JICA Ogata Sadako Research Institute for Peace and Development

JICA Ogata Research Institute Working Paper

Strengthening Teacher Support for Students to Improve Math Learning: Empirical Evidence on A Structured Pedagogy Program in El Salvador

Takao Maruyama

No. 222

July 2021

JICA Ogata Sadako Research Institute for Peace and Development

JICA Ogata Sadako Research Institute for Peace and Development

10-5 Ichigaya Honmura-cho

Shinjuku-ku

Tokyo 162-8433 JAPAN

TEL: +81-3-3269-3374

FAX: +81-3-3269-2054

Strengthening Teacher Support for Students to Improve Math Learning: Empirical Evidence on A Structured Pedagogy Program in El Salvador

Takao Maruyama*

Abstract

While recent debates on educational development focus on the learning crisis in primary education, the crisis in lower secondary education level is equally profound. Around 58 percent of school-age children worldwide enrolled in lower secondary education are not reaching the minimum proficiency level in mathematics. One of the approaches to improve student learning is a structured pedagogy program that provides schools with teaching and learning materials and other related interventions. The impact of teaching and learning materials on student learning depends upon the support of teachers for students. This study investigates the impact of additional components in a structured pedagogy program that tried to strengthen support of teachers to improve student math learning at the lower secondary level in El Salvador through a randomized controlled trial. The study tracked the same students for two years. While the average one-year impact of the additional component is estimated at around 0.18 standard deviations of test scores, the impact did not persist when the difference of interventions between the treatment and control groups disappeared in the second year of this research. Furthermore, a causal mediation analysis is conducted to investigate the possible causal path of the additional interventions on student math learning.

Keywords: Educational Development; Math learning; Lower secondary education, Latin America and Impact evaluation

---

* Graduate School of Humanities and Social Sciences, Hiroshima University, 1-5-1, Kagamiyama, Higashi-Hiroshima, Hiroshima 739-8529 Japan (tmaru@hiroshima-u.ac.jp)

I am grateful to the Ministry of Education in El Salvador for their understanding and intensive support to conduct this research. I appreciate Norihiro Nishikata, Kohei Nakayama, Eiichi Kimura, Satsuki Kawasumi, and Yuko Kawanami for their valuable support to conduct the field surveys, and Shin-ichi Ishihara, Hiromichi Morishita, Eiji Kozuka, Chie Esaki, Kazuro Shibuya, and Miyako Kobayashi for their continuous encouragement for this study. I am grateful to Takashi Kurosaki for his continuous guidance and support, Yutaka Arimoto, Yukichi Mano, Kensuke Teshima for their constructive comments to this study. I appreciate anonymous reviewers for their valuable comments. I also thank Sugashi Nagai, Koei Research & Consulting Inc., and his team members for their dedicated work on data management.

1. Introduction

Approximately 617 million primary and lower secondary school-age children worldwide are not reaching the minimum proficiency levels in reading and mathematics (UNESCO 2017). While recent debates on educational development focus on the learning crisis in primary education, the crisis in lower secondary education level is equally profound. The current status is far from the Sustainable Development Goals (SDGs) target of achieving at least minimum proficiency levels in reading and mathematics. Around 58 percent of school-age children are either finishing lower secondary education without acquiring the minimum proficiency or are out-of-school, and most of them are in low- or lower-middle-income countries (UNESCO 2017).

In Latin America, while enrollment in lower secondary education has expanded since the 2000s, the quality has stagnated over the years. Mathematics is a foundation of science, engineering, and technology, which are key drivers for a country's socio-economic development. Around a half of children at primary and lower secondary school-age are not mastering minimum proficiency level of mathematics in the region (UNESCO 2017). For example, results from a regional standardized assessment called Third Regional Comparative and Explanatory Study (TERCE) rank students in Chile highly (UNESCO, 2015); however, the country is below average among the participating countries in an international standardized assessment, the Trends in International Mathematics and Science Study (TIMSS) (Mullis et al. 2016). Through cross-country study, Hanushek and Woessman (2012) demonstrated that the low level in educational achievement was a source of poor growth performance in Latin America.

A recent systematic review of evidence in education suggests that a structured pedagogy program that provides teaching and learning materials and the other different types of interventions is an effective approach to improve student learning (Snilsveit et al. 2016). To address several challenges such as inadequately trained teachers and a lack of appropriate materials, curricula, and instructional approaches for improving learning, the structured pedagogy program includes different types of components like teacher training and lesson observations by school principals. Since little is known about the the effective combination of different components in the program, there is potential to further improve the effectiveness by integrating additional components that facilitate teachers to strengthen their support for students. One of the potential components is the assessment of student learning. In Argentina, diagnostic feedback of student test results for school principals and teachers improved student learning in mathematics and reading at lower secondary level. The feedback helped school principals manage the school to improve student learning, and teachers change their teaching practices in classrooms (de Hoyos et al. 2019). Another potential is strengthening study at home. While the quality of instruction in classrooms matters, study at home is also important for students to master what they have learned at school. Homework assignments are a traditional tool to enhance student learning at home (Cooper et al. 2006). However, simply including such components in a program might not work as intended, since there would be complementarities among different inputs in a package of interventions (Kerwin and Thornton 2020).

This experiment in El Salvador intends to bridge the gap in understanding surrounding the combination of different components in a structured pedagogy program. To improve student math achievement, the ministry of education implemented a project called "Project for the Improvement of Mathematics Teaching in Primary and Secondary Education" (hereinafter 'ESMATE project') with technical cooperation by Japan International Cooperation Agency (JICA). The ESMATE project developed a set of teaching and learning materials, i.e., teachers' guides, textbooks (hereinafter ESMATE textbooks) and workbooks,[1] then combined the provision

[1] Textbooks, workbooks and teachers' guidebooks developed by the project are posted at the following website of the Ministry of Education, Science and Technology in El Salvador. https://www.mined.gob.sv/materiales-educativos/item/1014902-esmate.html Appendix 1 of Maruyama and Kurosaki (2021) explains the page structure of textbook, and the relation among the different teaching and learning materials.

of the materials with the other types of interventions as a structured pedagogy program (hereinafter referred "ESMATE program"). While the ministry of education scaled up the ESMATE program nationwide for grades 7 through 9 (the lower secondary education level[2]), the ministry could not cover several parts of the program, including the distribution of student math workbook and diagnostic math test, because of the budget constraints. This study investigates the impact of the additional components in the ESMATE program on student math learning using a randomized controlled trial, and examine the mechanism by the causal mediation analysis.[3] The surveys tracked the same students for two years. During the first year, schools in the treatment group received the complete ESMATE program including the distribution of student math workbooks and diagnostic math tests, while a package of interventions that excluded these components was given to the control group. During the second year, there were no differences in the intervention between the two groups. Both groups received the ESMATE textbooks, teachers' guidebooks, and student math workbook.

This study makes three main contributions to the literature. First, this study examines the mechanism of the additional components in the structured pedagogy program through a review of the process data and a causal mediation analysis employing a method of Imai et al. (2010). The average causal mediation effect is the treatment effect that operates through a particular channel of a causal path. When there are several components in an intervention for the treatment group, it is hard to capture which component worked better through simple comparison with the control groups. But we can explore the mechanism using causal mediation analysis. In accordance with the additional components provided for the treatment group in year 1, the analysis takes three

[2] In El Salvador, the primary and lower secondary education (grades 1 through 9) are compulsory.

[3] The randomized controlled trial was conducted under an agreement between the Ministry of Education in El Salvador and JICA on June and October 2018. The survey and database construction were done by Koei Research & Consulting Inc., under a contract with JICA, in collaboration with the Ministry of Education in El Salvador. Hitotsubashi University Research Ethics Examination Committee reviewed the research plan. All the data used for this paper is provided by JICA.

possible mediators: students' time engaged in math learning in a lesson (called "student engaged time"); the frequency of homework assignments; and the provision of additional math class when the curriculum was delayed. Among the three mediators, only the average causal mediation effect of the provision of additional math class is positive and statistically significant.

Second, this study analyzes the heterogeneity of the impacts of the first-year impact with respect to student household economic status. In El Salvador, economic inequality is historically high. The Gini index was 54.5 in 1998 and gradually decreased with the modest economic growth of the country, reaching 38.6 in 2018 (World Bank 2021). Survey data of this study confirms that student math baseline test scores positively correlate with the household economic status. Based upon the baseline data on different types of student household asset, a composite index that represents the student household economic status by the principal component analysis was constructed. Even after controlling the heterogeneous impact by the baseline score level, the impact of additional components in the ESMATE program was smaller for students with higher household economic status.

Third, this study evaluates the accumulated impact of the first-year intervention on student math learning in the following year. Since students continue learning for years, it is critical to see whether students can advance their learning after receiving the intervention. This study tracked the same students for two years. In 2019, both groups received the same interventions by the ministry. The ministry continuously provided a set of math textbooks, teacher's guidebooks, and student workbooks. In year 1, the additional components in the ESMATE program improved student math learning by 0.18 standard deviations of test scores. Students in the treatment group responded correctly to the test items such as simple operation with positive and negative numbers, and direct and inverse proportions than the control group. But the average accumulated impact in the following year is estimated at nearly zero and not statistically significant. The gains in year 1 were not sufficient to yield an accumulated impact on math learning in year 2.

The remainder of this paper is organized into the following five sections. Section 2 describes the experimentation design and the contents of the ESMATE program. Section 3 reports the impacts of the additional components in the ESMATE program on student math learning. Section 4 investigates the mechanism of the additional components by the causal mediation analysis, and report the results. Then, section 5 discusses the findings, and section 6 concludes.

2. Experimentation design

(1) Contents of the ESMATE program

In 2018, the schools in the control group received a package of interventions to improve student math learning. The primary component of the package was a distribution of ESMATE textbooks and teachers' guidebooks. The ESMATE textbooks were designed to help students learn math through solving math problems. Each page of the textbooks includes the following four steps: (1) show the topic of the lesson; (2) pose example problems; (3) explain the general principle; and (4) provide exercises. Teachers organize a daily math lesson following the four steps. The package included the other components, namely, introductory training on the textbook for teachers and school principals, and mutual review meetings among teachers to improve teaching with the ESMATE textbook. Additionally, introductory training for the representatives of the parent association was also provided to strengthen family support for study at home.

For the treatment group, the ESMATE project additionally provided supplementary interventions to the abovementioned package, composed of the distribution of workbooks, development of annual math teaching plan at the introductory teacher training, initial on-site advice on lesson observations for school principals, and the distribution of math tests and mutual review meetings among teachers using the test results. The contents of the math workbook correspond to the ESMATE textbook. With the workbook, students can practice math problems that they learned in class. The annual teaching plan was a simple format for teachers to complete indicating which page in the ESMATE textbook will be taught on which date. Initial on-site advice on lesson observations for school principals aimed at improving the feedback for teachers being observed. Math tests were distributed to help teachers assess student math understanding. At the mutual review meetings of teachers organized inter-semesters, teachers review and discuss student test results with colleagues.

In January 2018, at the beginning of the school year, the ESMATE textbooks and teachers' guidebooks were distributed in the treatment group by the project. On the other hand, the distribution for the control group was delayed for roughly one month, because of a delay in the public procurement process by the ministry. In the control group, the schedule of introductory trainings for teachers, school principals, and representative of parents were also postponed for approximately one month in accordance with the delay of textbook distribution. In 2019 (year 2 of this research), the ministry provided a set of the teaching and learning materials including workbooks for both groups, and organized mutual review meeting of teachers.[4]

(2) Assessment of math learning

To assess their math learning, students completed written tests in all three rounds of surveys.[5] Survey teams administered the tests without the presence of teachers in the class room. To account for the progress following the curriculum, the test items differ across the three tests. The test items at the baseline survey were from math content in the primary education. The test items at the end-line were from math content in grade 7. The number of test items at the baseline and end-line surveys was respectively twenty in total, which included the problems posed in texts. The test

[4] The ESMATE project was finished at the end of June 2019. The ministry integrated the activities of the project into the plan and allocated necessary budget to continue the activities.

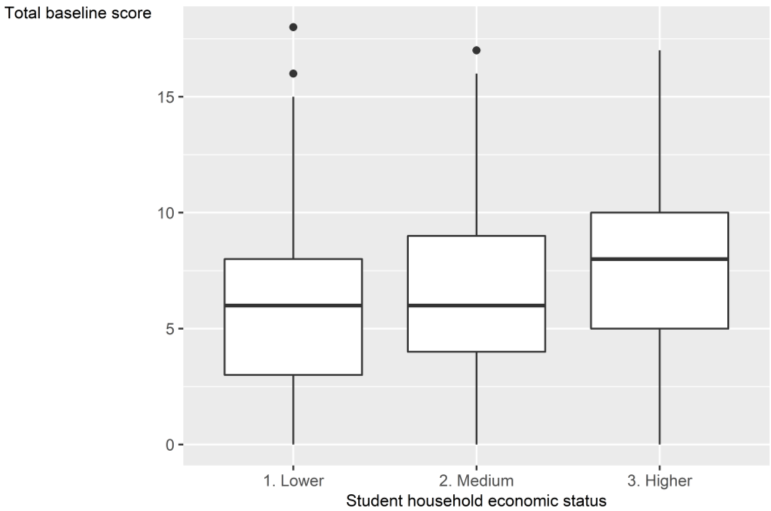
[5] The baseline survey was organized in January-March 2018, the end-line survey in September-October 2018, and the follow-up survey in September-October 2019.

items at the follow-up survey were mainly from math content in grade 8. The test at the follow-up survey was composed of 25 different items, of which five were the same from the end-line survey. The written test at each round of the surveys included the items to measure math learning by the cognitive skills (knowing, applying, and reasoning) and by the cognitive domains (number and operation, function, and geometry). The composition of test items is presented in Tables A-1 to A-3 in Appendix A.

The baseline test results demonstrate that the grade 7 students did not master the foundations of math including the basic four operations. For example, only 42 percent of the students were able to correctly answer the item "43 – 17". Responding to the item "2×2×2", 34 percent of students confused multiplication with addition, answering "6". On the division item "612÷102," half of the students did not write any response.

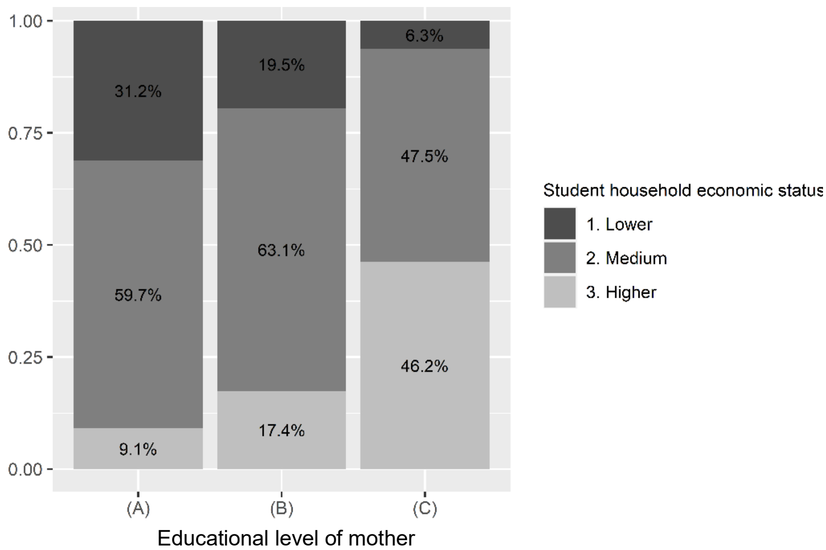
Through student interviews, the baseline survey collected information on different types of student household assets. A composite index of the different types of household assets was computed using the principal component analysis. The first principal component accounts for around 25 percent of the overall variance of the variables on different types of household assets, and all their loading coefficients are positive. Therefore, the first principal component was taken as the index of student household economic status. According to the first principal component, whole samples of students were divided into one of the three levels: lower; medium; and higher household economic status. The details of the principal component analysis are presented in Appendix B. As shown in Figure 1-1, student baseline test scores modestly correlate with household economic status. The follow-up survey additionally collected the information on parents' educational backgrounds through student interviews. Mothers' educational attainment positively correlates with household economic status (Figure 1-2), and mothers with higher educational attainment support student study at home more frequently in a week (Figure 1-3).

Figure 1-1: Boxplots of the baseline test scores by the student household economic status



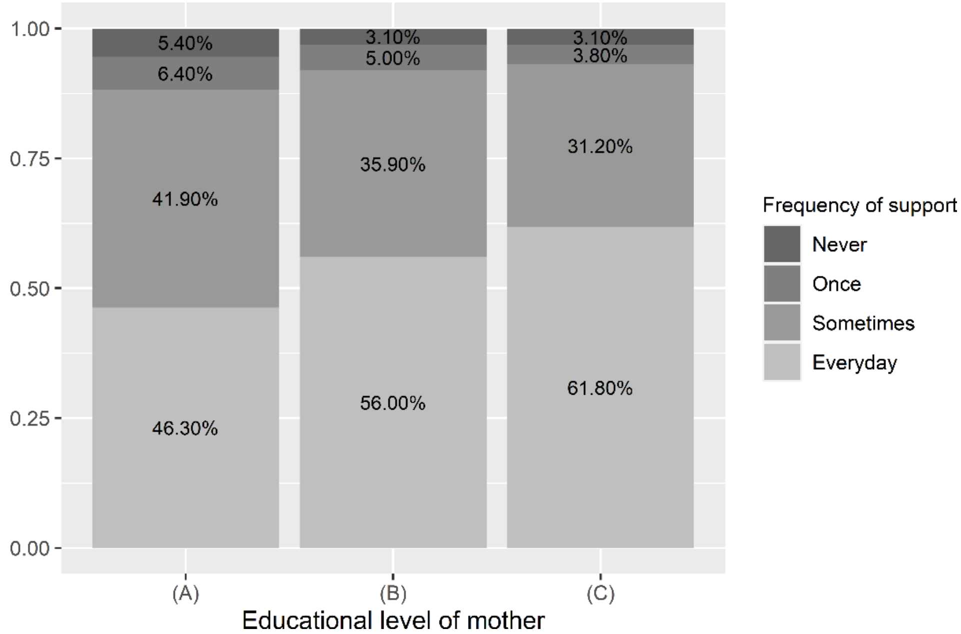
Note: Data source is the baseline survey of this research. N. of students in the lower household economic status is 534 (treatment group) and 488 (control group). N. of students in the medium household economic status is 1,317 (treatment group) and 1,279 (control group). N. of students in the higher household economic status is 472 (treatment group) and 437 (control group).

Figure 1-2: Student household economic status by mother's educational record



Note: Data source is the baseline and follow-up surveys of this research. Educational level of mother: (A) never go to school or primary (N.=583); (B) lower secondary (N.=843); (C) upper secondary or university (N.=861). The educational level of mother was surveyed through student interview at the follow-up survey. Among the surveyed students, 826 students answered that they did not know the educational record of mothers.

Figure 1-3: Frequency of family support for study at home in a week by mother's educational record



Note: Data source is the baseline and follow-up surveys of this research. Educational level of mother: (A) never go to school or primary (N. =583); (B) lower secondary (N.=843); (C) upper secondary or university (N.=861). The educational level of mother was surveyed through student interview at the follow-up survey.

(3) Sampling

This research originally targeted both the primary and lower secondary education levels, focusing on 2nd grade in primary education and 7th grade in lower secondary education. Considering the difference in the content of interventions and the educational level, this paper focuses on lower secondary education. Among 14 departments, Cabañas, La Union, San Miguel and San Vicente, situated in the central and eastern parts of the country, were selected. The educational outcomes such as enrollment and drop-out rates in those four departments are close to or below the national averages (Table 1).

Table 1: Educational statistics in the four departments

	National Average	Cabañas	La Union	San Miguel	San Vicente
Primary net enrollment rate (2015)	86.2%	89.0%	81.2%	85.7%	85.7%
Primary repetition rate (2014)	5.8%	6.7%	5.5%	5.4%	7.7%
Primary drop-out rate (2014)	6.4%	9.8%	8.5%	6.7%	7.7%
Secondary net enrollment rate (2015)	37.9%	25.4%	25.9%	35.5%	38.5%
Secondary repetition rate (2014)	4.9%	3.7%	4.9%	4.2%	4.3%
Secondary drop-out rate (2014)	8.5%	12.4%	11.5%	7.1%	8.0%

Source: The Ministry of Education in El Salvador

Public schools in the country are called "basic education public schools" and can include preschool, primary, and lower and upper secondary levels according to the local educational needs. Within the four departments, there were 612 basic education public schools offering lower secondary education. Since this research targets both primary and lower secondary education levels, 6 schools that did not have the primary education level were excluded. The country suffers from security problems due to the presence of gang members inherited from past civil conflicts. Intentional homicides per 100,000 were 61.8 in 2017 (World Bank 2021), the highest rate in the world. The schools were also affected by activities of gang members (USAID 2017). The schools located in an area severely affected by such activities, and those that were physically difficult to access were excluded from the sampling frame. Outside our experimental design, the Millennium Challenge Corporation (MCC) planned to distribute the ESMATE textbooks in 2018. The schools targeted by MCC were also excluded from our evaluation framework. As a result, the sampling frame was comprised of 369 basic education public schools. Of these, 250 basic education public schools were randomly sampled, with half randomly assigned to the treatment group and the other half to the control group (Table 2-1). Stratification variables in the randomization are department and urban status. If there were several classes of the targeted grades in the sampled school, one class was randomly selected. In the baseline survey, seven schools in the treatment group and four schools in the control group were excluded due to security reasons (Koei Research & Consulting Inc. 2018).[6] In addition to these eleven schools, there were no enrolled students in grade 7 at two schools in the control group.

Based on the educational census data collected by the ministry of education in the country, the characteristics of 612 basic education public schools in the four departments is compared to the sampling frame (Table 2-2). Because data from some schools are not included in the educational census survey data, the number of schools in Column (A) to (C) in Table 2-2 does not exactly match with that in Table 2-1. As shown in Table 2-2, the sampling frame represents well the original 612 basic education public schools in the four departments. The characteristics of the original sample of 250 schools and the remaining 237 schools after attrition of 13 schools were also equivalent between the two groups.

Table 2-1: Sampling frame of schools in the four departments

	Cabañas	La Union	San Miguel	San Vicente	Total
(A) N. of public schools (lower secondary)	105	146	248	113	612
(B) Schools with both cycle 1 and cycle 3 in (A)	104	144	247	111	606
(C-1) Schools without difficulty in access or security in (B)	64	68	164	105	401
(C-2) Schools not targeted by the MCC program (Sampling frame)	64	49	151	105	369
(D-1) Sampled schools (Treatment)	22	16	51	36	125
(D-2) Sampled schools (Control)	21	17	51	36	125

Source: Author.

6 At the end-line survey, three schools in the control group were additionally excluded because of security reasons (Koei Research & Consulting Inc., 2019). In the other two schools, the number of 7th grade students enrolled at the baseline survey was small, and all the students were absent at the end-line survey.

Table 2-2: Comparison of characteristics of schools in four departments (Cabanas, La Union, San Miguel, San Vicente)

Content	Public schools with cycle 3 (A)	Public schools with cycle 1 & 3 (B)	Sampling frame (C)	Sample (D)	Surveyed schools (E)	P-value (D)=(E)
Percentage of N. of schools in urban area	25.7	25.0	29.4	29.2	30.5	0.75
Average N. of students (grade 7) (both shifts)	26.7	26.9	28.6	29.5	30.0	0.88
Average N. of total students (grade 1 to 9) (both shifts)	214.6	216.4	227.7	233.7	237.3	0.88
Percentage of Male students in grade 7	51.6	51.6	51.5	50.7	50.8	0.95
Percentage of grade 7 students in morning shift	52.0	52.0	53.8	55.4	55.7	0.92
Percentage of grade 7 students repeated	4.5	4.4	4.4	4.5	4.6	0.91
Average age of grade 7 students	14.1	14.1	14.2	13.9	13.9	0.96
School infrastructure: Electricity	99.0	99.0	98.9	98.8	98.7	0.96
School infrastructure: Water	82.7	82.5	82.9	83.6	84.1	0.88
School infrastructure: Computer	77.4	77.2	80.7	80.0	80.3	0.93
School infrastructure: Internet	37.8	37.4	40.0	41.2	41.8	0.89
School infrastructure: Library	23.4	23.0	24.2	24.0	24.7	0.86
School infrastructure: Laboratory	9.2	8.6	9.2	9.6	9.6	0.99
School infrastructure: Kitchen	78.7	79.0	78.3	80.0	79.5	0.89
N. of schools	611	605	368	250	239	
N. of schools (which have grade 7)	606	601	366	248	237	

Note: 10% significance: *, 5% significance: **, 1% significance: ***.
(1) Data source is educational census survey data in El Salvador. Because the data of some schools are not available in the census survey data, the numbers of schools in Column (A) to (C) in this table do not exactly match with those in Table 2-1. (2) Values on school facilities are binary (Yes:1, No:0). (3) The p-values on number of students, percentage of students in morning shift, and percentage of students who repeated 2nd grade are the results of Wilcoxon rank sum test with stratified data (department dummy, urban/rural dummy). (4) The p-values on binary values show the results of chi-squares test with stratified data (department dummy, urban/rural dummy).

The student, teacher, and school characteristics of the remaining 237 schools are presented in Tables 3-1 to 3-3. The percentage of teachers who finished teacher training courses in the control group was higher than the treatment group by 10 percentage points (Table 3-2). The percentage of teachers teaching subjects other than mathematics was larger in in the treatment group than the control group by 9 percentage points (Table 3-2). In terms of the school characteristics, the number of 7th grade students in the treatment group is larger than the control group by 9 students (Table 3-3).[7] All the differences are statistically significant at the 10 percent level. While those several differences are noted in teacher and school characteristics, the overall characteristics of treatment and the control groups are well balanced, indicating successful randomization.

[7] The difference in the number of 7th grade students is a result of the number of classes of 7th grade in the sampled schools. Class size of the sampled class is equivalent between the treatment group (average: 24 students) and the control group (average 23 students).

Table 3-1: Comparison of characteristics of students (baseline survey)

Content	Treatment	Control	Mean Diff.	Adjusted Mean Diff.: (a)	Standard Error of (a)	P-value of (a)
Morning Shift (%)	58.30	63.29	-4.99	-3.92	5.39	0.46
Age	13.03	12.98	0.05	0.05	0.06	0.46
Sex (Male) (%)	48.80	50.09	-1.30	-1.47	1.78	0.40
N. elder brother/sister	1.76	1.79	-0.03	-0.03	0.08	0.66
N. younger brother/sister	1.13	1.11	0.02	0.02	0.04	0.70
Raw test score (Total points: 20)	6.58	6.63	-0.04	0.00	0.23	0.98
Standard deviations of raw test scores	3.55	3.57				
Asset of study						
Math textbook 2017 (%)	25.74	26.59	-0.85	-0.74	3.76	0.84
Math notebook 2017 (%)	90.40	87.70	2.70	2.18	2.58	0.39
Notebook only for Math 2017 (%)	89.24	85.57	3.67	3.18	2.76	0.25
Study desk at home (%)	95.01	95.37	7.12	-0.31	0.70	0.65
Asset of student household						
Smartphone (%)	86.48	85.39	1.09	0.89	1.64	0.58
Computer (%)	29.79	29.54	0.25	0.39	2.28	0.86
Refrigerator (%)	85.36	84.21	1.15	1.30	1.55	0.40
Car (%)	30.22	29.85	0.36	0.54	2.15	0.80
TV (%)	93.37	93.74	-0.37	-0.37	1.00	0.71
Tap water (%)	77.18	80.22	-3.03	-2.82	2.78	0.31
Electricity (%)	96.43	96.19	0.24	0.31	0.80	0.69
Flush Toilet (%)	52.35	55.63	-3.28	-2.86	2.79	0.30
Using wood for cooking (%)	61.08	62.43	-1.35	-1.22	3.19	0.70
Using gas for cooking (%)	91.78	92.42	-0.64	-0.63	1.17	0.58
Using electricity for cooking (%)	5.38	7.12	-1.74	-2.05	1.52	0.17
N. of schools	118	119				
N. of students	2324	2191				

Note:
(1) Data source is the baseline survey of this research. Adjusted mean difference in this table is obtained by regressing the value of each characteristics on treatment assignment dummy with controlling stratification variables (department and urban/rural dummies, and the interactions). Robust standard errors are clustered at school level. (2) Binary values are Textbook/Notebook (Yes:1, No:0), and Asset of Study and Family (Yes:1, No:0).

Table 3-2: Comparison of characteristics of teachers (baseline survey)

Content	Treatment	Control	Mean Diff.	P-value
Sex	0.53	0.55	-0.03	0.65
Age	39.34	39.94	-0.60	0.75
Total teaching period (years)	15.63	16.26	0.63	0.72
Academic Degree				
High school	0.03	0.02	0.01	0.64
Professorate	0.60	0.65	-0.05	0.47
Bachelor	0.35	0.32	0.03	0.65
Master	0.03	0.01	0.02	0.61
Doctor	0.00	0.00	0.00	0.95
Teacher qualification (1)				
Pedagogical Bachelor	0.03	0.03	0.00	0.99
Professor	0.75	0.85	-0.09 *	0.07
License in Education	0.19	0.14	0.04	0.37
Master's in Education	0.02	0.01	0.01	0.99
Doctorate in Education	0.00	0.00	0.00	0.95
Pedagogical Training Course	0.08	0.04	0.04	0.18
Teacher qualification (2)				
Basic Education Teacher (Cycle I and II)	0.15	0.15	0.00	0.98
Mathematics Specialty Teacher (Cycle III and High School)	0.60	0.59	0.01	0.83
Teacher specializing in other than math (Cycle III and High School)	0.18	0.23	-0.05	0.35
Class facility				
Board	0.98	1.00	-0.02	0.15
Teacher's desk	0.98	0.97	0.02	0.42
Teacher's chair	0.91	0.88	0.02	0.54
File cabinet / shelves	0.64	0.60	0.04	0.54
Working condition				
Teaching other subject	0.25	0.16	0.09 *	0.07
Only morning shift	0.42	0.50	-0.07	0.27
Only afternoon shift	0.52	0.45	0.07	0.27
Both shifts	0.06	0.06	0.00	0.99
N. of schools	118	119		

Note:
(1) Data source is the baseline survey of this research. 10% significance: *. (2) Values on class facilities are binary (Yes:1, No:0). (3) The p-values on age and total teaching period (years) is the result of Wilcoxon rank sum test with stratified data (department dummy, urban/rural dummy). (4) The p-values on binary values show the results of chi-squares test with stratified data (department dummy, urban/rural dummy).

Table 3-3: Comparison of characteristics of schools (baseline survey)

Content	Treatment	Control	Mean Diff.	P-value
Number of students				
N. of Student (7th grade) Morning Shift	39.75	30.58	9.17 *	0.06
N. of Student (7th grade) Afternoon Shift	20.87	20.65	0.22	0.91
N. of Student (Total)	257.59	238.83	18.76	0.99
Repetition and dropout rate				
Repetition rate (morning shift of 7th grade in 2017)	0.04	0.04	0.01	0.85
Repetition rate (afternoon shift of 7th grade in 2017)	0.04	0.05	-0.01	0.70
Dropout rate (morning shift of 7th grade in 2017)	0.09	0.09	0.00	0.66
Dropout rate (afternoon shift of 7th grade in 2017)	0.11	0.10	0.01	0.89
N. of teachers				
N. of vice school principal	1.10	1.14	-0.04	0.68
N. of teachers	10.59	10.19	0.40	0.61
School facility				
Electricity	1.00	1.00	0.00	0.95
Drinking Water	0.80	0.87	-0.07	0.16
Computer	0.95	0.93	0.02	0.59
Internet	0.43	0.48	-0.05	0.47
Internet use for students	0.67	0.70	-0.04	0.69
Library	0.25	0.25	0.00	0.97
Laboratory	0.12	0.07	0.05	0.17
Kitchen	0.78	0.81	-0.03	0.61
Canteen	0.81	0.73	0.07	0.18
Student support				
School lunch	0.97	0.97	0.00	0.99
Supplementary class (math)	0.23	0.16	0.07	0.18
Donor support within 5 years (except ESMATE)	0.92	0.94	-0.03	0.44
N. of schools	118	119		

Note:
(1) Data source is the baseline survey of this research.10% significance: *. (2) Values on school facilities are binary (Yes:1, No:0). (3) The p-values on number of students are the results of Wilcoxon rank sum test with stratified data (department dummy, urban/rural dummy). (4) The p-values on binary values show the results of chi-squares test with stratified data (department dummy, urban/rural dummy).

(4) Balance check of baseline scores

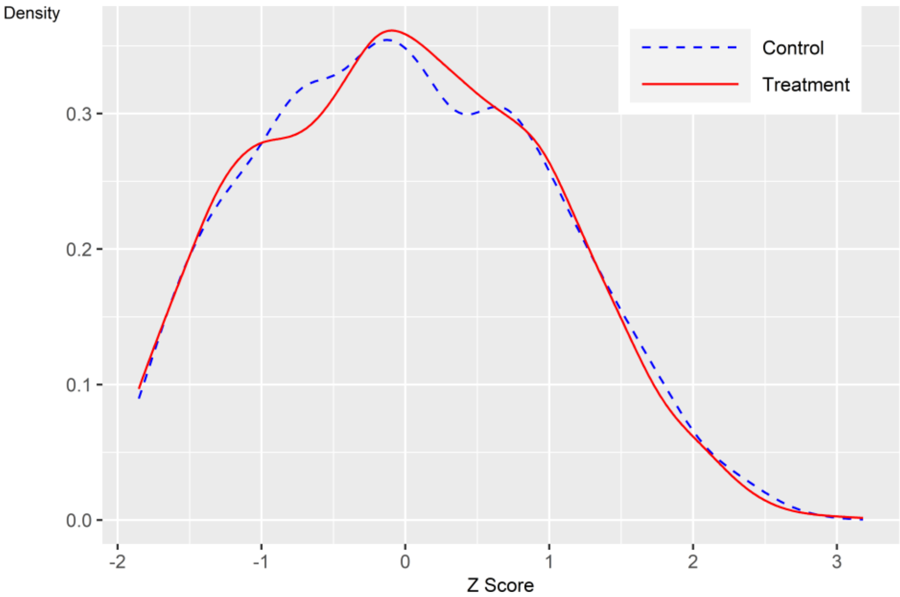
Because of logistical reasons, it was not possible to conduct the baseline survey before any component of the intervention package started. The baseline survey was started in mid-January 2018, just after the intervention in the treatment group began, and finished on March 1, 2018. The surveys of the treatment and control schools were conducted in parallel.

Density curves of the baseline scores are presented in Figure 2. The balance of baseline scores in the two groups is checked using the following equation:

$$Y_{ij0}=\alpha_0+\delta_0 Treatment_j+D_j\beta_{0D}+\varepsilon_{ij0} \ (1),$$

where Y_{ijk0} represents the math test baseline score for student *i* in school *j*. Test scores are standardized by the mean and standard deviations of the scores of students in the control group. *Treatment* is the treatment assignment. D_j is a vector of strata fixed effects constructed by department and urban status of school *j*. Robust standard errors are clustered at the school level. As presented in Table 4, δ_0 is nearly zero and not statistically significant, which indicates that the baseline scores are well balanced. All test items at the baseline survey were from the contents of primary education. Since the 7th grade ESMATE textbook does not include a review section of content learned in primary education and starts from new math content, it is plausible to think that the distribution of ESMATE textbooks at the beginning of the school year did not have an impact on the baseline scores in the treatment group.

Figure 2: Density curves of the baseline test scores (standardized scores)



Note: Data source is the baseline survey of this research.

Table 4: Comparison of the baseline test scores

	(1)
Treatment	-0.005
	(0.063)
Strata fixed effects	Yes
Adj. R^2	0.036
Num. obs.	4,515
N. Clusters	237

Note:
(1) Data source is the baseline survey of this research.
(2) Robust standard errors are clustered at the school level, and are in parentheses.
(3) The dependent variable is the standardized scores of the baseline test. Baseline test scores are standardized by the mean and standard deviations of baseline scores of the control group.
(4) Strata fixed effects are constructed by department and urban status.

(5) Student attrition

The same students were tracked through three rounds of surveys for two years. Among 4,515 students who were present at the baseline survey, 894 were absent at the end-line survey, and 1,370 were absent at the follow-up survey. The attrition rate was 19.4 percent at the end-line survey and 29.1 percent at the follow-up survey in the treatment group. In the control group, the attrition rate was 20.3 percent at the end-line survey, and 31.6 percent at the follow-up survey. The major reasons behind the attrition were students moving to the other schools or dropping out. At the end-line survey, 11 percent of students in the treatment group and 9 percent in the control group moved to another school or dropped out. In the following year, 12 percent in the treatment group and 11 percent in the control group moved to another school or dropped out.

To check whether differential attrition occurred between the two groups, the student attrition dummy is regressed on the treatment assignment, student characteristics (baseline), and strata fixed effects constructed by department and urban status. The results are shown in Table C in Appendix C. The attrition was not differential between the two groups. The attrition occurred more among boys, students with lower baseline scores, and older students. There was also higher student attrition among those who have more younger siblings.

Grade 7 students who repeated the same grade in year 2 of this research remained in the sample and were assigned the same math test taken by the grade 8 students. In the follow-up survey data, 1.9 percent of students in the treatment group and 2.3 percent of students in the control group repeated 7th grade.

3. Impacts on student math learning

(1) Estimation strategy

The impacts of the additional components in the ESMATE program on student math learning in year 1 and year 2 are estimated using following equation.

$$Y_{ijr}=\alpha_r+\gamma_r Y_{ij0}+\delta_r \text{Treatment}_j+ C_{ij}\beta_{rc}+ P_{kj0}\beta_{rp}+ S_j\beta_{rs}+ D_j\beta_{rD}+\varepsilon_{ijr} \quad (2)$$

In equation (2), Y_{ijr} represents the math test score for student *i* in school *j* at round *r* of the survey (r=0: baseline, r=1: end-line (year 1), and r=2: follow-up (year 2)). Test scores are standardized by the mean and standard deviations of test scores of students in the control group at each round of the survey. Sub-totals of test scores by the cognitive skills and cognitive domains are also used. C_{ij} is a vector of characteristics of student *i* at school *j* such as age, gender, shift at school (morning or afternoon), and the number of brothers and sisters, and characteristics of the family of student *i* such as the number of household asset types at the baseline. P_{mj0} is a vector of characteristics of teacher *k* at school *j*, who teaches mathematics to student *i* in the year 1, such as age, gender, and educational qualification at the baseline. S_j is a vector of characteristics of school *j* such as the number of students, school infrastructure and school meal, and characteristics of the school principal. D_j is a vector of strata fixed effects constructed by department and urban status of school *j*. Robust standard errors are clustered at the school level.

Equation (2) is not estimated using three-period panel data, but instead using either cross-section data composed of the control variables of r=0 and the dependent variable of r=1 for identifying the treatment effect in year 1, or cross-section data composed of the control variables of r=0 and the dependent variable of r=2 for identifying the accumulated treatment effect of year 1 in the following year. Both cross-section data are constructed from the balanced panel data either of r=0 and r=1 or r=0 and r=2. Then, the parameter δ_r in equation (2) is a scalar, not a vector of parameters.

To analyze the heterogeneous impacts by the baseline scores and the student household economic status, equation (2) is expanded by adding the interaction term of the baseline score or student household economic status index levels and the treatment assignment as

$$Y_{ijr}=\alpha_r+\gamma_r Y_{ij0}+(\delta_{rA}+\delta_{rB}X_{ij0})Treatment_j+C_{ij}\beta_{rc}+P_{kj0}\beta_{rp}+S_j\beta_{rs}+D_j\beta_{rD}+X_{ij0}\beta_{rx}+\varepsilon_{ijr} \quad (3).$$

In equation (3), X_{ij0} stands for either a vector of the baseline scores or student household economic status level dummies (lower, medium, or higher status).

(2) The one-year impact of the additional components in the ESMATE program

The regression results from equation (2) applied to the cross-section data composed of the independent variables of r=0 and dependent variable of r=1 are shown in Table 5-1. A portion of the teachers in charge of math class at the time of the baseline survey were changed during the 2018 school year. Thus, the teacher characteristics at the beginning and the end of school year were controlled respectively. The average impact of additional components in the ESMATE program is estimated at 0.18 standard deviations of test scores, which is statistically significant at the 1 percent level (Column (3) and (4) in Table 5-1).[8] Density curves of the end-line test scores are presented in Figure 3.

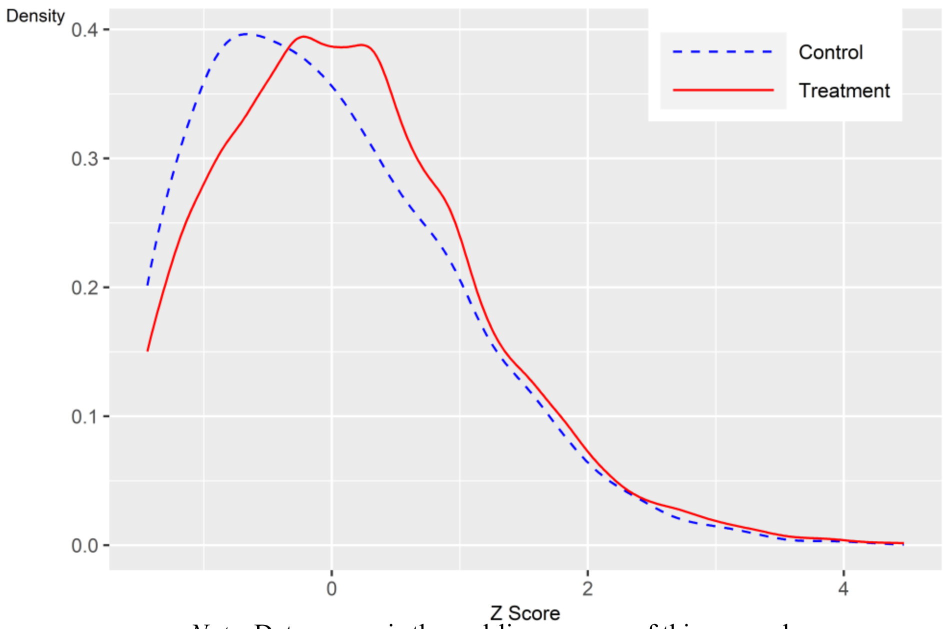
[8] I conducted the cost-effective analysis, following the methodology presented by J-PAL (Bhula, R. et al. 2020; Dhaliwal et al. 2014). The cost-effectiveness is measured as the ratio of the aggregated impact of the project (the average impact on student learning per student multiplied by the number of students impacted) to the aggregated cost of implementing the project. The cost-effectiveness is presented as the total standard deviations gained across the sample per 100 USD spent. The cost-effectiveness of the additional interventions of the ESMATE program, the total standard deviations gained across the sample per 100 USD spent, is estimated to be 2.80. The level of cost-effectiveness of the additional interventions in the ESMATE program is comparable to the other programs cited in Kremer et al. (2013).

Table 5-1: Average one-year impact of the additional components in the ESMATE program

	(1)	(2)	(3)	(4)
Treatment	0.164**	0.165***	0.184***	0.183***
	(0.077)	(0.061)	(0.058)	(0.057)
Z score baseline		0.515***	0.479***	0.472***
		(0.022)	(0.021)	(0.021)
Strata fixed effects	Yes	Yes	Yes	Yes
Student characteristics	No	No	Yes	Yes
School characteristics	No	No	Yes	Yes
Teacher characteristics (baseline)	No	No	Yes	No
Teacher characteristics (end-line)	No	No	No	Yes
Adj. R^2	0.030	0.281	0.330	0.332
Num. obs.	3,621	3,621	3,619	3,619
N. Clusters	232	232	232	232

**p < 0.05; *p < 0.1
(1) Data source are the baseline and end-line surveys of this research.
(2) Robust standard errors are clustered at the school level, and are in parenthesis.
(3) Student test scores are standardized by the mean and standard deviations of test scores in the control group at each round of survey.
(4) Strata fixed effects are constructed by department and urban status.

Figure 3: Density curves of the end-line test scores (standardized scores)



Note: Data source is the end-line surveys of this research.

The impacts across the cognitive domains and cognitive skills are estimated. While the impacts in the domains of "number and operation" and "function" are positive and statistically significant, the impact in the domain of geometry is not statistically significant (Table 5-2). In the former two domains, students in the treatment group correctly responded to the items such as simple operations with positive and negative numbers, and direct and inverse proportions than the control group.

Looking at the cognitive skills, while the impact on the knowing skill is statistically significant, the impact on the applying and reasoning skills is not. For example, the item on the knowing skill (No. 14 in the end-line test) asked students to fill a blank in a table that shows the relation between the number of units and the total price given the unit price (correct response rate: 49 percent in the treatment group and 40.8 percent in the control group). On the other hand, for the item (No. 16 in the end-line test) on the applying skill, which showed a simple graph of the linear function and asked students the slope coefficient, there was no a difference in the correct response rates (correct response rate: 5 percent in the treatment group and 3.9 percent in the control group).

Table 5-2: Average one-year impact of the additional components in the ESMATE program by the cognitive domains and the cognitive skills

	(1)-1 Number. & Operation	(1)-2 Function	(1)-3 Geometry	(2)-1 Knowing	(2)-1 Applying & reasoning
Treatment	0.169***	0.182***	0.079	0.187***	0.065
	(0.057)	(0.054)	(0.063)	(0.055)	(0.059)
Z score baseline	0.426***	0.332***	0.263***	0.464***	0.275***
	(0.020)	(0.019)	(0.022)	(0.020)	(0.022)
Adj. R^2	0.282	0.187	0.129	0.326	0.132
Num. obs.	3,619	3,619	3,619	3,619	3,619
N. Clusters	232	232	232	232	232

***$p < 0.01$
(1) Data source are the baseline and end-line surveys of this research.
(2) Robust standard errors are clustered at the school level, and are in parenthesis.
(3) Student test scores are standardized by the mean and standard deviations of test scores in the control group at each round of survey.
(4) Student, teacher, school characteristics, and strata fixed effects constructed by department and urban status are controlled in all the regressions but not shown.

As noted in the section 2. (1), the distribution of ESMATE textbooks and teachers' guidebooks was delayed by around one month in the control group. While the curriculum covered in the 2018 school year was limited in the control group due to the delay, approximately the same percentage of schools completed the units No. 1 through 4 in the curriculum.[9] The impact of the additional components in the ESMATE program are estimated, taking the sub-total of the end-line test scores corresponding to the unit 1 through 4, for the robustness check. The results are shown in Table 5-3. The estimated impacts remain at almost the same level.

Table 5-3: Robustness check of the one-year impact of the additional components in the ESMATE program

	(1)	**(2)**	**(3)**	**(4)**
Treatment	0.131*	0.132**	0.178***	0.178***
	(0.070)	(0.061)	(0.061)	(0.060)
Z score baseline		0.402***	0.370***	0.363***
		(0.022)	(0.022)	(0.022)
Strata fixed effects	Yes	Yes	Yes	Yes
Student Characteristics	No	No	Yes	Yes
School Characteristics	No	No	Yes	Yes
Teacher Characteristics (baseline)	No	No	Yes	No
Teacher Characteristics (end-line)	No	No	No	Yes
Adj. R^2	0.029	0.185	0.238	0.237
Num. obs.	3,621	3,621	3,619	3,619
N. Clusters	232	232	232	232

***p < 0.01; **p < 0.05; *p < 0.1
(1) Data source are the baseline and end-line surveys of this research.
(2) Robust standard errors are clustered at the school level, and are in parenthesis.
(3) Dependent variable is the sub-total of the end-line test from the item No. 1 through No. 9 that corresponds to the unit 1 through 4 in math curriculum.
(4) Student test scores are standardized by the mean and standard deviations of test scores in the control group at each round of the survey.

(3) Heterogeneity in the one-year impact of the additional components

The heterogeneity of the average treatment effect by the baseline scores and the household economic status is investigated by applying equations (3). The results are presented in Table 6.

9 In both groups, around 90 percent of schools finished the unit No. 1 through 4 among eight units in the curriculum at the end-line survey.

As shown in Column (1), the net impact for higher economic status is close to zero. Since the impacts for students with the lower and median economic status are positive and statistically significant, the additional components reduced the disparity in learning outcomes by household economic status. On the other hand, the impact of the additional treatment is not heterogeneous by the baseline test scores (Column (2) in Table 6).

Table 6: Heterogeneity in the impact of the additional components in the ESMATE program

	(1)	**(2)**	**(3)**	**(4)**
(Intercept)	-0.013	0.013	-0.015	-0.037
	(0.461)	(0.459)	(0.461)	(0.462)
Treatment	0.228***	0.183***	0.229**	0.186***
	(0.062)	(0.058)	(0.063)	(0.058)
Treatment × Lower economic status	-0.018		-0.017	
	(0.069)		(0.069)	
Treatment × Higher economic status	-0.207***		-0.208***	
	(0.068)		(0.068)	
Treatment				-0.041*
× Student economic status index				(0.021)
Treatment × Z score baseline		0.000	0.048	0.007
		(0.041)	(0.071)	(0.040)
Treatment × Z score baseline				-0.019
× Student economic status index				(0.016)
Lower economic status	-0.126***	-0.137***	-0.127***	
	(0.044)	(0.037)	(0.044)	
Higher economic status	0.173***	0.068**	0.173***	
	(0.048)	(0.038)	(0.048)	
Student economic status index				0.064***
				(0.013)
Z score baseline	0.478***	0.477***	0.474***	0.473***
	(0.021)	(0.028)	(0.028)	(0.028)
Adj. R^2	0.333	0.332	0.333	0.333
Num. obs.	3,619	3,619	3,619	3,619
N. Clusters	232	232	232	232

***p < 0.01; **p < 0.05; *p < 0.1
(1) Data source are the baseline and end-line surveys of this research.
(2) Robust standard errors are clustered at the school level, and are in parenthesis.
(3) Student test scores are standardized by the mean and standard deviations of test scores in the control group at each round of survey.
(4) Student, school, teacher (end-line) characteristics, and strata fixed effects constructed by department and urban status are controlled but not shown.
(5) In the column (1) to (3), according to the student household economic status index, student samples are divided to lower; medium; and higher household economic status. The student economic status index is the first principal component. Refer to Appendix B for the details.

Since the baseline test scores and the household economic status level are correlated with each other, both interaction terms are included in equation (3). As shown in Column (3) in Table 6, the estimated values of heterogeneous impact remain at almost the same level. Then, the triple interaction term of the treatment assignment, student household economic status index,[10] and baseline test scores are added in the equation. The coefficient of the triple interaction term is close to zero and not statistically significant.

(4) The accumulated impact of the additional components in the following year

In year 1, the additional components in the ESMATE program improved student math learning by 0.18 standard deviations of test scores. The accumulated impact of the first-year intervention is investigated by applying equation (2) to the cross-section data composed of the control variables of r=0 and dependent variable of r=2.[11] In the introductory training for school principals, they were advised to continuously assign the same teacher for two years to the surveyed students. Although similar guidance was given to the school principals in the control group, those in the treatment group followed the guidance better. The rate of assignment of the same teacher to the surveyed students in the treatment group was 85.9 percent in year 2, which was higher than in the control group (75.9 percent). Thus, teacher assignment in year 2 correlates with the treatment assignment, which can attenuate the estimated impact when simply controlling the teacher characteristics in both years. Then the teacher characteristics in year 1 are controlled at first and compared the result with the estimated value controlling teacher characteristics in year 2. Both

[10] Instead of student household economic status levels (lower; medium; and higher), the student household economic status index (first principal component from the principal component analysis on the different types of household assets) is used. Please refer to Appendix C for the details on the principal component analysis.

[11] Several variables of student characteristics are added to the control variables in year 2, such as the school shift, the availability of ESMATE textbooks, and whether the student repeated the same grade.

estimates are nearly zero and not statistically significant (Column (3) and (4) in Table 7).[12] [13]

Density curves of the follow-up test scores are presented in Figure 4.

Table 7: Accumulated impact of the first-year intervention (the additional components in the ESMATE program) in the following year

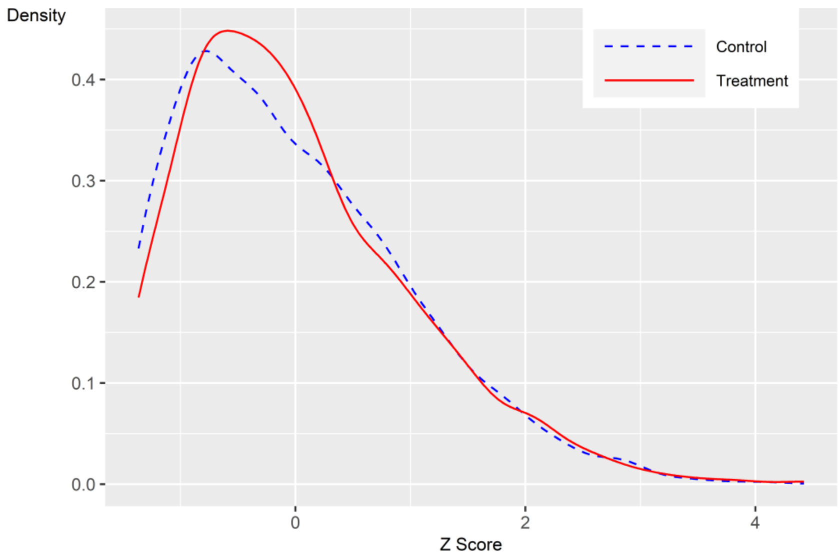
	(1)	**(2)**	**(3)**	**(4)**
Treatment	0.037	0.054	0.017	0.024
	(0.076)	(0.061)	(0.059)	(0.058)
Z score baseline		0.475***	0.438***	0.434***
		(0.021)	(0.022)	(0.021)
Strata fixed effects	Yes	Yes	Yes	Yes
Student characteristics	No	No	Yes	Yes
Teacher characteristics (2018)	No	No	Yes	No
Teacher characteristics (2019)	No	No	No	Yes
School characteristics	No	No	Yes	Yes
Adj. R^2	0.020	0.237	0.294	0.292
Num. obs.	3,143	3,143	3,117	3,117
N. Clusters	229	229	228	228

***$p < 0.01$
(1) Data source are the baseline, end-line and follow-up surveys of this research.
(2) Robust standard errors are clustered at the school level, and are in parenthesis.
(3) Student test scores are standardized by the mean and standard deviations of test scores in the control group at each round of survey.
(4) Student, teacher, and school characteristics, and the strata fixed effects constructed by department and urban status are controlled in all the regressions but not shown.

[12] As noted in the section 2, the percentage of students who repeated grade 7 in year 2 of this research was 1.9 percent in the treatment group, and 2.3 percent in the control group. In order to check whether the estimated impact was influenced by those who repeated the same grade, the regression analysis is conducted with the sub-sample that excluded those students. The results are almost at the same level with Table 7. The point estimate is 0.013 (standard error: 0.059; and p-value: 0.823).

[13] For the robustness check of the estimate of accumulated impact, the regression analysis is conducted, controlling both teacher characteristics in year 1 and 2. The estimated value remains at the same level (0.024 standard deviations (standard error: 0.058), which is not statistically significant).

Figure 4: Density curves of the follow-up test scores (standardized scores)



Note: Data source is the follow-up surveys of this research.

To examine the accumulated impact from different angle, the impact is estimated across several categories of math content: (a) content learned in grade 7 (item No. 1 through 5 of the math test at the follow-up survey); (b) content learned in grade 8 using knowledge and skills introduced in grade 7 (No. 6 through 10 and No. 12 through 15); (c) new content in grade 8 (No. 17 through 20 and 22 through 25); and (d) applying and reasoning on content learned in the grade 8 (No. 11, 16 and 21).[14] All the estimated values, including the value of the accumulated impact on the content learned in grade 7, are nearly zero and not statistically significant.[15]

[14] The content of math items is described in Table A-3 in Appendix A.

[15] The estimated values of the impact on each category of items are as follows; (a) -0.018 (standard error: 0.061); (b) 0.015 (standard error: 0.057); (c) 0.062 (standard error: 0.064); and (d) -0.063 (standard error: 0.047).

4. Causal mediation analysis of the first-year intervention

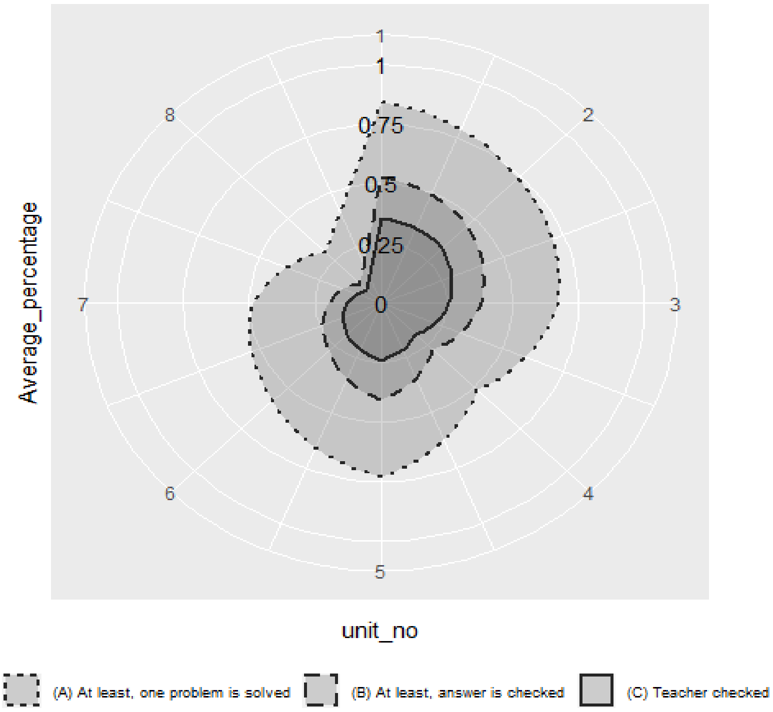
(1) Intermediate outcomes of the first-year intervention

In year 1 of this research, the treatment group received the additional interventions in the ESMATE program, composed of (a) the distribution of workbooks, (b) an annual math teaching plan developed at the introductory teacher training, (c) initial on-site advice for school principals in initial lesson observation, and (d) the distribution of math tests and a mutual review meeting with the student math test results. As presented in section 3. (1), the additional components improved student math learning by 0.18 standard deviations of test scores. This section analyzes the mechanism of the additional components in year 1, beginning with a review of the intermediate outcomes on the teaching practices. The details of the analysis of intermediate outcomes are presented in Appendix D of this paper.

Since teachers can use math workbooks to assign students homework regularly, the distribution of math workbooks would increase the frequency of homework assignment. The percentage of teachers who assigned math homework four or more times in a week was larger in the treatment group (69.5 percent) than the control group (52.6 percent). On the other hand, student study at home with workbooks and the homework checks by teachers were not satisfactory. At the follow-up survey, the survey team randomly selected up to 3 students in grade 7 and checked their workbooks. As shown in Figure 5, most pages of student workbooks were not checked by teachers. The result indicates that teachers did not provide sufficient feedback and guidance for students after homework.

At the beginning of school year, teachers in the treatment group developed an annual math teaching plan, which they regularly reviewed during periodical teacher meetings. During the inter-semestrial review meeting of teachers, they checked their progress and revised it considering the remaining number of lessons in the school year. The process might have facilitated teachers to identify the delay of curriculum and provide additional math class. The percentage of teachers who provided additional math class was larger in the treatment group (52.5 percent) than the control group (37.7 percent).

Figure 5: Spider chart of the status of student self-check and teacher check of workbook (7th grade)



Note: N. of students: 183 (only the treatment group)

Data source is the follow-up survey of this research.

The survey of this research also conducted math lesson observations at the end-line survey, which measured the length of time when students engaged in math learning. Surveyors counted the length of time throughout a lesson when half of students in a class either solved a math problem, referred to their notebook or textbook, or consulted with each other. The percentage of the engaged time in a lesson was slightly longer in the treatment group (22.4 percent) than the control group (18.9 percent).

The process data on teaching practices indicate several mediators for the improvement of student math learning in year 1, i.e., the volume of engaged time in a math lesson, the frequency of homework assignments, and the provision of additional math class. The next section discusses the average causal mediation effects of those mediators.

(2) Estimation strategy of the causal mediation effect

Causal mediation analysis employing the method of Imai et al. (2010) and taking the abovementioned three possible mediators sheds light on the mechanism underlying the impact on student math learning in year 1. Since the treatment group received a set of additional components, a simple strategy of using the treatment assignment as an instrument to estimate the impact via respective mediator violates the exclusion restriction assumption.

Following Imai et al. (2010), the conceptual framework of the causal mediation analysis is introduced. The mediation effect, λ_i, for student *i* can be described as

$$\lambda_i(t) \equiv Y_i(t, M_i(1)) - Y_i(t, M_i(0)),$$

where *t* represents the treatment assignment dummy that takes 0 or 1, and $M_i(t)$ is a potential value of mediator. The causal mediation effect $\lambda_i(t)$ is a change in Y_i due to the change in $M_i(0)$ from $M_i(1)$, holding the treatment status of student *i* constant at *t*. Then, the direct effect, ζ_i, for student *i* is defined as

$$\zeta_i(t) \equiv Y_i(1, M_i(t)) - Y_i(0, M_i(t)).$$

The direct effect represents the treatment impact that is not brought via the mediator $M_i(t)$. The average treatment impact can be decomposed to the average causal mediation effect and average causal direct effect.

Imai et al. (2010) introduces the following assumptions, called "Sequential Ignorability (SI)," to identify the average causal mediation effect;

SI assumption (1): $Y_i\{(t, m), M_i(t)\} \perp T_i \mid X_i = x$, and

SI assumption (2): $Y_i(t, m) \perp M_i(t) \mid T_i = t, X_i = x$.

SI assumption (1) indicates that the treatment assignment is independent of the potential mediator and potential outcome. The assumption is satisfied in a randomized controlled trial, since the treatment assignment T_i is random. On the other hand, SI assumption (2) requires that the mediator is independent of the potential outcome, given the individual treatment assignment and pre-treatment covariates (X_i). The assumption is strong since it will be violated once there are any unobservable pre-treatment covariates or post-treatment covariates that affect both the mediator and outcome variable.

Under SI assumptions, the average causal mediation effect is estimated using the following two equations,

$$M_{ij} = \alpha^1 + \varphi^1 Y_{ij0} + \tau^1 \text{Treatment}_j + C_{ij}\xi_c^1 + P_{kj0}\xi_p^1 + S_k\xi_s^1 + D_k\xi_D^1 + \omega_{ij}^1 \quad (6\text{-}1), \text{ and}$$

$$Y_{ij} = \alpha^2 + \tau^2 \text{Treatment}_j + \theta^2 M_{ij} + \varphi^2 Y_{ij0} + P_{kj0}\xi_p^2 + S_k\xi_s^2 + D_k\xi_D^2 + C_{ij}\xi_c^2 + \omega_{ij}^2 \quad (6\text{-}2).$$

where M_{ij} is a mediator, i.e., the engaged time in a math lesson; the frequency of math homework assignments; and the provision of supplementary math classes. The average causal mediation effect is expressed as the product of coefficients $\hat{\tau}^1$ and $\hat{\theta}$.

Since the impact of a mediator on an outcome variable might vary between the treatment and the control groups, equation (6-2) is expanded by adding the interaction term of the treatment assignment and the mediator as,

$$Y_{ij}=\alpha^3+\varphi^3 Y_{ij0}+\tau^3 Treatment_j+\theta^3 M_{ij}+\eta^3 (M_{ij}*Treatment_j)+$$

$$+ P_{kj0}\xi_p^3+ S_k\xi_s^3+ D_k\xi_D^3+C_{ij}\xi_c^3+\omega_{ij}^3 \quad (6\text{-}3).$$

Then, the heterogeneous causal mediation effect by the treatment group is expressed as the product of coefficients $\hat{\tau}^1$ and $(\hat{\theta}+\hat{\eta})$.[16]

Since the unobserved pre-treatment confounders of teachers or school principals might have affected both mediator and student math learning outcomes, the sensitivity analysis proposed by Imai et al. (2010) is conducted to see how the estimates changes when the assumption does not hold. When the SI assumptions hold, the error terms (ω) in (6-1) and (6-2) or (6-1) and (6-3) will not be correlated. The correlation of the error terms ρ is a parameter, which will be used in the sensitivity analysis. A larger value of ρ in the absolute term results in a larger bias in the estimation of the average causal mediation effect.

(3) Average causal mediation effects

The results of causal mediation analysis are presented in Table 8. Among the three mediators examined, only the provision of additional math class has a positive average causal mediation effect on student math learning. The impact is estimated at 0.029 standard deviations of math test scores, statistically significant at the 5 percent level (Column (1)). The average causal mediation effect represents 15.8 percent of the total impact. When examining the heterogeneity by the treatment and the control groups, while the average causal mediation effect on the control group is nearly zero and not statistically significant, the impact on the treatment group is 0.047 standard deviations, statistically significant at the 5 percent level.

[16] The R package "mediation" (Tingley et al., 2014) was used to estimate the average causal mediation effects. The effects are estimated by quasi-Bayesian Monte Carlo approximation.

Table 8: Results of the causal mediation analysis

	(1) Additional math class	(2) Frequency of homework	(3) Engaged time
Panel A (OLS)			
Impact of the treatment on the mediator	0.230***	0.376***	0.046***
	(0.061)	(0.086)	(0.019)
Adj. R. square	0.399	0.416	0.370
Panel B (Mediation analysis)			
ACME (average)	0.029**	0.026	-0.010
	[0.002; 0.060]	[-0.009; 0.070]	[-0.037; 0.010]
ADE (average)	0.158***	0.155**	0.193***
	[0.052; 0.260]	[0.035; 0.270]	[0.079; 0.300]
Total Effect	0.187***	0.181***	0.183***
	[0.078; 0.300]	[0.066; 0.290]	[0.071; 0.290]
% of impact mediated by ACME	14.6%	13.9%	-4.64%
ACME (Treatment)	0.047**	0.012	-0.013
	[0.007; 0.100]	[-0.040; 0.070]	[-0.045; 0.010]
ADE (Treatment)	0.177***	0.141**	0.190***
	[0.064; 0.290]	[0.018; 0.260]	[0.074; 0.300]
% of impact mediated (Treatment)	24.6%	6.44%	-6.12%
ACME (Control)	0.010	0.040*	-0.074
	[-0.030; 0.050]	[-0.001; 0.090]	[-0.040; 0.020]
ADE (Control)	0.139**	0.168***	0.196***
	[0.029; 0.250]	[0.043; 0.300]	[0.083; 0.310]
% of impact mediated (Control)	4.78%	21.4%	-3.16%
P value on hypothesis of ACME (Treatment) = ACME(Control)	0.159	0.387	0.705
Num. obs.	3,619	3,619	3,619
N. Clusters	232	232	232

Note: (1) Data source are the baseline and end-line of this research. ***p < 0.01; **p < 0.05; *p < 0.1 (2) Mediators analyzed: Provision of additional math class when curriculum was delayed; mediator analyzed in 9-2: Frequency of homework assignment; 9-3: Percentage of engaged time in a math lesson. (3) Panel A presents the OLS estimates of the treatment impact on the mediators, and Panel B shows the results of causal mediation analysis. (4) Student, teacher, and school characteristics, and the strata fixed effects constructed by department and urban status are controlled in all the regressions but not shown.

On the other hand, neither the average causal mediation effect through the engaged time in math lessons nor the frequency of homework assignments are statistically significant (Column (2) and (3)). In terms of the heterogeneity by the status of the treatment assignment, while the frequency of homework assignments had a positive average causal mediation effect for the control group (0.041 standard deviations, statistically significant at the 10 percent level), the effect is not statistically significant in the treatment group. The results indicate that while the homework assignment improves student math learning, the math workbook distributed in the treatment group did not function as intended.

(4) Sensitivity analysis

The estimates of the average causal mediation effect in the previous section are under the SI assumption that the correlation of error terms in equation (6-1) and (6-2), and that of (6-1) and (6-3) is respectively zero. However, unobserved characteristics of teachers might have affected both the mediator and student math learning, which could bring a bias in the estimates. Therefore, the sensitivity analysis on the estimates of the average causal mediation effect that had a statistically significant value is conducted. The results of the sensitivity analysis are presented in Figure E-1 and E-2 in Appendix E. The sensitivity parameter is the correlation of error terms, ρ . A negative value of the correlation of error terms produces upward bias in the estimate of the average causal mediation effect, and a positive value produces downward bias. As shown in the figures, the estimates of the average causal mediation effect are sensitive to the changes in the correlation. For example, when the value of correlation is 0.10, the average causal mediation effect of the provision of additional math classes for the treatment group becomes zero.

Though the error terms in equation (6-1), (6-2), and (6-3) are not observable, the correlation of the residuals from equations (6-1) and (6-2), and those from equations (6-1) and (6-

3), can be checked. Scatter plots of the residuals are presented in Figure E-3-1 to E-3-6 in Appendix E. If the pattern of scatter plots is not random, it suggests that there would be confounders that are not captured well in the regression equations above. The correlations between the residuals are nearly zero, and the scatter plots in those figures do not show a non-random pattern.

Among the three mediators in the causal mediation analysis, one of them could be a post-treatment confounder that affects both the other mediator and student math learning. The correlation between the percentage of engaged time in math lessons and the provision of math classes in the treatment group is checked. The increase in engaged time in a math lesson was not correlated with the provision of additional math classes (the correlation value: 0.02). On the other hand, the provision of math classes and the frequency of math homework assignments were weakly correlated in the treatment group (the correlation value: 0.14). Logically, it is plausible to think that when additional math class was provided, homework would have been assigned in the additional math class, which resulted in the increase in the frequency of homework assignments. On the other hand, it is difficult to imagine that frequent math homework caused a delay in math curricular, which brought the provision of additional math lesson.[17]

[17] The correlation between the provision of additional math classes and homework assignments might attenuate the impact of the mediator on student math learning outcomes. To check the robustness of the impact of the provision of additional math class on student learning outcomes, I included the other two mediators in the control variable to estimate the coefficient of the provision of math class in equation (6-2). The coefficient remained at almost same level. Without controlling the other two mediators, the coefficient of the provision of additional math classes in equation (6-2) was 0.139 (standard error: 0.062; p-value: 0.02). After including the other two mediations in equation (6-2), the coefficient of the provision of additional math classes is 0.131 (standard error: 0.061; p-value: 0.03). The other two mediators are checked in the same manner, and their coefficients also remained at almost same level.

5. Discussions

(1) Mechanism of the additional components in the ESMATE program

While the results of causal mediation analysis suggests that the provision of additional math class is the mediator for the improvement of math learning, the mediator can be a proxy for the shift of teacher attention during math lessons towards students with lower math achievement. At the inter-semestrial meetings, teachers in the treatment group checked the percentage of students in their classes who received low scores on the diagnostic test. The process might enable them to better understand the status of math learning among their students, and strengthen their supports in math lessons. The primary objective of the additional classes was catching up following the delay in the curriculum. When teachers focused on students with low academic achievement in daily math lessons, it might have caused a delay in teaching plans.

The average causal mediation effect accounted only for 15.8 percent of the total impact of the additional components. One of the possible reasons for this is the technical limit on lesson observations in the survey. At the end-line survey, surveyors counted the length of student engaged time during a lesson when half of students in the class were either solving a math problem, referring to their notebook or textbook, or consulting with each other. But the student engaged time surveyed might not have captured well the quality of teaching including the shift of teacher attention to those with lower math achievement.

(2) Future challenges for the ESMATE program

While the additional components in the ESMATE program improved math learning in year 1, this study demonstrates several challenges for further improvement of the program. First, the additional component did not work to improve learning in geometry. The result is in line with the other research on a different structured pedagogy program in Costa Rica. In this case, Berlinski and Busso (2017) investigated the impact of a structured pedagogy program in mathematics at the lower secondary level. The program included a set of teacher manuals and student workbooks, and teacher training aimed at changing teaching practices to improve student math learning in geometry. While teachers and students utilized the distributed materials, the students in the control group obtained higher test scores than the treatment group.

The additional components in the ESMATE program included the initial on-site support for school principals in math lesson observation, but this did not impact the frequency of lesson observations or suggestions from school principals. Classroom observation feedback and expert coaching can improve language and math through changing the practices of teachers (Bruns et al. 2018). The background and reasons behind null impact on the lesson observations by school principals is worth further research to improve the effectiveness of the ESMATE program.

Third, the accumulated impact of the first-year interventions in the following year is estimated at nearly zero and not statistically significant. At the end of grade 7, most students in both groups had difficulty solving simple equation problem like “solve the equation, $2x=8$” (the correct response rates: 17.8 percent in the treatment group, and 14.3 percent in the control group). Since math content in grade 8 involve equations, the low-level understanding of linear equations in the grade 7 hampered student math learning in the following year. In the follow-up survey, the correct response rates to simple simultaneous equations were less than 10 percent in both groups. The root cause of the struggle can be found in the initial learning level in grade 7. At the beginning of 7th grade, 62 percent of the students in the treatment group correctly responded to the item “answer the number in □ of equation, $2\times\square=8$”. While the structure of the item looks similar to the simple equation item above, the correct response rate to the latter was significantly lower. The result indicates that students in grade 7 had difficulty bridging math learning from the primary to the lower secondary level, when math content becomes more theoretical and abstract.

6. Conclusions

While recent debates on educational development focus on the learning crisis in primary education, the crisis in the lower secondary education level is equally profound. Through lesson observations in several countries in Latin America, Bruns and Luque (2015) argued that poor student learning results could be directly linked to the failure of teachers to keep students engaged in learning. One of the approaches to improve student learning is the structured pedagogy program that combines different types of interventions with the provision of teaching and learning materials. A structured pedagogy program in mathematics, the ESMATE program, was developed by the ministry of education in El Salvador with technical cooperation from JICA. While the ministry scaled up the ESMATE program nationwide in the lower secondary education level in 2018, the ministry could not offer some parts of the program because of budget constraints.

This study investigated the impact of the additional components on student math learning through a randomized controlled trial. The average one-year impact of the additional components in the ESMATE program is estimated at around 0.18 standard deviations of test scores. While the additional components improved math learning for students with lower and medium household economic status, they did not have the impact on students with higher household economic status. The mechanism of the additional components was then studied using causal mediation analysis, taking three mediators; student engaged time in math learning during a lesson; the frequency of homework assignments, and the provision of additional math classes when the curriculum delayed. Among the three mediators, the average causal mediation effect of the provision of additional math classes is positive and statistically significant. Since the provision of additional math classes can be a proxy for the shift of attention of teachers in math lessons, the results should be interpreted with a caution.

At the second year, both groups received the ESMATE textbooks, teachers' guidebooks, and student math workbooks. The accumulated impact of the first-year intervention in the following year is estimated at nearly zero and not statistically significant. At the end of grade 7, most students in both groups had difficulty solving simple equation problems. Since math content in grade 8 often involve equations, the low-level understanding of linear equations in grade 7 hampered progress in learning in the following year. The root cause of the low achievement level is found in the beginning of the lower secondary education. Students entered lower secondary education without the foundations of mathematics, and then they struggled with math content as it becomes more theoretical and advanced. To improve math learning in lower secondary education, it is essential to increase support for students to catch up on foundational math concepts and to transition smoothly in math learning from the primary to the lower secondary level, at the early stage in the lower secondary education.

Reference

Berlinski, S., and M. Busso. 2017. "Challenges in educational reform: An experiment on active learning in mathematics." *Economics Letters* 156): 172-75.

Bruns, B., and J. Luque. 2015. *Great Teachers: How to Raise Student Learning in Latin America and Caribbean*. Washington DC: World Bank.

Bruns, B., L. Costa, and N. Cunha. 2018. "Through the looking glass: Can classroom observation and coaching improve teacher performance in Brazil?" *Economics of Education Review* 64: 214-50.

Bhula, R., M. Mahoney, and K. Murphy. 2020. "Conducting cost-effectiveness analysis (CEA)." https://www.povertyactionlab.org/resource/conducting-cost-effectiveness-analysis-cea.

Cooper, H., J.C. Robinson, and E.A. Patall. 2006. "Does Homework Improve Academic Achievement? A Synthesis of Research, 1987-2003." *Review of Educational Research* 76 (1): 1-62.

de Hoyos, R., A.J. Ganimian, and P.A. Holland. 2019. "Teaching with the Test: Experimental Evidence on Diagnostic Feedback and Capacity Building for Public Schools in Argentina." *The World Bank Economic Review* 00 (0): 1-22.

Dhaliwal, I., E. Duflo, R. Glennerster, and C. Tulloch. 2014. "Comparative Cost-Effectiveness Analysis to Inform Policy in Developing Countries: A General Framework with Applications for Education." In *Education Policy in Developing Countries*, edited by P. Glewwe, 285-338. Chicago: University of Chicago Press.

Fan, H., J. Xu, Z. Cai, J. He, and X. Fan. 2017. "Homework and students' achievement in math and science: A 30-year meta-analysis, 1986 – 2015." *Educational Research Review* 20: 35-54.

Hanushek, E. A., and L. Woessmann. 2012) "Schooling, educational achievement, and the Latin American growth puzzle." *Journal of Development Economics* 99 (2): 497-512.

Imai, K., L. Keele., and D. Tingley. 2010. "A General Approach to Causal Mediation Analysis." *Psychological Methods* 15 (4): 309-34.

Kerwin, Jason, T., and R.L. Thornton. 2020. "Making the grade: The sensitivity of education program effectiveness to input choices and outcome measures." *Review of Economics and Statistics* vol. 103 (2): 251-64.

Koei Research & Consulting Inc. 2018. *Impact Evaluation for the Project for the Improvement of Mathematics Teaching in Primary and Secondary Education in El Salvador: Baseline Survey Report*. JICA. Unpublished.

Koei Research & Consulting Inc. 2019. *Impact Evaluation for the Project for the Improvement of Mathematics Teaching in Primary and Secondary Education in El Salvador: End-line Survey Report*. JICA. Unpublished.

Kremer, M., C. Brannen, and R. Glennerster. 2013. "The Challenge of Education and Learning in the Developing World." *Science* 340: 297.

Maruyama, T., and T. Kurosaki. 2021. "Developing textbooks to improve math learning: Empirical Evidence from El Salvador." JICA RI working paper No.217, JICA RI.

Ministerio de Educatión, El Salvador. 2014. *EDUCACIÓN DE EL SALVADOR EN CIFRAS 2009-2014*.

Ministerio de Educatión, El Salvador. 2016. "Bases de Datos por Centros Escolares - Censo Escolar: Inicial Directorio 2016." El Salvador. https://www.mined.gob.sv/estadisticas-educativas/item/6116-bases-de-centros.html.

Ministerio de Educatión, El Salvador. 2018. "Bases de Datos por Centros Escolares - Censo Escolar: Inicial Directorio 2018." El Salvador. https://www.mined.gob.sv/estadisticas-educativas/item/6116-bases-de-centros.html.

Ministerio de Educatión, El Salvador. 2018. "Bases de Datos por Estudiante - Censo Escolar Inicial: Año 2018." https://www.mined.gob.sv/estadisticas-educativas/item/7154-bases-de-datos-spss.html.

Mullis, I.V.S., M.O. Martin, P. Foy, and M. Hooper. 2016. *TIMSS 2015 International Results in Mathematics*. TIMSS & PIRLS International Study Center, IEA.

Snilstvelt, B., J. Stevenson, R. Menon, D. Phillips, E. Gallagher, M. Geleen, H. Jobse, T. Schmidt, and E. Jimenez. 2016. *The Impact of education programmes on learning and school participation in low- and middle- income countries*. London: International Initiative for Impact Evaluation.

Tingley, D., T. Yamamoto, K. Hirose, L. Keele, and K. Imai. 2014. “mediation: R Package for Causal Mediation Analysis.” *Journal of Statistical Software* 59 (5).

UNESCO. 2015. *Executive summary: Third Regional Comparative and Explanatory Survey, Learning Achievements*. Santiago: Regional Bureau for Education in Latin America and Caribbean, UNESCO.

UNESCO. 2017. *More than One-half of Children and Adolescents Are Not Learning Worldwide. Fact Sheet No. 46*. (UIS/FS/2017/ED/46).

USAID. 2017. *EVALUACIÓN DEL SECTOR EDUCACIÓN DE EL SALVADOR USAID/EL SALVADOR*.

World Bank. 2021. “World development indicators.” https://data.worldbank.org/indicator/

Appendix A

Table A-1: Composition of math test (baseline)

Item No.	Item content	Cognitive domain	Cognitive skill
1	To add two-digit numbers with carrying to the position of tens	NO	Knowing
2	To subtract a two-digit number from a two-digit number with borrowing from the position of tens	NO	Knowing
3	To find the product of three 2's	NO	Knowing
4	To operate division of three-digit number by three-digit number without remainder	NO	Knowing
5	To operate multiplication and division of one-digit numbers successively	NR	Knowing
6	To subtract a number with one decimal place from a number of the same type without borrowing	NO	Knowing
7	To operate division without residue of numbers with one decimal place	NO	Knowing
8	To add two proper fractions with the same denominator without reduction	NO	Knowing
9	To find the product of two proper fractions without reduction	NO	Knowing
10	To find the quotient of two proper fractions without reduction	NO	Knowing
11	To find the least number which can be divided without remainder both by 4 and by 6	NO	Knowing
12	To find the number which gives 3 when subtracting it from 7 (the item is presented in the form of equation representing the unknown number by ☐)	NR	Applying
13	To find the number which gives 8 when multiplying 2 by that number (the item is presented in the form of equation representing the unknown number by ☐)	NR	Applying
14	To find how many times 4 is 8 (the unknown number is represented by ☐)	NR	Knowing
15	To find the unknown number in the table which shows the relation between the quantity of the same goods and their total price	NR	Knowing
16	To find the unknown number in the table which shows the relation between the quantity of workers and the quantity of days which is needed to finish the same work	NR	Applying
17	To draw the development view of a die (an illustration of cutting it is attached)	G	Reasoning
18	To complete the figure which shows the left-hand side of a symmetric form	G	Knowing
19	To find the area of a rectangle from the length of its base and height	QM	Knowing
20	To find the sum of the length of the sides of the rectangle of the previous item	QM	Knowing

Note: Test items are developed in this research. NO: Number and Operation; QM: Quantity and Measurement; NR: Numerical and Quantitative Relation; G: Geometry

Table A-2: Composition of math test (End-line)

Item No.	Item content	Cognitive domain	Cognitive skill
1	To write a negative number which corresponds to a point in the number line	NE	Knowing
2	To add a one-digit negative number and a one-digit positive number (the answer is positive)	NE	Knowing
3	To subtract a one-digit negative number from a one-digit positive number (the answer is a two-digit number)	NE	Knowing
4	To multiply a one-digit positive number by a one-digit negative number (the answer is a one-digit number)	NE	Knowing
5	To divide a one-digit negative number by a one-digit negative number	NE	Knowing
6	To find the greatest common divisor of three integers	NE	Knowing
7	To find the expression of the sum of two numbers, one of which is represented by the letter x	NE	Knowing
8	To add two polynomials of degree one in one variable	NE	Knowing
9	To find the value of a polynomial of degree one in one variable when the variable is substituted by a number	NE	Knowing
10	To solve an equation of the first degree (by transposition)	NE	Knowing
11	To solve an equation of the first degree (by dividing both sides by the coefficient of the variable)	NE	Knowing
12	To solve an equation of the first degree (by transposition and division)	NE	Knowing
13	To find the unit price from the sum of the money, the number of units and the change	NE	Knowing
14	To fill a blank in the table which shows the relation between the number of units and the total price knowing the unit price	F	Knowing
15	To fill a blank in the table which shows the relation between the length of base and height of rectangles of the same area	F	Knowing
16	Given the graph of proportion, to find the coefficient of the variable in the equation	F	Applying
17	To find the total number knowing the proportion and the number of one part in a pie chart	S	Applying
18	To draw a figure which is symmetric to a given triangle with respect to a line	G	Knowing
19	When two circles have two points in common, to find the angle formed by the line which connects the centers and the line which connects the common points	G	Reasoning
20	To find the number of the lateral face of a prism whose base is rectangle	G	Knowing

Note: Test items are developed in this research. NE: Number and Expression; F: Function; G: Geometry; S: Statistics

Table A-3: Composition of math test (Follow-up)

Item No.	Item content	Cognitive domain	Cognitive skill
1	To subtract a one-digit negative number from a one-digit positive number (the answer is two-digit number)	NE	Knowing
2	To add two polynomials of degree one in one variable	NE	Knowing
3	To find the value of a polynomial of degree one in one variable when the variable is substituted by a number	NE	Knowing
4	To solve an equation of the first degree (by transposition and division)	NE	Knowing
5	Given the graph of proportion, to find the coefficient of the variable in the equation	F	Applying
6	To arrange the terms of a polynomial in one variable according to their degree	NE	Knowing
7	To multiply two monomials	NE	Knowing
8	To add two polynomials of degree one in two variables	NE	Knowing
9	To solve a system of equations of the first degree in two variables	NE	Knowing
10	To solve a written problem in which two unknown numbers appear	NE	Knowing
11	To find the values of two constants so that two systems of equations of the first degree in two variables have the same solution	NE	Applying
12	To find the y-intercept of the graph of a function of the first degree	F	Knowing
13	To fill the table which represents some values of variables of a function of the first degree	F	Knowing
14	To find the rate of change of a function of the first degree from a table of the values of their variables	F	Knowing
15	Given the values of two pairs of variables and one variable of the third pair of a function of the first degree, to find the value of another variable of the third pair (the function represents the relation between the height of water and the time)	F	Knowing
16	To find the equation of the function of the first degree whose graph is symmetric to the graph of the given function with respect to the abscissa	F	Reasoning
17	To find the value of the corresponding angle to the given one formed by two parallel lines and a transversal	G	Knowing
18	To find the applicable condition for congruence of given triangles	G	Knowing
19	To find the length of one side of a parallelogram when the length of two consecutive sides are given	G	Knowing
20	To find the measure of an internal angle of a parallelogram when the measure of its opposite angle is given	G	Knowing
21	To find the measure of an angle formed by a diagonal and one side of a rectangle when the measure of an angle formed by its diagonals is given	G	Applying
22	To find the name of a solid generated by a right-angled triangle when it is rotated on one of its catheti	G	Knowing
23	To find the volume of a pyramid whose base is square from the length of the side of its base and its height	G	Knowing
24	To find the number of rods of a certain range of length from a frequency distribution table	S	Knowing
25	To find the midpoint of the modal class of a frequency distribution table	S	Knowing

Note: (1) Test items are developed in this research. NO: Number and Expression; F: Function; G: Geometry; S: Statistics. (2) No. 1 through 5 are the same items from the end-line survey.

Principal component analysis of the different types of student household assets

The baseline survey collected information on the following different types of student household assets through student interviews: (a) smartphone; (b) computer; (c) refrigerator; (d) car; (e) television; (f) access to tap water; and (g) flush toilet. In addition, the survey also collected information on the access to electricity and the use of wood for cooking. The data on student household assets consists of dummy variables which take the value of 1 in the case that the student household possess each category of asset that contributes to the wealth of a household. In case of the cooking fuel, as the use of wood is an indicator of the lack of wealth (more wealthy households use gas or electricity for cooking), the non-use of woods in cooking takes the dummy variable representing the asset. A composite index of student household economic status was computed by the principal component analysis.[18] The statistics of first principal component is presented in Table B-1. The first principal component accounts for 25.4 percent of the variances, and all nine asset types have positive loading coefficients, indicating that they contribute to household wealth.

Table B-1: Statistics of first principal component

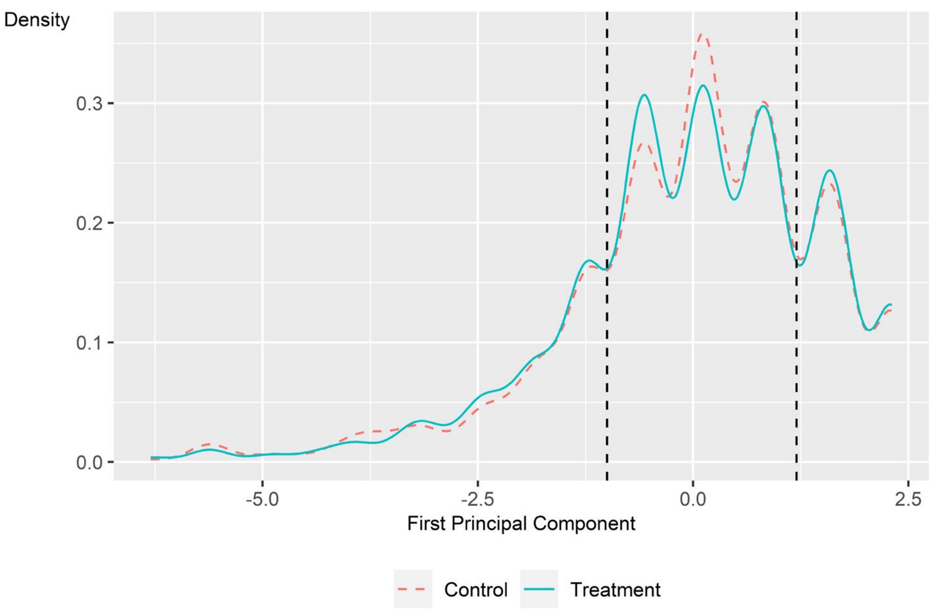
	Coefficient	Correlation with the first principal component
Smartphone	0.238	0.360
Computer	0.390	0.590
Refrigerator	0.402	0.608
Car	0.361	0.546
Television	0.348	0.526
Access to tap water	0.268	0.405
Electricity	0.320	0.484
Flush toilet	0.340	0.513
Non-use of wood in cooking	0.296	0.447

Note: Data Source is the baseline survey of this research.

[18] For Figure 1-1 and 1-2 in the paper, the student household economic status index was computed with the baseline sample. For the analysis of heterogeneous impacts in section 3. (2) in the paper, the index was computed with the sample of students who were present both at the baseline and end-line surveys. In this appendix, the statistics of the principal component analysis are reported with the sample of students who were present both at the baseline and end-line surveys.

Then, Figure B-1 presents the density curves of the first principal component. Based upon the distribution of the first principal component, the whole sample of students is divided into three groups: lower; medium; and higher household economic status. The dashed lines in Figure B-1 show the threshold values in the first principal component for the groupings. The students with lower household economic status represent around 22 percent respectively in the treatment and control groups. The students with medium household economic status represent around 56 percent in the two groups. The students with higher household economic status represent around 22 percent in the two groups.

Figure B-1: Density curves of the first principal component



Note: Data Source is the baseline survey of this research.

Table B-2 shows characteristics of each household economic status. From lower to higher household economic status, the possession rates of different type of household assets increased. The possession rates of different types of household assets were well balanced between the treatment and the control groups.

Table B-2: Possession rates of different types of household assets by household economic status

Item	Treatment group	Control group	Difference	p value
Lower household economic status (average possession rate %)				
Smartphone	64.6	63.7	0.9	0.844
Computer	2.9	2.6	0.3	0.817
Refrigerator	48.7	44.6	4.1	0.392
Car	2.9	3.9	-1.0	0.428
Television	74.3	75.0	-0.7	0.852
Access to tap water	49.2	54.6	-5.4	0.342
Access to electricity	85.5	83.0	2.5	0.441
Flush toilet	20.7	21.6	-0.9	0.783
Non-use of wood in cooking	12.4	12.6	-0.2	0.926
Number of students	421	388		
Medium household economic status (average possession rate %)				
Smartphone	90.8	89.4	1.4	0.584
Computer	20.3	21.2	-0.9	0.535
Refrigerator	94.8	94.1	0.7	0.516
Car	24	23.1	0.9	0.335
Television	99.2	99.2	0	0.696
Access to tap water	82.8	84.8	-2	0.613
Access to electricity	99.8	99.8	0	0.529
Flush toilet	53.2	57.5	-4.3	0.191
Non-use of wood in cooking	36.5	36.0	0.5	0.896
Number of students	1,038	984		
higher household economic status (average possession rate %)				
Smartphone	98.8	97.9	0.9	0.335
Computer	88.4	85.1	3.3	0.157
Refrigerator	100	100	0	1.000
Car	78.3	80.3	-2.0	0.584
Television	100	100	0	1.000
Access to tap water	97.6	98.4	-0.8	0.492
Access to electricity	100	100	0	1.000
Flush toilet	89.9	94.7	-4.8	0.104
Non-use of wood in cooking	80.4	77.1	3.3	0.443
Number of students	414	375		

Note: Data Source is the baseline survey of this research. The p-values show the results of the chi-squares test with the stratified data (department and urban status).

Table C: Attrition analysis

	OLS End-line (1)-1	Logit End-line (1)-2	OLS Follow-up (2)-1	Logit Follow-up (2)-2
(Intercept)	-0.568***		-0.576***	
	(0.084)		(0.093)	
Treatment	-0.013	-0.013	-0.033	-0.034
	(0.020)	(0.021)	(0.022)	(0.024)
Z score baseline	-0.030***	-0.032***	-0.044***	-0.047***
	(0.006)	(0.007)	(0.007)	(0.008)
Age	0.054***	0.042***	0.068***	0.070***
	(0.006)	(0.005)	(0.007)	(0.008)
Sex (Male=1)	0.022**	0.023**	0.029**	0.031**
	(0.011)	(0.011)	(0.013)	(0.014)
N. of elder brother/sister	0.000	0.000	-0.000	-0.000
	(0.003)	(0.003)	(0.003)	(0.003)
N. of younger brother/sister	0.015***	0.013***	0.009*	0.009*
	(0.005)	(0.004)	(0.005)	(0.005)
Shift (Morning=1)	0.001	0.001	-0.031	-0.032
	(0.026)	(0.024)	(0.031)	(0.032)
Repeated in 2017	0.089***	0.080***	0.076***	0.065***
	(0.020)	(0.019)	(0.021)	(0.022)
N. of asset types	-0.009*	-0.009*	-0.007	-0.008
	(0.005)	(0.005)	(0.006)	(0.006)
La Union	0.073**	0.085*	0.069*	0.075*
	(0.037)	(0.045)	(0.039)	(0.042)
San Miguel	0.098***	0.101***	0.025	0.028
	(0.037)	(0.039)	(0.038)	(0.039)
San Vicente	0.073**	0.077**	-0.004	-0.002
	(0.028)	(0.032)	(0.031)	(0.032)
Urban	0.013	0.006	0.025	0.028
	(0.040)	(0.050)	(0.056)	(0.061)
La Union×Urban	-0.041	-0.029	-0.078	-0.075
	(0.052)	(0.053)	(0.067)	(0.060)
San Miguel×Urban	-0.075	-0.060	-0.021	-0.023
	(0.052)	(0.048)	(0.062)	(0.064)
San Vicente×Urban	-0.049	-0.040	-0.049	-0.057
	(0.048)	(0.050)	(0.057)	(0.058)
Adj. R^2	0.087		0.085	
Num. obs.	4,513	4,513	4,513	4,513
F statistic	19.973		26.376	
N. Clusters	237		237	
Log Likelihood		-2,056.072		-2,572.858
Deviance		4,112.144		5,145.715

***p < 0.01; **p < 0.05; *p < 0.1

(1) Data source is the baseline survey of this research.
(2) Dependent variable is a dummy that takes 1 for students who were absent respectively at the end - line or follow-up survey. Robust standard errors are clustered at the school level, and are in parenthesis.
(3) Coefficients of the logit regression show marginal effects.

Appendix D

Analysis of the impacts of the additional components in the ESMATE program on teaching practices in year 1

Several possible intermediate outcomes on the one-year impact of the additional components were reviewed, taking teaching practices in math class, the provision of additional math class, the frequency of homework assignments, and the frequency of lesson observations and suggestions by school principals.

1. Teaching practices in math class

This research conducted math lesson observations at the end-line survey, which measured the length of time when students engaged in math learning. Surveyors counted the length of time in a lesson when half of students in a class were solving math problems, referring to their notebook or textbook, or consulting with each other. The percentage of the engaged time in a lesson in the treatment group was slightly longer than the control group (22.4 percent in the treatment group, and 18.9 percent in the control group). The percentage of engaged time is regressed on the treatment assignment, characteristics of teachers and schools, strata fixed effects constructed by department and urban status. The estimated impact is around 5 percent point (standard error: 2.3 percent point), which is statistically significant at the 5 percent level. Since the standard length of a lesson is 45 minutes, the difference is equivalent to around 2 to 3 minutes.

The lesson observation also checked the following instructional routine of teachers in math class: (a) teacher posed math problem; (b) teacher checked student notebooks; (c) teacher walked around the classroom to check notebooks; (d) teacher advised students to consult with each other; (e) teacher told students to check their answers; (f) teacher instructed students to try incorrect problems again; and (g) teacher assigned homework. The surveyor noted those points according to the speech and/or behavior of teachers. The results were similar in the treatment and the control groups except the point (f) above. A larger percentage of teachers in the treatment group instructed students to try again if they answered a question incorrectly than in the control group (74.6 percent in the treatment group, and 63.2 percent in the control group). The dummy variable of instructional routine (a) to (f) is regressed respectively on the treatment assignment, characteristics of teachers and schools, and strata fixed effects. Only the regression result on point (f) is statistically significant. The estimated impact is 13.1 percent point (standard error: 7.1 percent point), which is statistically significant at the 10 percent level.

2. Provision of additional math class when curriculum was delayed

At the beginning of school year, teachers in the treatment group developed an annual math teaching plan, and regularly reviewed it at their periodical meeting of teachers. The annual math teaching plan is a simple one-page format with a year-long calendar that defines which page of textbook is to be taught on which day. During inter-semestrial review meetings of teachers, they checked their progress, and revised plans considering the remaining number of lessons in the school year. The process might facilitate teachers to identify any delays in their teaching plan and provide additional math classes. The percentage of teachers who provided additional math classes was larger in the treatment group than in the control group (52.5 percent in the treatment group, and 37.7 percent in the control group).

The survey asked teachers whether they provide additional math classes when the curriculum was delayed (Yes: 1; No: 0). The dummy is regressed on the treatment assignment and characteristics of students, teachers and schools, and strata fixed effects. The estimated impact is 20.9 percentage point (standard error: 7.6 percent point), which is statistically significant at the 1 percent level.

3. Frequency of homework assignments

Since the math workbook was designed to correspond to the contents of the ESMATE textbook, teachers can use it to assign homework for students. Then, the distribution of math workbooks might have increased the frequency of homework assignment. The percentage of teachers who assigned math homework four or more times in a week was larger in the treatment group (69.5 percent) than the control group (52.6 percent). The survey asked teachers about the frequency of homework assignment in a week by the following four levels (1: Never; 2: Once a week or less; 3: Two or three times a week; and 4: Four or more times a week). The variable is regressed on the treatment assignment and characteristics of teachers and schools and strata fixed effects. The estimated impact is 0.41 level (standard error: 0.096), which is statistically significant at the 1 percent level.

The survey also asked students about the daily duration of study at home using the following five levels (1: Never; 2: Less than 15 minutes; 3: Between 15 to 30 minutes; 4: Between 30 minutes and 1 hour; and 5: More than one hour). The variable is regressed on the treatment assignment and characteristics of students, teachers, and schools, and strata fixed effects. The estimated impact is nearly zero, and not statistically significant.

4. Lesson observations and suggestions from school principal to teachers

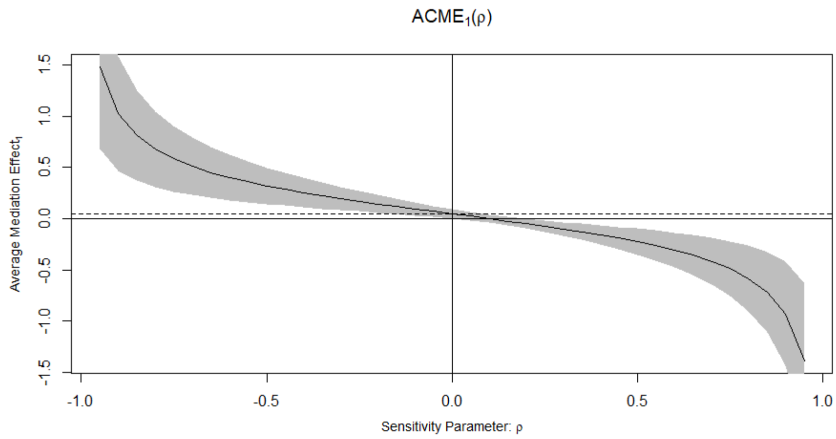
The additional components in the ESMATE program for the treatment group include initial on-site advice for school principals, which intended to increase the frequency of their lesson observations and suggestions for teachers. The survey conducted interviews with teachers about math teaching and support received by the school principal. There was no difference between the two groups in the support by school principal, i.e., frequencies of lesson observations and suggestions provided during those observations.

At the end-line survey, teachers reported on the frequency of math lesson observations and suggestions by school principal using the following five levels (1: Never; 2: Once in a year; 3: Twice in a year; 4: More than three times in a year but less than once in a month; 5: More than once a month). The frequency of math lesson observations and suggestions is regressed on the treatment assignment, characteristics of students, teachers and schools, and strata fixed effects. Both the estimated impact in the frequency of lesson observations and suggestions are around 0.20 to 0.25 but not statistically significant.

Appendix E

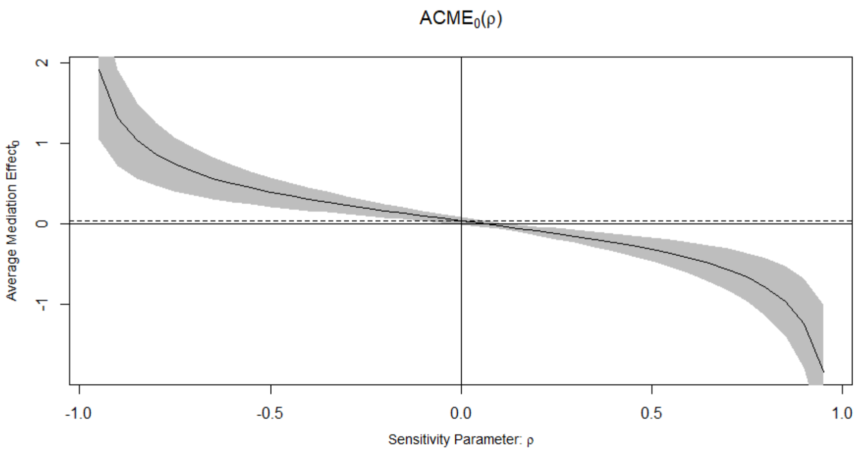
Sensitivity analysis plots

Figure E-1: Plot of sensitivity analysis on the average causal mediation effect of the provision of additional math class for the treatment group



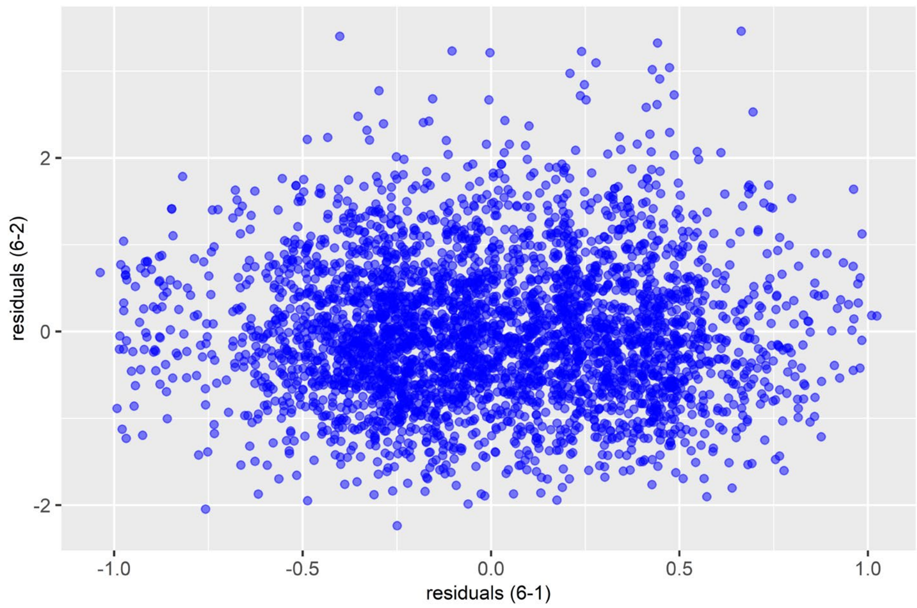
Note: (1) The solid line represents the estimated average causal mediation effect for the given mediator and for different values of ρ, while the shaded area represents the 95 percent confidence interval. The dashed line represents the estimated level of average causal mediation effect when ρ =0. (2) The value of ρ for which the average causal mediation effect becomes zero is 0.10. (3) Data source is the baseline and end-line surveys.

Figure E-2: Plot of the sensitivity analysis on the average causal mediation effect of the frequency of homework assignment for the control group



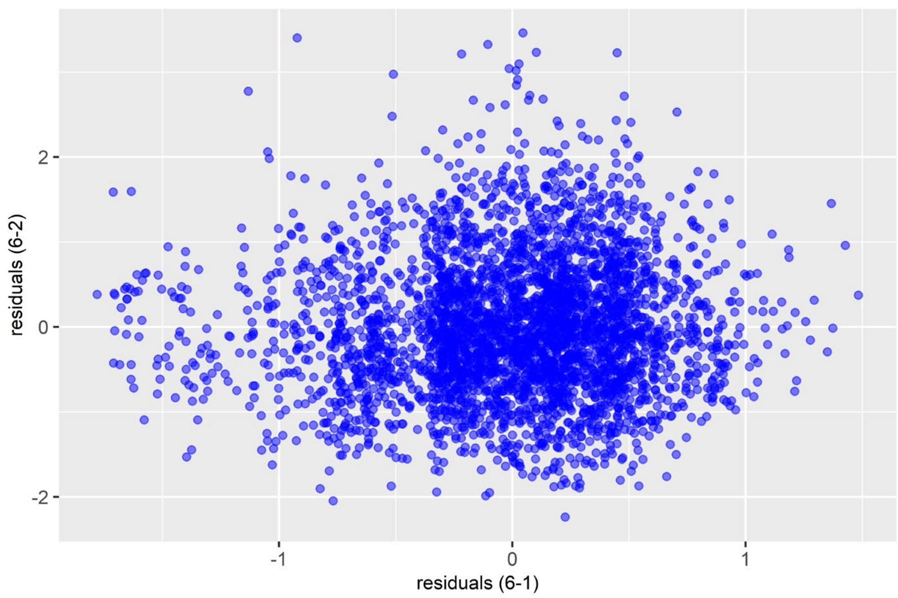
Note: (1) The solid line represents the estimated average causal mediation effect for the given mediator and for different values of ρ, while the shaded area represents the 95 percent confidence interval. The dashed line represents the estimated level of average causal mediation effect when ρ =0. (2) The value of ρ for which the average causal mediation effect becomes zero is 0.10. (3) Data source is the baseline and end-line surveys.

Figure E-3-1: Scatter plot of residuals in (6-1) and residuals in (6-2) in the causal mediation analysis of the provision of additional math class



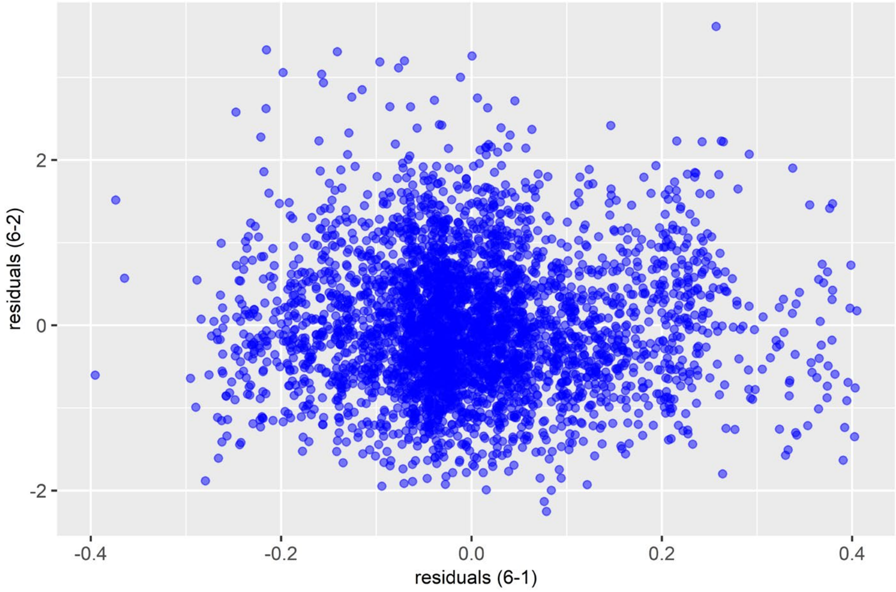
Note: Data source is the baseline and end-line surveys.

Figure E-3-2: Scatter plot of residuals in (6-1) and residuals in (6-2) in the causal mediation analysis of the frequency of homework assignments



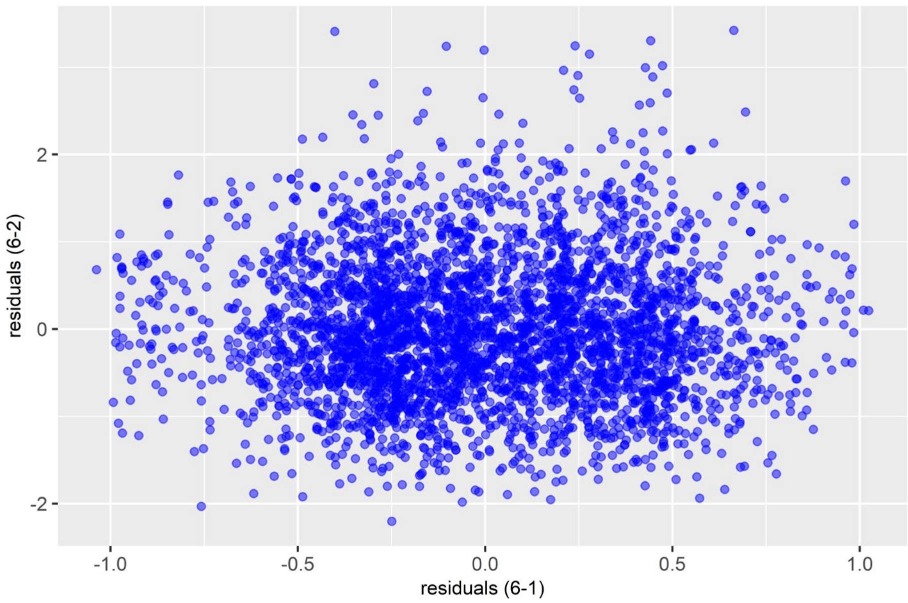
Note: Data source is the baseline and end-line surveys.

Figure E-3-3: Scatter plot of residuals in (6-1) and residuals in (6-2) in the causal mediation analysis of the percentage of engaged time in math lessons



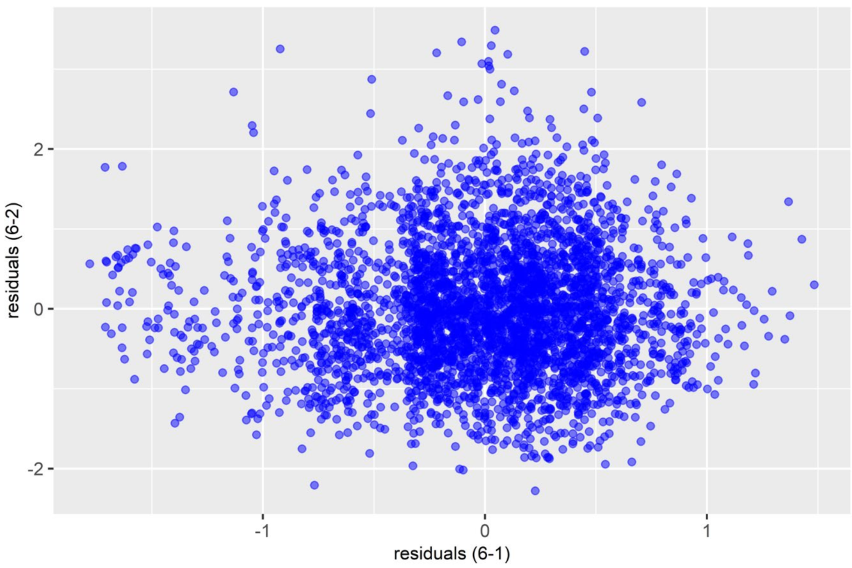
Note: Data source is the baseline and end-line surveys.

Figure E-3-4: Scatter plot of residuals in (6-1) and residuals in (6-3) in the causal mediation analysis of the provision of additional math class



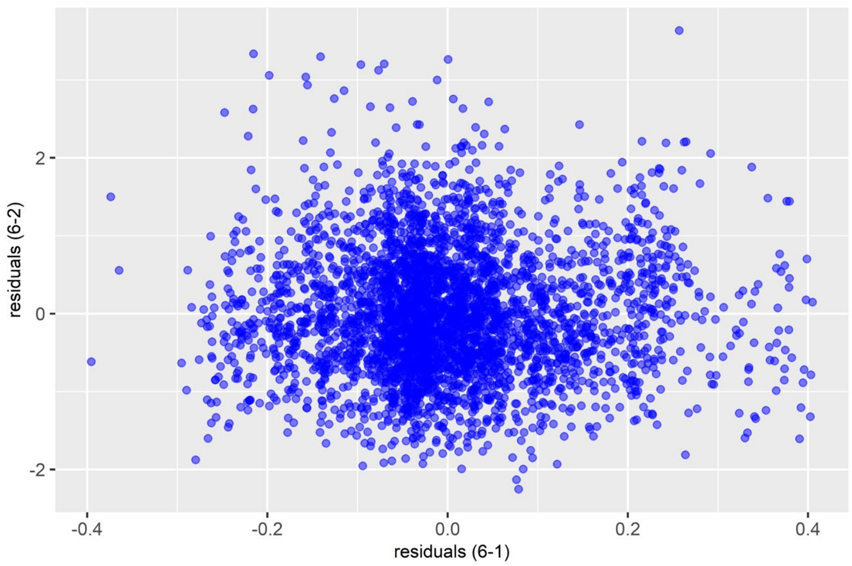
Note: Data source is the baseline and end-line surveys.

Figure E-3-5: Scatter plot of residuals in (6-1) and residuals in (6-2) in the causal mediation analysis of the frequency of homework assignment



Note: Data source is the baseline and end-line surveys.

Figure E-3-6: Scatter plot of residuals in (6-1) and residuals in (6-3) in the causal mediation analysis of the percentage of engaged time in math lessons



Note: Data source is the baseline and end-line surveys.

Abstract (in Japanese)

要約

近年、教育開発分野では初等教育段階における学習の危機が盛んに議論されるが、前期中等教育段階における学習の危機も同様に深刻である。世界の前期中等教育学齢期の子どもの約 6 割の数学の学力は、同教育段階で最低限到達すべき水準に達しておらず、その大半は低所得国及び低中所得国の子ども達である。それら子ども達の数学の学習改善を図るアプローチの一つとして、教授・学習教材の提供や教員研修等の複数のコンポーネントの介入パッケージからなる"Structured pedagogy program"がある。"Structured pedagogy program"による生徒の数学の学習成果への効果は、提供された教材を教員が有効に活用して生徒への学習支援を強化することにより、さらに向上されうる。本研究は、エルサルバドル国における算数教科書の配布を含む"Structured pedagogy program"を事例とし、生徒の数学テスト結果をもとにした教員間の定期的な振り返り等の介入を組み入れることにより、介入パッケージによる生徒の数学学習改善への効果が向上するかをランダム化比較試験により検証する。本研究では、公立学校の前期中等 1 学年の生徒を 2 年間にわたって追跡した。調査 1 年次に、介入による効果は、生徒の数学テストの 0.18 標準偏差と推定され、その効果は家計の経済的地位の中程度から低い層に見られた。他方で、対照群においても、教員間の振り返り等の同様の介入が行われた調査 2 年次には、介入群と対照群の間で生徒の数学の学習成果における差はみられなかった。さらに、本研究では、調査 1 年次のデータに焦点をあて、媒介分析（Causal mediation analysis）の手法を用い、"Structured pedagogy program"に組み入れられた教員間の定期的な振り返り等の介入から学習成果向上への因果経路を分析した。

キーワード：教育開発、数学学習、前期中等教育、中南米地域、インパクト評価

Working Papers from the same research project

JICA-RI Working Paper No. 217
Developing Textbooks to Improve Student Math Learning:
Empirical Evidence from El Salvador
Takao Maruyama and Takashi Kurosaki