Evaluation of the Impact of the Adaptive Platform for Mathematics on Learning Achievements[1]

Marcelo Perera and Diego Aboal[2]

October 2017

[1] This is a report of the consultancy assignment contracted by Plan Ceibal to the Center for Economic Research. We want to thank Helena Rovner for the initial impulse to undertake this project and the enormous collaboration from the many technicians from Plan Ceibal, in particular Cecilia Hughes, Inés Méndez, Marcela Gómez, Yanedy Pérez and Emiliano Pereiro. In turn, this study would not have been possible without the information generously provided by the Department of Research and Educational Statistics of CODICEN/ANEP and the collaboration and valuable comments from Andrés Peri and Santiago Cardozo. Notwithstanding the aforementioned, the contents of this document are the exclusive responsibility of its authors.

[2] perera@cinve.org.uy , aboal@cinve.org.uy

Summary

Since year 2013, the [Uruguayan] educational system has been using an adaptive learning tool for mathematics: Plataforma Adaptiva de Matemática (PAM) [The Adaptive Platform for Mathematics]. PAM's content has been adapted to the national curriculum and is a tool that - based on an analysis of students' experiences - offers personalized feedback according to each student's skill level. The use of PAM has been spreading throughout the educational system. By 2016, approximately half of all students in the 3rd through 6th years of primary education had used the platform. The purpose of this study is to identify the effect of the use of PAM on the learning gain in mathematics based on longitudinal data from a sample of students in primary education. The results show a positive effect of 0.20 standard deviations on the learning gain. This is the first evidence at a country-wide level of the impact of a pedagogical tool of this type that shows that the possibility of improving the quality of education through the use of technology is a real alternative.

1. Introduction

In 2013, Plan Ceibal [(the Uruguayan digital learning initiative)] provided the Adaptive Platform for Mathematics (PAM) for students and teachers in primary education. This digital tool is designed for content development and for learning concepts and methods. The platform offers the teachers different resources for creating their lessons, establishing learning objectives and supplying individual or group tasks for their students in order to enrich and to deepen their learning practices. PAM offers more than 100.000 activities that give personalized assistance to each student and adapt to his or her level of knowledge. The platform gives immediate feedback to the student after each answer, offering help and theoretical materials and showing alternative solutions to the problems[3].

Even though we continue to observe considerable expansion in the usage of PAM in primary education, there is still considerable variability in the tendency to use it and in the intensity of use of the platform by teachers and students.

The purpose of this study is to identify the effect that using PAM had on learning gains in mathematics based on longitudinal data from a sample of the cohort of students who were studying in the 3rd year of primary education in 2013. The students' proficiency levels in mathematics and reading were evaluated at the end of 2013 and again three years later. Combining this data with data about the usage of PAM obtained from the platform's data base, we analyzed whether using the tool was a factor associated with the gain in the students' results on the tests.

Currently, we know relatively little about what impact using information and communication technologies (ICT) in education really has on learning. Even less is known about its impact in other dimensions such as the social behavior or a student's eventual performance in adult life. This is particularly relevant in the developing countries where concerns about the quality of education have led to the increasing adoption of ICT in class. A general review of high quality studies on a worldwide level shows mixed evidence regarding the impact of using technology in education (for example, see Bulman and Fairlie, 2016).

This suggests technology's potential in education depends on the details of the intervention. These details determine to what extent a specific intervention is capable of overcoming the restrictions in a student's learning process.

Even though systematic investigations on the effects of technology-assisted learning are rare, some experimental evaluations, such as those by Banerjee et al. (2007) and Muralidharan et al. (2016), have provided solid evidence of the high relative impact of such interventions on students in primary and secondary schools in developing countries. One outcome of these studies is that the educational tools that accounted for the greatest gains appear to be those that use technology to personalize the learning process.

In Muralidharan et al. (2016), the impact of an intervention was evaluated. The intervention included using educational software (for mathematics and language) called Mindspark. Like PAM by

---

[3] The platform is developed by the company bettermarks (http://uy.bettermarks.com) based in Germany. It has been adapted to the national curriculum and currently covers the topics from the 3rd year of primary to the 4th year of secondary education. Further information about PAM and some experiences in schools in Uruguay can be found on the following site: https://www.youtube.com/user/canalceibal/search?query=PAMhttps://www.youtube.com/user/canalceibal/search?query=PAM

Plan Ceibal, Mindspark is designed for personalized student instruction[4]. This kind of software employs algorithms for error recognition and the possibility to give students immediate, personalized feedback (hints, alternative solutions, activities corresponding to their level, etc.). The authors of the study found a positive impact on the learning of mathematics and language for the sample of students in middle school in India. They also state that the relative gain is greater for students with a lower academic level. In their opinion, the main mechanism behind the efficacy of the intervention is the capability of the tool to direct the instruction in a personalized form, given the extremely heterogeneous nature of the students' ability levels[5].

The magnitude of the impact that such interventions have on learning suggests that the possibility of increasing productivity in education through the use of technology is a real option.

PAM is an example of a digital learning tool that adapts to a student's ability level. This analysis provides the first evidence of its impact in the way it was applied in Uruguay. The tool was made available to all teachers and students in the educational system; however, the decision whether and how to use it was left to them. Even though its usage has been spreading, the specific pedagogical use cases are still relatively unknown.

Our strategy to identify the impact of PAM is based on three pillars: i) the possibility to count on longitudinal data about learning, ii) the variability in the usage of PAM between 2014 and 2016 and iii) the possibility of using the gain in reading to control the potential selection bias due to the presence of unobservable factors which may affect the general performance of the student.

Although this is not an experimental study, and therefore the internal validity of the evaluation is based on some assumptions, it has an important advantage over experimental studies in terms of their external validity. This is the possibility of analyzing an intervention at a national level, contemplating effects generally ignored in the studies using small scale interventions[6].

The remainder of this paper is organized as follows: in section 2 we present a brief description of the data used. In section 3 we discuss the methodology applied to estimate the effects of PAM, with a special emphasis on the challenges in identifying a causal relation. In section 4 we present the results. In section 5 we summarize the main findings of the study and comment on them. In the Appendix we give some details regarding the empirical model we used and we present additional estimations as a proof of the robustness of the results we obtained.

---

[4] The students used the software outside of class time for 90 minutes. Half of the time they used it in free activities and the other half under the guidance of an assistant in groups of 12-15 students.

[5] Adapting the instruction to the student's level of learning also explains the high effectiveness of other interventions documented in the literature. Such is the case of differentiated support for students with low performance (for example, see the evaluation of the "Balsakhi" Program in India in Banerjee et al., 2007) or of grouping students according to their skill levels (see the assessment of the impact of "tracking" in Kenya in Duflo et al., 2011)

[6] A discussion of scaling programs that have proven effective in experimental studies on a small scale can be seen in Banerjee et al. (2016).

2. The data

This analysis is based on two sets of data. One set comes from a data base of the national evaluations on learning, provided by the Department of Investigation and Educational Statistics of the Central Directive Council of the National Administration for Public Education (DIEE/CODICEN/ANEP). The second set is from the data base provided by Plan Ceibal and generated from the platform's data base. It contains information on the use of PAM at the student level. The two data bases were connected by means of a unique student identifier, allowing the analysis of the relationship between performance on the tests and the use of the educational tool[7].

The first set of data is a sample taken from two learning assessments in mathematics and reading, carried out with the cohort of students who were in their third year of primary school in 2013. The first of these assessments was given at the end of 2013 and the second at the end of 2016, when the majority of those children were in their 6th year of primary school. The first evaluation corresponds to the Third Regional Comparative and Explanatory Study (TERCE), a large-scale study of learning achievements coordinated by the Latin American Laboratory for Assessment of the Quality of Education (LLECE) by UNESCO, in which Uruguay and 14 other countries in Latin America took part (see UNESCO-OREALC, 2016 for details on the design of the instruments of evaluation)[8]. The second learning assessment of was applied by DIEE/CODICEN/ANEP in 2016 on the same cohort of students who participated in TERCE in 2013, that means the majority of these students were attending the last year of their primary education. The second evaluation was based on the same methodology applied by LLECE/UNESCO in the regional evaluations. The DIEE created a procedure for comparing the results from the 3rd year test to the metric of the 6th year test, thereby allowing the analysis of the students' learning progress between 2013 and 2016 in absolute terms, both in mathematics and reading.

The table below shows the number of students who were evaluated in both mathematics and reading in 2013 and in 2016. This group of students, for which the learning gain between both evaluations could be measured, make up the sample used for this analysis. All the results presented below are based on the information from the 2143 students attending 237 public or private schools all over the country who were evaluated in mathematics and reading in 2013 and 2016[9].

---

[7] The databases managed by the team responsible for this study were previously anonymized in order to protect the identity of the students in the sample.

[8] The TERCE evaluation focuses on the curriculum of each country. All documentation on this evaluation can be found at: http://www.unesco.org/new/en/santiago/education/education-assessmentllece/third-regional-comparative-and-explanatory-study-terce/

[9] The contrasts of differences between the original TERCE sample and the sample finally analyzed were not significant, so bias cannot be expected due to loss of sample.

Table 1. The number of students evaluated in mathematics and reading and the number of schools that participated in the evaluation in each year

	Students	Schools
Evaluation 2013	3761	242
Evaluation 2016	2349	237
Evaluations 2013 and 2016(*)	**2143**	**237**

(*) The actual number of students evaluated is slightly larger (2174), however 31 students could not be found in the PAM data base. Note: the number of students is a little bit larger if it includes the students who participated in only one of the evaluations, either in mathematics or reading. The data given in this table corresponds to the students who took both of the tests.

The second set of data comprises indicators of PAM usage, gathered by the platform itself and referring to the years 2013, 2014, 2015 and 2016. In particular, we obtained the total number of exercises performed each year by the students who participated in the learning assessments[10]. There is also another indicator for the year 2016 regarding the participation of the student in the Competition in Mathematics and the number of exercises performed by the student during the competition. In this way it is possible to reconstruct a measure of the number of exercises performed outside of the competition, which is the variable used to approximate PAM usage in 2016.

Three alternative user definitions were employed: a user as a student who performed at least one, at least 10 or at least 20 exercises in a given year or as the (annual) average over a given period. The percentages of students falling into these categories are represented in Table 2. If we concentrate on the usage in 2016 outside of the competition, we find that 44% of the students performed at least one exercise on the platform, 39% performed 10 or more and one third of the students performed 20 or more exercises. No matter which definition of user we consider, there is verifiably strong growth in the usage of the platform over the period of interest. For example, the ratio of the students who performed 20 or more exercises grew by a factor of 3.5, increasing from 11% to 39%.

Table 2. Percentage of the student users of PAM who performed at least 1, 10 or 20 exercises per year for the given period. Sample of 2143 students from 237 schools

	% of users		
Year	1 or more exercises	10 or more exercises	20 or more exercises
2016 (outside the competition)	44.4	39.0	33.8
2016 (total number of exercises)	48.2	42.7	39.0
2015 (total number of exercises)	42.9	37.9	28.8
2014 (total number of exercises)	36.4	20.7	14.1
2013 (total number of exercises)	20.7	13.4	11.3
2014-2016 (total number of exercises)	67.5	51.8	42.7

Note: the percentage of users in 2014-2016 refers to the percentage of students who performed at least "N" exercises in the annual average

[10] For the year 2016, there is an indicator that identifies a student's participation or non-participation in the Competition in Mathematics as well as the number of exercises performed during the competition. This made it possible to reconstruct the number of exercises performed outside the competition. The latter was the variable finally used to approximate the educational use of PAM during the year 2016.

Table 3 shows some measures of the variable number of exercises performed by students on the platform. If we consider the year 2016 and the activity that took place outside of the competition, we find that a quarter of the users performed up to 22 exercises, half of them performed up to 84 exercises and 75% performed up to 274 exercises. The median of the exercises performed by the students who worked on PAM has increased over the years and especially in 2016.

Therefore, we can confirm that over the analysed period, the range of platform usage increased both in its extent (number of users) and in its intensity (number of exercises per student).

Table 3. Percentiles of the number of exercises performed by the students who used PAM in the given period. Sample of 2143 students from 237 schools

	Annual number of exercises (performed by students who did 1 or more exercises)			
Year	Percentile 25%	**Median**	Percentile 75%	Percentile 90%
2016 (outside the competition)	22	**84**	274	402
2016 (total number of exercises)	25	**102**	461	931
2015 (total number of exercises)	17	**39**	140	546
2014 (total number of exercises)	5	**12**	39	67
2013 (total number of exercises)	7	**24**	61	89
2014-2016 (total number of exercises)	11	**43**	167	329

Note: the number of exercises in 2014-2016 refers to the annual average of exercises over the three years

3. Empirical strategy

The identification of the causal effect of an intervention in the educational area poses several challenges. First is the need to have good measurements of the variables of interest, in particular of the progress the students have made in their own learning. The estimation of the effect of PAM represented in this study is based on reliable data. On one hand, the measurement of the student's learning progress originates from evaluations based on international standards and is a representative sample on national level of the cohort of all students who attended the 3rd year of primary school in 2013. On the other hand, the measurement of the usage of PAM relies on objective information, provided from the data base of the platform itself.

The second challenge is the strategy for identifying the causal effect itself, that is, the way in which the data is analyzed in order to find out whether there is a causal relationship between the usage of the educational tool and the learning outcomes. This strategy is particularly important in observational studies, since it is not possible to manipulate the treatment experimentally, which would allow for the inference of a causal effect through the simple comparison of the treatment group and the control group. On the contrary, PAM is a digital tool which was made available to all teachers and students in the educational system; however, the decision whether and how the tool would be used was left to them and possibly to the adults responsible for the child. Practically, considering the pupils who attended the 6th year in 2016, there are as many children who performed an exercise on the platform, as there are those who did not. Let us consider what would happen if PAM were used by the students who are most motivated or skilled, or whose usage is encouraged by the most effective teachers. By simply comparing the users (treatment group) and the non-users (control group), we would risk assigning a false causal effect to PAM. Expressed in statistical terms, we would obtain a biased estimation of the impact produced by the educational tool.

Our starting point is working with the gains in mathematics achieved by the students between 2013 and the end of 2016. Once we had defined the indicator of the usage of PAM from the years 2014 to 2016, we were interested in whether the learning gain for students using PAM was different from the gain for students who did not use PAM. Nevertheless, this comparison can be biased if there are other differences between the users and non-users which affect their performance in mathematics and have nothing to do with PAM. Our strategy is to compare the gains between users and non-users by “controlling” or “conditioning” for this type of variable.

The most complex problem however is the potential difference between users and non-users in unobservable dimensions. For example, the educational variables preceding the evaluation period which partially account for the student’s current performance, or dimensions which are difficult to measure such as the student’s motivation to learn, natural ability or the effectiveness of his or her teachers. What can we do to “control” such variables if we do not have any measures for them? In the first place, the possibility of having two consecutive measurements of learning progress for the same students is practically a necessity for evaluating the impact of a digital tool in the educational area. This allows us to use the initial scores as an indirect method for controlling the educational factors which had contributed to the student’s performance in the base year of the evaluation. In the second place, under the plausible assumption that PAM only affects performance in mathematics, we suggest using the gain in reading achieved by a student as a measure to summarize the unobservable variables for the student and the teacher which have influenced his or her general performance in both mathematics and reading.

In order to clarify this last idea, imagine the following situation. Let us suppose that the usage of PAM does not have any effect on the learning gains, but the most motivated students (who are motivated for other reasons than using of PAM) were more inclined to use it. In turn we know that motivation is an explanatory factor of the student's gain in both mathematics and reading. If we proceed and compare the learning gain of the students who use PAM with the gain of those who do not, we will find that the first group performed better. In this case we would erroneously assign PAM as the causal factor of that difference, when in reality this result is due to bias by omission of the underlying causal factor which is motivation. Now, consider a slightly different measurement: let us compare the difference between the gain in mathematics and the gain in reading for PAM users and non-users. In this last case we will not find an advantage in favor of the PAM users, since they [(the motivated students)] outperform non-users in both mathematics and reading and therefore both effects are compensated for in this new measurement. In this case we conclude correctly (according to our starting assumption) that the causal effect is zero. On the other hand, if we suppose that PAM has a positive impact on performance in mathematics and not in reading, it can be reasoned that the first measure (gains in mathematics) will lead to a positive result larger than the real effect (i.e. biased upwards), whereas the second measure (gain in mathematics minus gains in reading) will allow us to infer the true effect of PAM on learning in mathematics.

Taking the previous reasoning about the possible problems of comparing users and non-users to the extreme, we could argue that the differences between the students who used PAM and the rest consisted in attributes or abilities that are only related to mathematics. If this is the case, these differences cannot be controlled by the gain in reading and the estimations could still be positively biased. An alternative methodology is the application of an instrumental variable, in other words a variable which indirectly measures (is highly correlated to) the usage of PAM and which at the same time is not related to this type of unobservable variable. In this report we also present an analysis aligned with this methodological focus. We suggest measuring the usage of PAM by a student by means of the usage of PAM performed by his or her classmates. Although this can be seen as the application of an instrumental variable, in the current study we present it as a proxy for the teacher's usage of the platform[11]. Nevertheless, this instrumental variable will not correct the possible bias due to the presence of a specific type of unobservable characteristics of the teacher. For example, the teachers who encourage the usage of PAM may be the teachers who are on average most effective in teaching mathematics, but not necessarily the best in teaching reading[12]. Finally, even though our aim is to perform the most rigorous analysis given the available data, we cannot completely discard a subtle argumentation such as the one stated above that links special characteristics of teachers with the usage of PAM, which could bias our estimates.

In short, our identification strategy consists in comparing the gain in mathematics between users and non-users of PAM, controlling by a set of information summarized in the following variables: the test results obtained in 2013, the gain in reading, the socioeconomic level, the gender, the region (Montevideo/interior), the school type, the participation in the competition and the usage of PAM in

[11] The estimates that include a measure of group use correspond to the reduced form that directly links the gain in mathematics with the instrument. Strictly speaking, the estimation by instrumental variables is not reported, but its results are very similar in terms of significance and sign of the impact. The presentation of the method of instrumental variables exceeds the scope of this report.

[12] We would need a valid instrument for teacher usage.

2013[13]. From a methodological point of view, the analysis is performed by means of an approach by regression, estimating equations as shown in Appendix. In addition, different definitions of users depending on the number of performed exercises are used and then compared with non-users in order to find out if the resulting effect increases with increasing intensity of the usage of PAM.

A separate discussion is needed for the definition of PAM users. As mentioned above, we present results under the assumption that a PAM user is a student who has performed a given number of exercises per year (at least one, at least 10 or at least 20 exercises per year). Another aspect is whether we have to consider the use by the student or the use by the teacher. The available data corresponds to use by the student and shows an important intra-group correlation, which suggests that the decision of the teacher to incorporate PAM is a determining factor in the student's use. The discussion whether to include a measure of student usage or a measure of usage by a group of students (as a proxy for usage by the teacher) depends on the mechanism by which PAM can affect learning. This mechanism on the other hand depends on the way the teacher incorporates the tool into teaching. In the current study, we confine ourselves to presenting results based on measures of student usage and of mean usage by the class.

[13] It is important to note that PAM was introduced in 2013, which means in the base year of our evaluation. This is the reason why the usage in 2013 is included as a variable of an additional control in the estimations. An alternative is to exclude the students who used PAM in 2013 and perform the analysis with the rest of the students. This was done and it did not change the fundamental conclusions of this study.

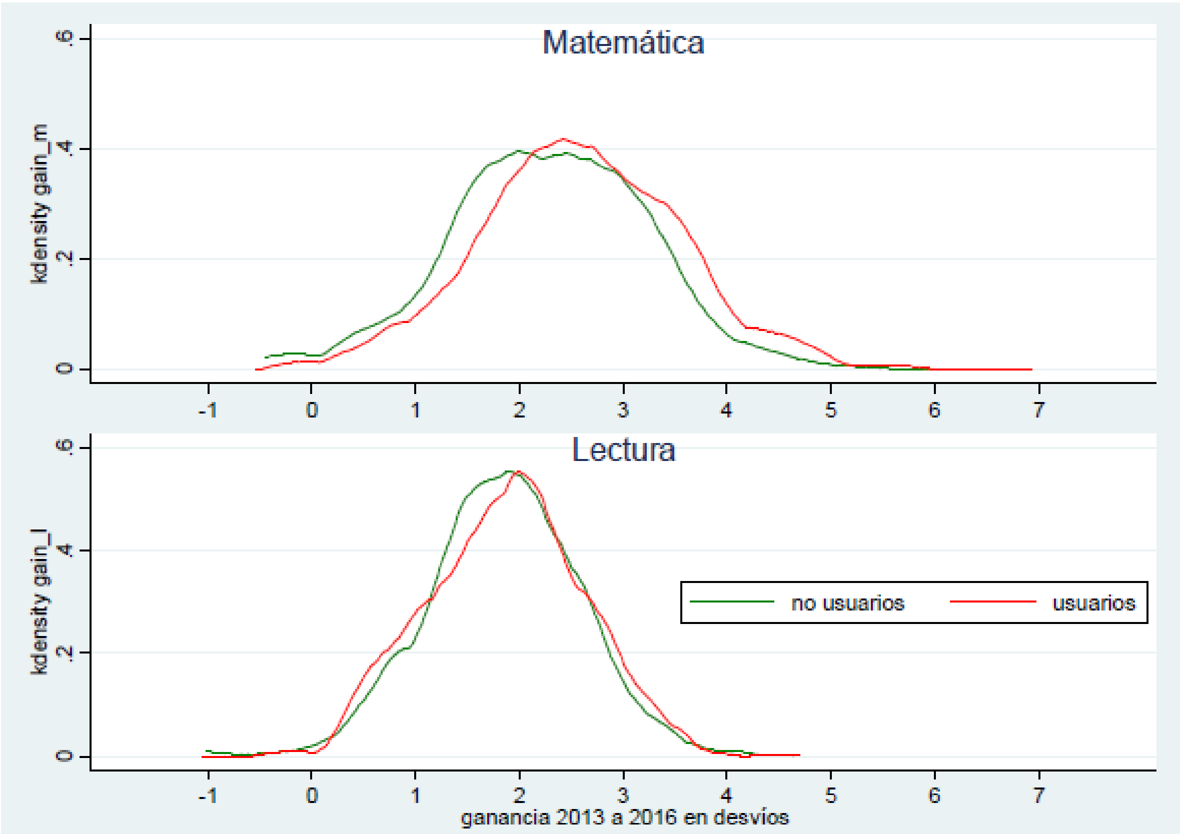
4. Results

All estimates result from standardized values of the test scores. Therefore, the gains in math and reading are expressed in standard deviations (SD). To have a reference value, the average gain in mathematics by students in the sample over their three years of schooling was approximately 2.6 SD, while the average gain in reading was 1.9 SD.

The diagram below shows the density function of the gain in mathematics (upper diagram) and of the gain in reading (lower diagram). In each case the function belonging to PAM-users (red line) is distinguished from the one belonging to non-users (green line). For this diagram the users are defined as the students who performed exercises between the years 2014 and 2016.

The density functions show the concentrations of students (vertical axis) for the different values the gain in score can assume (indicated on the horizontal axis). The graphs show that the condition of being a PAM-user is particularly perturbing for the gain in mathematics but not for the gain in reading. There is a clear difference in the distribution of the gain in mathematics between users and non-users. The density function for the students who used the platform is shifted to the right, that is, towards higher scores. This is not observed in reading.

Diagram 1. Distribution of the gain in mathematics (upper diagram) and of the gain in reading (lower diagram) comparing PAM-users (red line) and non-users (green line)



Note: Users are defined as students who performed exercises between 2014 and 2016

Now, let us compare the average gains between users and non-users in each of the subjects. That is, instead of seeing the distribution of the gains, let us concentrate on the median of each of those distributions and compare users and non-users. This comparison is a measure of the shift in the previously considered density functions. Furthermore, a statistical contrast is also performed to determine whether such a shift is statistically significant. All of this is presented in Table 4 for different definitions of a platform user.

What Table 4 shows is very clear. In the first place, the platform users show an average gain in mathematics that is consistently higher than the one obtained by non-users. Second, there are no significant differences between users and non-users with respect to the gain in reading (except for one of the user definitions where the difference is less than half the observed difference in mathematics).

So far, the data is consistent with the hypothesis that this educational tool is associated with learning mathematics but not with reading.

Table 4. Average gain 2013-2016 in Mathematics and Reading comparing PAM-users and non-users according to different user definitions

	Gain in Mathematics (standard deviations)			Gain in Reading (standard deviations)		
	User	Non-user	Difference	User	Non-user	Difference
Exercises in 2016						
1 or more	2.63	2.33	0.30***	1.91	1.86	0.05
10 or more	2.59	2.37	0.22***	1.94	1.85	0.09***
20 or more	2.57	2.40	0.17***	1.90	1.87	0.03
Exercises between 214 and 2016 (*)						
1 or more	2.56	2.24	0.32***	1.88	1.88	0.00
10 or more	2.60	2.31	0.30***	1.90	1.86	0.05
20 or more	2.56	2.38	0.18***	1.90	1.87	0.03

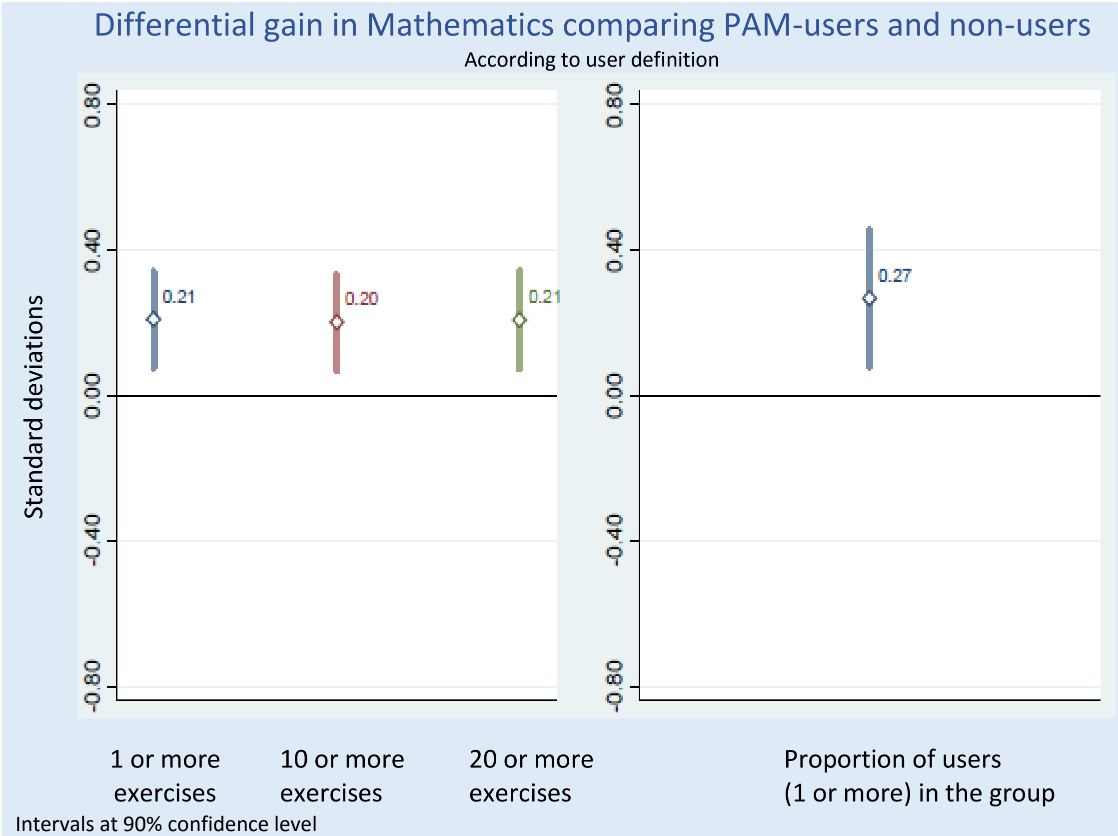
(*) The number of exercises between 2014 and 2016 indicates the annual average of exercises over those three years
Significance: * p <0.10, ** p <0.05, *** p <0.01

Below, we present a more ambitious estimate of the differences between users and non-users. This includes comparing the gain in mathematics as well as controlling the possible differences between the two groups by a set of observable variables (initial score, gender, socioeconomic level, region, type of school) as well as by the gain in reading (see the rationale for the latter in section 3). This is implemented by means of a regression model. The result is presented in Table 5. The first three columns report the advantage in mathematics obtained by students using the platform according to the three user definitions. The left chart of Diagram 2 illustrates these three estimates. The differences are significant and positive. The estimated coefficients indicate that PAM has an impact of 0.20 SD on the gain in mathematics.

Column 4 of Table 5 changes the user condition from a student's attribute to an attribute of his or her group. Methodologically, this involves changing the variable of interest in the regression model; the indicator variable of the user is replaced by a variable indicating the proportion of the students in the group who are users. This method for measuring the use of the platform can be interpreted as a measure of use on the part of the classroom teacher.

The coefficient of this variable is illustrated in the chart on the right-hand side in Diagram 2. It turns out to be significant and positive while showing that the higher the use at a group level, the greater the advantage in mathematics obtained by the students in the group[14]. This result allows us to presume the importance of using the tool in a group motivated by the teacher and to presume the existence of some type of external effect at the class level[15].

Diagram 2



A final analysis consists of contrasting the relationship between the intensity of platform use – measured as the number of exercises performed – and the gain in mathematics. The analysis is similar to the previous one with the difference that the users are grouped according to the number of exercises performed during the year (column 5 of Table 5 and Diagram 3). Similarly, we can define the intensity of group usage by defining four sections of the median number of exercises performed by the students in the group (column 6 of Table 5 and Diagram 4). It is important to state that the relationship between the number of exercises and the gain in mathematics is not a relationship that is clearly revealed in the data.

Let us take a look at the results regarding the intensity of group usage (Diagram 4). The coefficients of the 4 sections are positive and significant. A non-linear relationship between the number of

[14] The magnitude of the coefficient indicates that the gain from belonging to a group where 100% of the students use PAM is 0.27 SD with respect to a group where no one uses it.

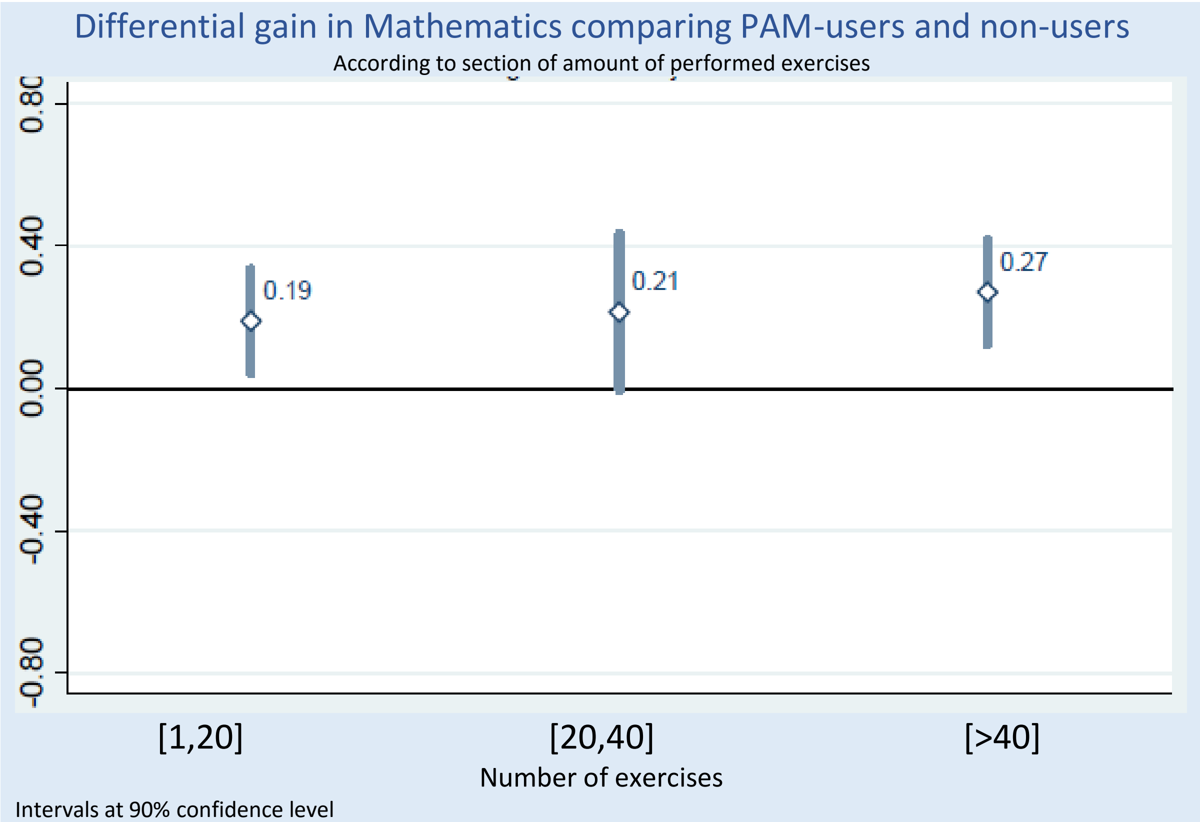
[15] Working with group data reduces the sample size because for some private schools in the sample this data is not available.

exercises and the gain in mathematics is indicated; the maximum impact is achieved by groups with a median number of exercises between 40 and 100. Table A4 from the Appendix shows an alternative estimation of the relationship between the number of exercises and the gain. The results are qualitatively similar and indicate the positive effect of the number of exercises.

In any case, identifying the relation between the number of exercises and the learning gain within the data is a complex task. It is certainly an unstable relationship because it depends on many factors such as the way in which the teacher includes the tool in his or her pedagogical approach or whether or not the teacher guides the student in the use of the platform[16]. Therefore, the above results should not be interpreted as an estimation of the optimal number of exercises.

The tables in the Appendix present additional estimations. The positive relation between the condition of being a platform user and the gain in mathematics is a robust result with respect to different specifications.

Diagram 3



[16] If this distinction is relevant, the existence of competitions generates a distortion in the measure of the intensity of usage. Concerning the data from 2016, although working with the number of exercises outside of the competition, it is possible that this measure may not completely correct the number of exercises induced by the competition. In fact, even with this corrected measure, we observe that the participants of the competition have on average a larger number of exercises than the rest.

Diagram 4

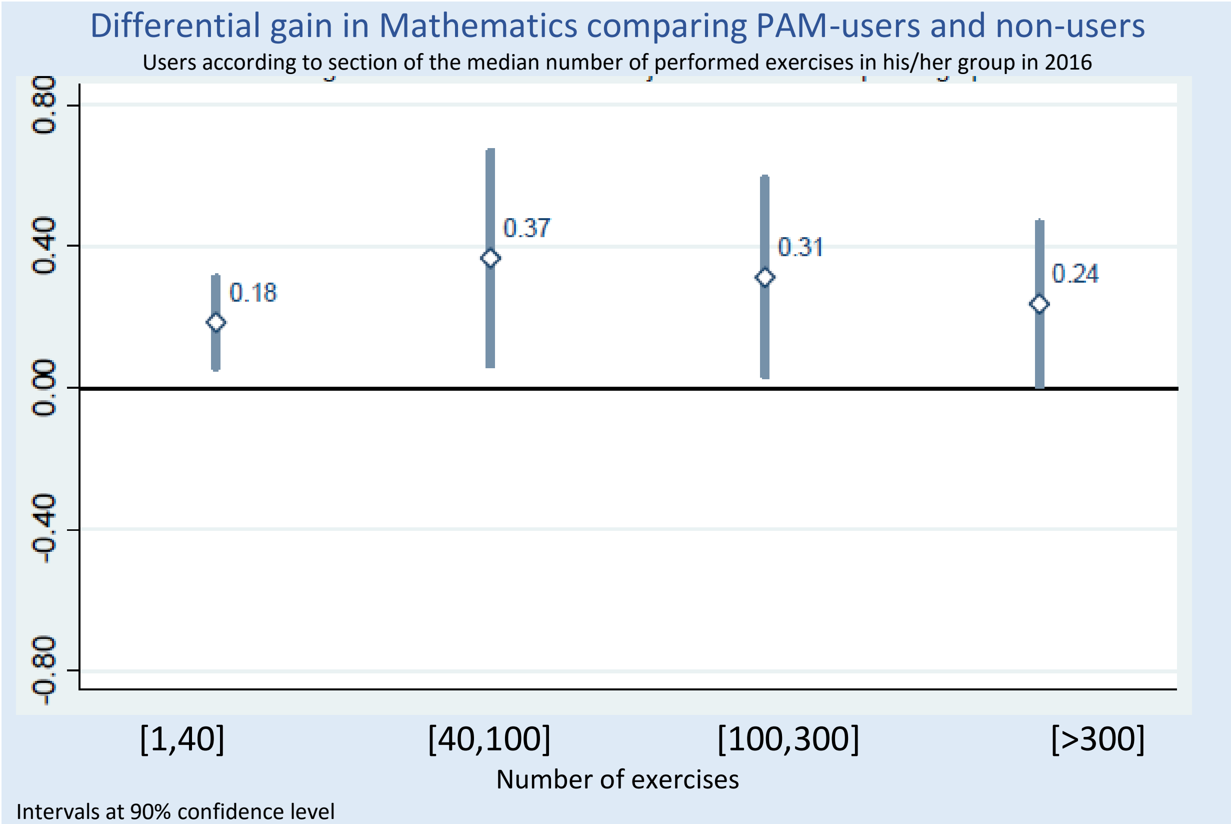

Differential gain in Mathematics comparing PAM-users and non-users
Users according to section of the median number of performed exercises in his/her group in 2016
0.80
0.40
0.00
-0.40
-0.80
0.18
0.37
0.31
0.24
[1,40]
[40,100]
[100,300]
[>300]
Number of exercises
Intervals at 90% confidence level

Table 5: Estimation of the PAM effect according to different user definitions and sections of amount of exercises

	(1) DMat	(2) DMat	(3) DMat	(4) DMat	(5) DMat	(6) DMat
[>1]	0.209**					
	(0.081)					
[>10]		0.201**				
		(0.080)				
[>20]			0.209**			
			(0.082)			
%[>1]grupo				0.267**		
				(0.113)		
[1,20]					0.189**	
					(0.089)	
[20,40]					0.214	
					(0.132)	
[>40]					0.270***	
					(0.089)	
g[1,40]						0.184**
						(0.075)
g[40,100]						0.365**
						(0.181)
g[100,300]						0.312*
						(0.167)
g[>100]						0.236*
						(0.136)
error_est	0.68	0.68	0.68	0.68	0.68	0.68
escuelas	237	237	237	210	237	210
alumnos	2143	2143	2143	1842	2143	1842

Dependent Variable: Gain in Mathematics
Control Variable: pts Mat and Lit 2013, Gain Reading, Usage in 2013, Gender, NSE, Mdeo, TC, Campeonato
Columns (5) and (6) include the interaction with the variable competition
Clustered standard errors per school in parentheses
Level of significance: *p<0.10, **p<0.05, ***p<0.01

[Translation note: escuelas – schools; alumnos – students; grupo – group; NSE: nivel socioeconómico – socioeconomic level; TC: tipo de colegio – type of school]

Columns (1), (2) and (3) consider a PAM-user identifier defined as students who performed at least 1, 10 and 20 exercises, respectively, in 2016 (outside the competition).
Column (4) considers the proportion of students in the group who performed exercises in 2016 (outside the competition).
Column (5) includes a set of dummy variables that identify three sections of the amount of exercises performed by the student in 2016 (outside the competition).
Column (6) includes a set of dummy variables that identify four sections of the median number of exercises of the group in 2016 (outside the competition).

5. Conclusions

This study was aimed at identifying the effect of using PAM on the gain in learning mathematics based on longitudinal data from a sample of primary school students.

The available data made it possible to apply an empirical strategy under less restrictive assumptions. For this, it was essential to be able to access longitudinal data about the learning, to observe an important variability in the usage of PAM by the students, and to have the possibility of using the gain in reading to control a possible bias of selection that is usually assumed in this type of analysis.

The results are robust with respect to the different specifications of the empirical model and with respect to the way a platform user is defined.

The use of PAM has a positive effect on learning gains. The estimated effect amounts to about 0.20 standard deviations of the learning gain in math.

A reference value of this result is the average progress achieved by the same students after three years of schooling which is estimated at 2.6 standard deviations.

An appropriate assessment of this result should balance this impact with the cost of the analyzed intervention, which is known as a cost-effectiveness analysis. Although this is not part of the present study, it is reasonable to assume that PAM is a highly cost-effective policy.

When analyzing the impact of the platform with a definition of usage at the group level, we observe that the higher the usage in the class, the greater the benefit the students obtain. This result allows us to assume the importance of using the tool for a group as induced by the teacher and to presume the existence of some type of externality at the class level.

It is important to note a few precautions concerning the scope of these results. One should not lose sight of the nature of the intervention being evaluated. PAM was made available to teachers and students with few guidelines as to how it should be used[17]. The effective implementation is relatively unknown and it may be presumed that there were different forms of usage. For example, the data available for this study does not distinguish between the use during class and use outside of the class schedule. The teachers who have incorporated the tool have made their own assessment of how and how much to use it and, consequently, could have altered other pedagogical actions. We have found that the impact of all this has been positive.

Therefore, it is not inferred that the usage *per se* has the caused the impact. The impact can reflect differences in content, pedagogy, learning time, etc. linked to the decision to use PAM.

The results show that PAM is an opportunity to improve the teaching of math. The ways in which the tool is used for pedagogical purposes vary and are relatively unknown. A deeper study of these practices may shed light on this point in the future.

Finally, it is worthwhile to point out, as Muralidharan et al. (2016) do, that the effectiveness of the incorporation of technology into the education should not be interpreted as a loss of importance of the teacher in education. On the contrary, when the technology is incorporated to perform routine tasks (such as skill classification) and intensive tasks in data analysis (such as identifying patterns in

[17] Yes, there have been incentives to use it by Plan Ceibal such as promotional and informative offerings, workshops, courses for teachers and competitions for students.

students´ responses and providing differentiated feedback and orientation), it can complement and enhance the teacher's efforts.

For example, teachers can dedicate more time to those aspects of education where they have a comparative advantage over technology, such as support for group learning strategies that can help develop social and other non-cognitive skills.

6. References

Duflo, E., Dupas, P., and M. Kremer. (2011).

Peer Effects, Teacher Incentives, and the Impact of
Tracking: Evidence from a Randomized Evaluation in Kenya. American Economic Review, 101: 1739-1744

Banerjee, Abhijit Vinayak, Shawn Cole, Esther Duflo, and Leigh Linden. 2007.

"Remedying Education: Evidence from Two Randomized Experiments in India." Quarterly Journal of Economics 122 (3): 1235-64.

Banerjee, Abhijit Vinayak, Rukmini Banerji, James Berry, Esther Duflo, Harini Kannan, Shobhini Mukerji, Marc Shotland, and Michael Walton. 2016.

"Mainstreaming an Effective Intervention: Evidence from Randomized Evaluations of 'Teaching at the Right Level' in India." CEPR Discussion Paper 11530, Center for Economic Policy Research, London.

Bulman, G. and R.W. Fairlie. 2016.

"Technology and Education: Computers, Software and the Internet," in Eric Hanushek, Stephen Machin, and Ludger Woessmann, eds., Handbook of the Economics of Education, Elsevier, pp. 239-280.

Linden, Leigh L. 2008.

"Complement or Substitute? The Effect of Technology on Student Achievement in India." Edited by Michael Trucano. InfoDev Working Paper 17, World Bank, Washington, DC.

Muralidharan, Karthik, Abhijeet Singh, and Alejandro J. Ganimian. 2016.

"Disrupting Education? Experimental Evidence on Technology-Aided Instruction in India." NBER Working Paper 22923, National Bureau of Economic Research, Cambridge, MA.

UNESCO-OREALC. 2016.

Technical Report. Third Regional Comparative Study and Explanatory, TERCE. Santiago, Chile.

7. Appendix

Formalization of the empirical model

The effect of the use of PAM is derived from the estimation of the following equation:

$$\boldsymbol{T^{m}_{ij(t)} - T^{m}_{ij(t-1)} = \delta\left(T^{l}_{ij(t)} - T^{l}_{ij(t-1)}\right) + (1-\rho^{m})T^{m}_{ij(t-1)} + \left(1-\rho^{l}\right)T^{l}_{ij(t-1)} + \beta PAM_{ij} + \gamma X_{ij} + \epsilon_{ij}} \quad \textbf{(1)}$$

$\boldsymbol{T^{a}_{ij(t)}}$ is the score in topic a (m = mathematics, l = language) obtained by child i from school j in the evaluation of year t (2013, 2016), therefore the term on the left-hand side is the gain in mathematics. PAM is a measure for the use of the platform for some year or some time period between $t-1$ and t, X is a vector of characteristics of the children and schools (gender, socioeconomic level, participation in the competition in 2016, all-day school, capital / interior) and $\boldsymbol{\epsilon_{ij}}$ is a random error that can be correlated between observations of the same school but is independent between observations of different schools (i.e. clustered errors per school).

A variation of the above equation, for which results are also presented in this report, is one where instead of controlling by the gain in reading, by inclusion of the term $\boldsymbol{\delta\left(T^{l}_{ij(t)} - T^{l}_{ij(t-1)}\right)}$ on the right-hand side of the equation, one directly specified an equation where the dependent variable (the left-hand side) is the difference between the gain in mathematics and the gain in reading: $\boldsymbol{(T^{m}_{ij(t)} - T^{m}_{ij(t-1)}) - \left(T^{l}_{ij(t)} - T^{l}_{ij(t-1)}\right)}$:

$$\boldsymbol{(T^{m}_{ij(t)} - T^{m}_{ij(t-1)}) - \left(T^{l}_{ij(t)} - T^{l}_{ij(t-1)}\right) = (1-\rho^{m})T^{m}_{ij(t-1)} + \left(1-\rho^{l}\right)T^{l}_{ij(t-1)} + \beta PAM_{ij} + \gamma X_{ij} + \epsilon_{ij}} \quad \textbf{(2)}$$

The columns 1 and 2 of the tables A1, A2 and A3 in this appendix show the results of such a specification. Although equation (2) may be more intuitive in aiming at explaining the differential gain between mathematics and reading, we believe it is more appropriate to work with equation (1) which involves controlling by the gain in reading but in a more flexible way.

In both equations, the estimate for the coefficient $\boldsymbol{\beta}$ is the measure of the mean effect of PAM on the gain in learning mathematics.

Table A1: Estimation of the "PAM-use" effect = having performed exercises between 2014 and 2016

	(1) DMat-DLec	(2) DMat-DLec	(3) DMat	(4) DMat
uso_1416	0.3242**	0.2164*	0.3194***	0.2496**
	(0.1289)	(0.1112)	(0.1013)	(0.1008)
error_est	1.19	0.82	0.94	0.68
escuelas	237	237	237	237
alumnos	2143	2143	2143	2143

Dependent variables in columns
(1) and (2): Gain Mathematics – Gain Reading
(3) and (4): Gain Mathematics
Control variables in columns:
(2): pts Mat and Lit 2013, Usage 2013, Gender, NSE, Mdeo, TC, Champ.
(4): pts Mat and Lit 2013, Gain Reading, Usage 2013, Gender, NSE, Mdeo, TC, Competition
Clustered standard errors per school in parentheses
Level of significance: *p<0.10, **p<0.05, ***p<0.01

Table A2: Estimation of the "PAM-use" effect = having performed exercises in each year 2016, 2015 and 2014

	(1) DMat-DLec	(2) DMat-DLec	(3) DMat	(4) DMat
uso_2016f	0.1744	0.1814**	0.2602*	0.2092**
	(0.1330)	(0.0704)	(0.1402)	(0.0809)
uso_2015	0.1384	-0.0239	0.0756	0.0017
	(0.1704)	(0.0908)	(0.1352)	(0.0849)
uso_2014	0.1064	0.1351**	0.0582	0.1035
	(0.1673)	(0.0685)	(0.1307)	(0.0758)
error_est	1.19	0.81	0.94	0.68
escuelas	237	237	237	237
alumnos	2143	2143	2143	2143

Dependent variables in columns
(1) and (2): Gain Mathematics – Gain Reading
(3) and (4): Gain Mathematics
Control variables in columns:
(2): pts Mat and Lit 2013, Usage 2013, Gender, NSE, Mdeo, TC, Champ.
(4): pts Mat and Lit 2013, Gain Reading, Usage 2013, Gender, NSE, Mdeo, TC, Competition
Clustered standard errors per school in parentheses
Level of significance: *p<0.10, **p<0.05, ***p<0.01

Table A3: Estimation of the "PAM-use" effect = proportion of the students in the group of 2016 who performed exercises in this year

	(1) DMat-DLec	(2) DMat-DLec	(3) DMat	(4) DMat
puso_~2016	0.0578	0.2089*	0.2383	0.2667**
	(0.2031)	(0.1120)	(0.1863)	(0.1133)
uso_2015	0.0367	0.0031	0.0187	0.0263
	(0.1912)	(0.1041)	(0.1545)	(0.0955)
uso_2014	0.0628	0.1538**	0.0138	0.1115
	(0.1853)	(0.0756)	(0.1510)	(0.0825)
error_est	1.19	0.81	0.95	0.68
escuelas	210	210	210	210
alumnos	1842	1842	1842	1842

The same notes as below Table A2.

Table A4: “Intensity of the PAM-use" = (log of 1 +) number of performed exercises

	(1) DMat	(2) DMat	(3) DMat
lej1	0.0411*		
	(0.0248)		
lej1_2016f		0.0518**	
		(0.0215)	
lej1_grupo			0.0437**
			(0.0211)
error_est	0.68	0.68	0.68
escuelas	237	237	210
alumnos	2143	2143	1842

Dependent variable: Gain Mathematics
Control variables: pts Mat and Lit 2013, Gain Reading, Usage 2013, Gender, NSE, Mdeo, TC, Competition
Clustered standard errors per school in parentheses
Level of significance: *p<0.10, **p<0.05, ***p<0.01

lej1 = logarithm of 1 + the annual average number of exercises between 2014 and 2016
lej1_2016f = logarithm of 1 + the number of exercises in 2016 (outside the competition)
lej1_grupo = log of 1 + the median number of exercises of the group 2016 (outside the competition)