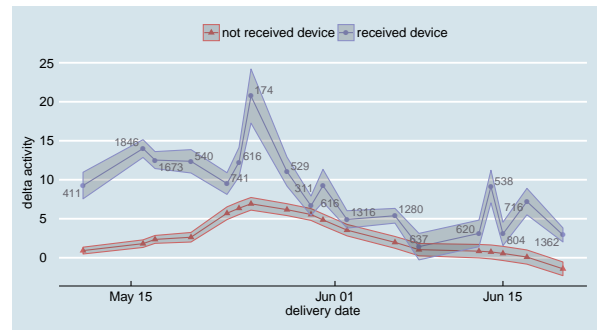

Figure 4: Increase between the average monthly activity in CREA2 after and before each delivery date t, for students who received (blue) and did not receive (red) a device on t, during 2016. The outer bands correspond to 2 standard errors for each mean estimate. The numbers listed by each blue point report how many students received a device on t.
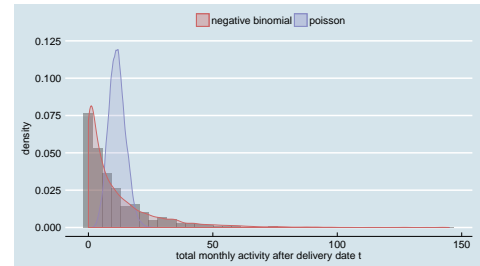

Figure 5: Normalized histogram of the monthly activity after t (for delivery date 2016-05-11).

with the number of such students), and the red curve shows students who received a device (at least one month) *after* t. If giving students a new device has a positive impact on CREA2 activity levels, then we expect the blue curve to be higher than the red curve (which it is).

2.2 Methodology

This is a quasi-experimental study enabled by the *delayed treatment* design [10] that was used in deploying new devices to students: Almost every student in every school who participates in Plan Ceibal eventually receives a new device; hence, there is no selection bias as to who enrolled in the program. In particular, (almost all) students within the same grade of the same school receive their devices on the same date, but these dates are essentially random *across* schools. In particular, the delivery dates are independent of the classroom curriculum and students' prior activity on the CREA2 and PAM platforms. In our analysis, we thus study the effect of device delivery at each delivery date separately and then average these estimates to estimate the average treatment effect across all dates. We do note, however, that our analysis is not immune to all possible confounds, e.g., a relationship between the date of device delivery and whether the school is located in an urban or rural environment.

2.2.1 Data model

Each learner's activity in the CREA2 and PAM platforms consists of *count* data. Suitable models for counts include the Poisson and the negative binomial distributions. The

advantage of the negative binomial is that the variance of the distribution can be set independently of its mean to account for overdispersion of the data. Figure 5 shows the normalized histogram of the total CREA2 activity after a given delivery date t, overlaid onto Poisson and negative binomial probability density functions fit to the histograms using maximum likelihood estimation (Poisson log-likelihood = -70647.08, negative binomial log-likelihood = -22066.75). This comparison shows the clear overdispersion of the considered data, which makes the negative binomial a more accurate approximation than the Poisson model.

Multi-level modeling: Because for each delivery date we are considering students in the same school, and possibly in the same classroom, the activity data for them will be correlated. Hence, a multi-level modeling approach is employed where the classroom effect on the student's activity, often determined by the teacher, will be modeled as a random effect. We only consider deviations of the *intercept* of a classroom from the overall intercept; random *slopes* are not considered. Therefore, we propose to model student i's activity N months after the delivery date t (i.e., between $t + ((N - 1)$ months) and $t + (N$ months)) as a negative-binomial random variable with expected value A_{it} given by:

$$\log(A_{it}) = e_t + \gamma_{0t}b_{it} + \gamma_{1t}d_{it} + C_t, \qquad (1)$$

(capital letters denote random variables and lower-case denote fixed values). The case $N = 1$ corresponds to the activity during the month right after the delivery date. We define a "month" to be 4 weeks (28 days). e_t is the baseline activity in the same time period considered for A_{it}. b_{it} is student i's activity during the month before the delivery date t. d_i is a boolean variable taking value 1 if student i got a new device on t and 0 otherwise. The fixed effects γ_{0t} and γ_{1t} represent the effect on the activity N months after t, of the activity during the month before t, and the device delivery, respectively. The random effect of classrooms is represented by the random variable C_t, assumed to follow a zero-mean Gaussian distribution and standard deviation σ_{C_t}. A nested classroom-school random effect was also explored, but it was discarded because the results were very close.

Gender: Oliver & Goerke [17] found that female students (in Australia, Ethiopia and Malaysia) reported different attitudes towards educational technology than their male counterparts. Might device deployment affect Uruguayan girls and boys differently in terms of CREA2 activity? To investigate, we extended the Model 1 with a boolean variable g_{it} representing the student's gender as well as an interaction between g_{it} and the device delivery variable d_{it}.

Treatment effect: When computing the device delivery effect of a given delivery date t, we compare students who received a device on t (treatment group), to students who received a device on $t^* > t + (N$ months) (control group). In particular, we make sure to exclude from the control group those students whose treatment occurred within N months of students in the treatment group. This analysis is consistent with a delayed treatment design.

2.2.2 Combining per-date estimates
For each of the M considered delivery dates, we compute the maximum likelihood estimator (MLE) of the device effect

γ_{1t}, as well as its associated standard error $SE_{\gamma_{1t}}$. Because different number of students receive/do not receive their devices on each date, the standard errors $SE_{\gamma_{1t}}$ will vary across dates. We model the M estimates $\{\gamma_{1t}\}_{t=1,...,M}$, as independent samples of Gaussian random variables with equal mean γ_1 and different standard deviations $SE_{\gamma_{1t}}$. Then, the MLE of the device delivery effect $\hat{\gamma}_1$ is given by averaging the individual γ_{1t}'s weighted by the inverse square of their standard errors. From $\hat{\gamma}_1$ and its standard deviation we can compute confidence intervals, and perform a t-test to assess the statistical significance of the device delivery effect [2].

To ensure that the treatment effect estimates $\{\gamma_{1t}\}_{t=1,...,M}$ across delivery dates are statistically independent, each group of students belonging to the same classroom is used to estimate the treatment effect for one delivery date only. That is, all the classrooms under consideration are partitioned over delivery dates, and the treatment effect for each date is computed only from the students assigned to that date. Some of these students will be in the treatment group (those who received the device that day) and others will be in the control group (those who received the device later).

To partition students across delivery dates, we used a greedy algorithm whereby one classroom is assigned to a delivery date at a time: For each classroom, one of the M delivery dates t is chosen with probability p_t, which is inversely proportional to the total number of students already assigned plus the total potential number of students that could be assigned to each date – thus favoring dates not yet assigned and with few potential students. We ran this procedure 100 times and selected the assignment with smallest variance in the number of students assigned per date, which helps to avoid possible numerical issues in model estimation.

Implementation: Models were fit using the *glmer.nb* function of the R *lme4* package. To detect possible convergence problems, each experiment was run using several different optimizers and consistent results were verified [1].

3. RESULTS I: LMS INTERACTIONS
Table 2 shows the estimated effects (averaged over all delivery dates) of delivering a new device on the number of CREA2 interaction events, for 2015 and 2016. Since the computed effects are in *logarithmic* scale (see Equation 1), a log-effect of 0 corresponds to $\exp(0.0) = 1.0$ in the original scale, i.e., no impact on CREA2 activity, whereas 0.76 equals $\exp(0.76) = 2.1$ in the original scale, i.e., a 110% activity increase. Table 3 shows the average, over all delivery dates, of the rest of Model 1 parameters in original scale.

Even after accounting for class-specific random effects as well as students' prior baseline activity levels, **we observe a clear CREA2 activity boost in the 1 month following the device delivery** (first row in Table 2): a relative effect of $\exp(0.76) = 2.14$ (114% increase) for 2015 and $\exp(0.40) = 1.49$ (49% increase) for 2016. In other words, while the absolute increases are low (due to the low overall CREA2 activity usage – Figure 3), the relative effect is high. Results for 2016 tend to be less noisy because the activity levels are larger compared to 2015. Though present in both years, the effect clearly decreases from 2015 to 2016.

	2015		2016	
	all class.	**class.≥10 stud.**	**all class.**	**class.≥10 stud.**
N-months	$\hat{\gamma}_1$ $SE_{\hat{\gamma}_1}$	$\hat{\gamma}_1$ $SE_{\hat{\gamma}_1}$	$\hat{\gamma}_1$ $SE_{\hat{\gamma}_1}$	$\hat{\gamma}_1$ $SE_{\hat{\gamma}_1}$
1	0.76 0.12 ***	0.69 0.14 ***	0.40 0.04 ***	0.41 0.04 ***
2	0.84 0.13 **	0.80 0.15 *	0.20 0.04 **	0.19 0.04 *
3	0.83 0.27	0.50 0.29	0.31 0.14	0.28 0.13

Table 2: Effects of delivering a new device to each student on their CREA2 activity for academic years 2015 and 2016, over a period of $N = 1, 2, 3$ months, in logarithmic scale. Weighted averages, together with their corresponding standard errors, are reported. Significance codes: 0 (***), 0.001 (**), 0.01 (*), 0.05 (.), 0.1, () 1.

	Fixed eff.		**Random eff.**		**Fixed eff.**		**Random eff.**
year	e	γ_0	σ_C	**year**	e	γ_0	σ_C
2015	1.17	1.06	3.18	**2016**	5.92	1.02	2.88

Table 3: Average of the model parameter estimates among delivery dates in the original scale.

Temporal evolution: The second and third rows of Table 2 show the estimates of the effect of device delivery in students' monthly CREA2 activity 2 and 3 months after the delivery date, respectively. The effect is still present two months after the delivery date, and it appears to be stable in 2015 and to decrease in 2016. The effect is not statistically significant three months after delivery; this may due to the small number of samples available at that time. (Note that examining $N > 3$ was not possible since there were too few students who had not yet received a device who could serve as a control group.)

Classroom size: No significant differences are observed when comparing the estimates considering all classrooms or only those with at least 10 students (see Table 2).

Highly active students: We also conducted the analysis on only those students who accessed CREA2 on at least 25 different days in 2016. (Note that 2015 data could not be analyzed due to small sample size.) The results were consistent with what we found above for all students, with 43% activity increase right after receiving the new device.

Activity change: To investigate what *kinds* of CREA2 activities were affected, Figure 6 shows the percentage of the monthly activity increase, at different time points (the month before t, the month after t, and two and three months after t), of the students who received the device on t relative to those who received it after t. For instance, the average increase in the number of comments posted during the month following t, by the students who received their new device on that date, was 70% larger than that of students who received their device later in the year. For some activity types, the boost is larger and remains longer in time (e.g., comments posted, item submissions). Note that the sum over all activities at the first time-point ($t - 1$) should be close to zero, denoting similar total activity for all users before device delivery.

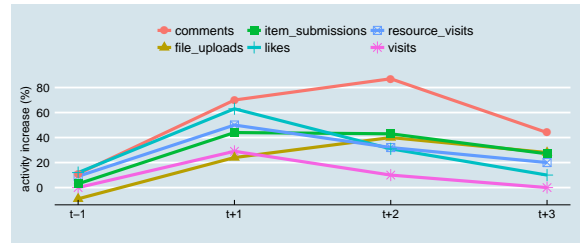**Gender effect**: Using the extended model to support pos-

Figure 6: Monthly activity increase (%), at different time points, of students who received the device on t relative to those who did not receive it.
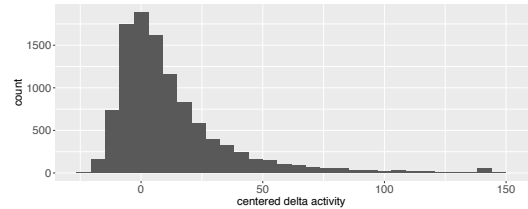

Figure 7: Histogram of Δ CREA2 activities of treated students (over all delivery dates) w.r.t. median activity of untreated students.

sible gender effects, we found a clear difference on the total activity: girls performed about 32% more total CREA2 activity compared to boys in 2015 (and 28% more in 2016). We observed no significant difference, however, in the effect of device delivery between boys and girls.

Who drives the effect?: Was the the boost in CREA2 activity driven by a large increase among a small number students? To explore this question, we calculated the change (Δ_i) in CREA2 activity level for each student i before/after treatment, minus the median change in activity level for all untreated students. We then computed a histogram over the Δ_i values over all students and delivery dates in 2016. The histogram is shown in Figure 7. Since there is a positive and statistically significant treatment effect, the mean of the histogram is greater than 0. The histogram also shows smooth gradation from small effects to large effects and provides evidence that the average treatment effect is due to increased activity levels among many students, not just a few.

4. RESULTS II: MATH PROBLEM ATTEMPTS

Similar to the analysis of device deployment on students' CREA2 LMS activity, we used the same model (Eq. 1) to assess the potential impact on the number of *attempted exercises* in the PAM math ITS provided by Plan Ceibal.

Results: A positive ($\hat{\gamma}_1 = 0.63$, $SE_{\hat{\gamma}_1} = 0.14$) and statistically significant ($p < 0.01$) effect is observed during the one month following the device delivery, with a 87% increase in the total attempted exercises ($\exp(0.63) = 1.87$). The effects observed two and three months later are small and not statistically significant, suggesting that the boost disappears over time.

5. CORRECTNESS OF MATH EXERCISES

In addition to the number of *attempted* math exercises, we also explored whether receiving a new device helps students to complete exercises *correctly*. Possible reasons include: (1) the new device has a better user interface that helps students avoid careless data-entry errors; and (2) the learners' improved user experience encourages them to practice more often and thereby improve their math skills.

To assess the impact of device deployment on correctness $(0 - 1$ scale) of submitted exercises, it was not possible to use the same methodology as for LMS activity. The reason is that PAM data on correct/incorrect exercises are available only for 2017, and during this year only a small number of 6th grade students received a device. Hence, we resorted to a correlational analysis in which we estimated the change in exercise correctness *without* a control group. In particular, we estimated the treatment effect based on the average accuracy N months after device delivery minus the average accuracy N months before delivery, and averaged across all treatment dates. Because no control group is used (unlike in the previous analyses), there may be confounding factors affecting this analysis.

Results: None of the average delta accuracies for 4th, 5th and 6th grade, computed either on individual students (Δacc_S) or on classrooms (Δacc_C), was statistically significant.

6. DISCUSSION
The results suggest that receiving new devices resulted in a strong relative **increase in learners' CREA2 LMS activity**: 114% in 2015 and 49% in 2016. Within sensitivity analyses based on academic year (2015 and 2016), classroom size, and students' baseline activity levels, we found that the trends were similar: new devices result in increased LMS activity. Moreover, the boost in activity persists up to 3 months after device delivery. We note again that the *absolute* average CREA2 activity levels were very low; hence, the increase may only amount to a few extra logged events (about 10 extra actions per student per month).

Receiving new devices not only increases the activity but also **alters the kind of activities** performed in the platform (Section 3). The fact that device delivery increases (w.r.t. learners who did not receive a device) the number of *comments* and *resource_visits* even more than just *visits* (which reflects merely accessing the CREA2 web page) suggests that learners are engaging more *substantively* with the LMS after receiving their new device.

Within the math ITS, we observed that, during the 1 month after receiving a new device, **learners attempted to solve more (88%) math problems**. However, the results were not statistically significant two and three months after delivery, suggesting that the impact is short-lived. We found no evidence (given the limited data available in 2017) that new devices resulted in higher *accuracy* of attempted exercises.

6.1 Possible explanations
Novelty: Receiving a brand-new device could potentially increase students' motivation to use them, but the effect might diminish over time. In our data, we do observe that the LMS activity boost, as well as the boost in number of attempted ITS math exercises, declines over time (though

more strongly for the ITS than for the LMS) after receiving a device declines – which suggests possible novelty effects.

Availability: Oftentimes, devices are not available to students because of recurrent failures. Hence, receiving a new device not only means having a new, more performant one but having a working device at all. It is possible that a student who suddenly (due to device deployment) regains access to a working computer might resume CREA2 activity at a much higher level after receiving it. The strong activity gains we observe are compatible with this hypothesis (though they cannot directly confirm it).

7. SUMMARY AND CONCLUSIONS
We conducted a quasi-experimental analysis (on 24,000 learners over 2 years) to estimate the treatment effect of giving OLPC students new computers. We harnessed the facts that (1) all students were eventually treated, so that there was no selection bias, and (2) the device deployment schedule was random w.r.t. a variety of potential confounds (e.g., students' prior LMS/ITS activity). The main results include:

(1) When students receive a new device, they interact more with their schools' LMS and engage more (attempt more exercises) with their math ITS, compared to learners who had not yet received a device upgrade. To the extent that increased engagement with educational content and practice in solving exercises contributes to students' learning [7, 19, 8], OLPC programs should try to provide students with up-to-date devices in a timely and cost-effective manner.

(2) While conducting these analyses we discovered that the *absolute* baseline activity levels of many learners in the examined dataset were very small. This raises the question of whether teachers are receiving proper training on how to use online learning resources effectively and how to instruct and encourage their learners to engage with them.

(3) Our study indirectly raised the question of how often a new device delivery simply replaces a device that had *broken*. For researchers who wish to assess the potential benefits of OLPC programs, it is important to take into account how many students truly have access to a working device (not just a broken one). For administrators, it underlines how technical support may play an important role in ensuring the success of large-scale educational computing initiatives.

(4) The fact that new device deployment increases CREA2 and PAM activities – even if the effects are transient – is evidence that learners' activities *can* be incited to engage more with educational platforms. One way is through renewed hardware, as explored in this paper. Another way is to help teachers to *use* these platforms more effectively [22].

Future research: It would be interesting to explore whether novelty, new features, or replacing broken hardware contributes more to the overall treatment effect; to this end, it would be useful to ask learners themselves about how they perceive and interact differently with new devices.

8. REFERENCES
[1] BATES, D., MÄCHLER, M., BOLKER, B., AND WALKER, S. Fitting linear mixed-effects models using

lme4. *J Stat Softw 67*, 1 (2015), 1–48.

[2] BORENSTEIN, M., HEDGES, L. V., HIGGINS, J. P., AND ROTHSTEIN, H. R. A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods 1*, 2 (2010), 97–111.

[3] CRISTIA, J., IBARRARAN, P., CUETO, S., A., S., AND SEVERIN, E. Technology and child development: Evidence from the one laptop per child program. *IZA Paper no. 6401* (2012).

[4] DE MELO, G., MACHADO, A., AND MIRANDA, A. The impact of a one laptop per child program on learning: Evidence from uruguay. *IZA Paper no. 8489* (2014).

[5] DÍAZ, A., NUSSBAUM, M., AND VARELA, I. Orchestrating the xo computer with digital and conventional resources to teach mathematics. *J Comput Assist Lear 31*, 3 (2015), 202–219.

[6] FAJEBE, A. A., BEST, M. L., AND SMYTH, T. N. Is the one laptop per child enough? viewpoints from classroom teachers in rwanda. *Information Technologies & International Development 9*, 3 (2013), pp–29.

[7] FENG, M., HEFFERNAN, N., AND KOEDINGER, K. Addressing the assessment challenge with an online system that tutors as it assesses. *User Modeling and User-Adapted Interaction 19*, 3 (2009), 243–266.

[8] FENG, M., AND ROSCHELLE, J. Predicting students' standardized test scores using online homework. In *Proceedings of the Third (2016) ACM Conference on Learning@ Scale* (2016), ACM, pp. 213–216.

[9] GALLOP, R., SMALL, D. S., LIN, J. Y., ELLIOTT, M. R., JOFFE, M., AND TEN HAVE, T. R. Mediation analysis with principal stratification. *Statistics in medicine 28*, 7 (2009), 1108–1130.

[10] HEATH, L., KENDZIERSKI, D., AND BORGIDA, E. Evaluation of social programs: A multimethodological approach combining a delayed treatment true experiment and multiple time series. *Evaluation Review 6*, 2 (1982), 233–246.

[11] HOURCADE, J. P., BEITLER, D., CORMENZANA, F., AND FLORES, P. Early olpc experiences in a rural uruguayan school. In *CHI'08 extended abstracts on Human factors in computing systems* (2008), ACM, pp. 2503–2512.

[12] IMBENS, G. W., AND LEMIEUX, T. Regression discontinuity designs: A guide to practice. *Journal of econometrics 142*, 2 (2008), 615–635.

[13] KRAEMER, K. L., DEDRICK, J., AND SHARMA, P. One laptop per child: vision vs. reality. *Communications of the ACM 52*, 6 (2009), 66–73.

[14] LAI, C., WANG, Q., AND LEI, J. What factors predict undergraduate students' use of technology for learning? a case from hong kong. *Computers & Education 59*, 2 (2012), 569–579.

[15] LOWTHER, D. L., ROSS, S. M., AND MORRISON, G. M. When each one has one: The influences on teaching strategies and student achievement of using laptops in the classroom. *Educ. Tech. Research and Development 51*, 3 (2003), 23–44.

[16] NEGROPONTE, N., BENDER, W., BATTRO, A., AND CAVALLO, D. One laptop per child. In *National Educ. Computing Conference in San Diego, CA.*

Retrieved April (2006), vol. 5, p. 2007.

[17] OLIVER, B., AND GOERKE, V. Undergraduate studentsâĂŹ adoption of handheld devices and web 2.0 applications to supplement formal learning experiences: Case studies in australia, ethiopia and malaysia. *International Journal of Education and Development using ICT 4*, 3 (2008).

[18] ROSENBAUM, P. R., AND RUBIN, D. B. The central role of the propensity score in observational studies for causal effects. *Biometrika 70*, 1 (1983), 41–55.

[19] SALES, A. C., AND PANE, J. F. Exploring causal mechanisms in a randomized effectiveness trial of the cognitive tutor. In *Proceedings of the 8th International Conference on Educational Data Mining* (2015).

[20] SHANK, D. B., AND COTTEN, S. R. Does technology empower urban youth? the relationship of technology use to self-efficacy. *Computers & Education 70* (2014), 184–193.

[21] SIMON, S. Laptops may change the way rural peru learns. *Weekend Edition* (2008).

[22] THERIAS, E., BIRD, J., AND MARSHALL, P. Más tecnologia, más cambio?: Investigating an educational technology project in rural peru. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (2015), pp. 447–456.

[23] VIRIYAPONG, R., AND HARFIELD, A. Facing the challenges of the one-tablet-per-child policy in thai primary school education. *Education 4*, 9 (2013), 23–32.

[24] WARSCHAUER, M., ZHENG, B., NIIYA, M., COTTEN, S., AND FARKAS, G. Balancing the one-to-one equation: Equity and access in three laptop programs. *Equity & Excellence in Education 47*, 1 (2014), 46–62.