Individualized Self-learning Program to Improve Primary Education: Evidence from a Randomized Field Experiment in Bangladesh

BEP x Kumon Project Research Team:
Y. Sawada, M. Mahmud, M. Seki, A. Le, and H. Kawarazaki

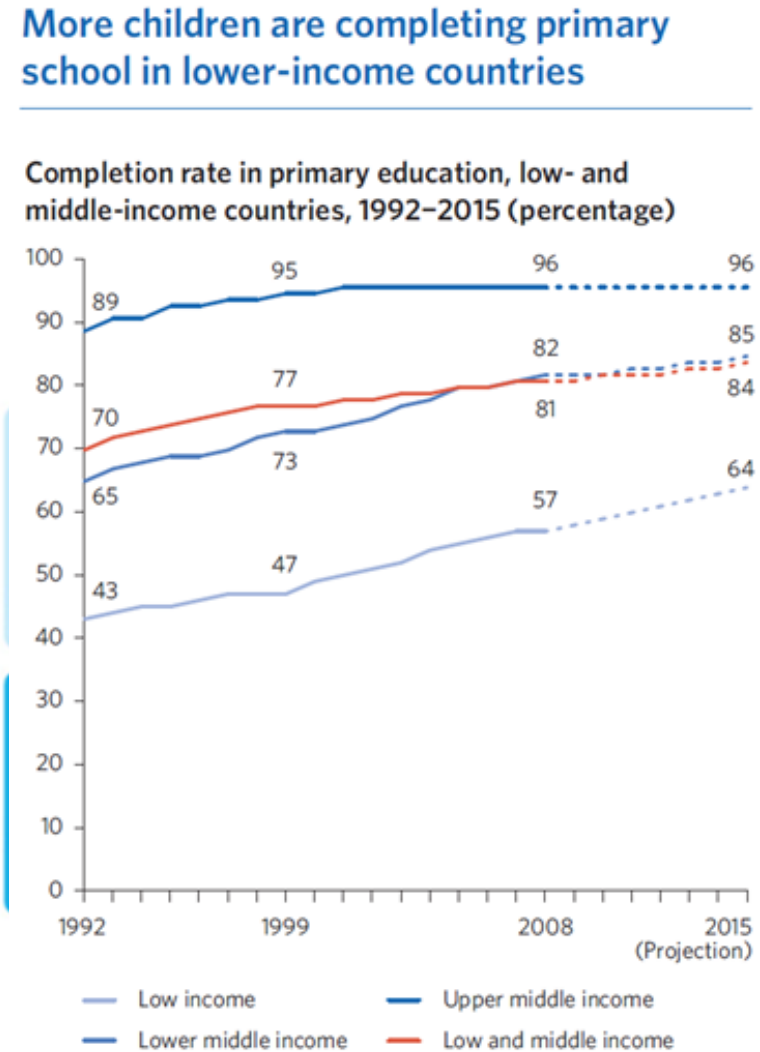
ASSA - AEA Session
January 7th 2018

Disclaimer

The views expressed are those of the authors. They do not necessarily reflect the views of the affiliated institutions.

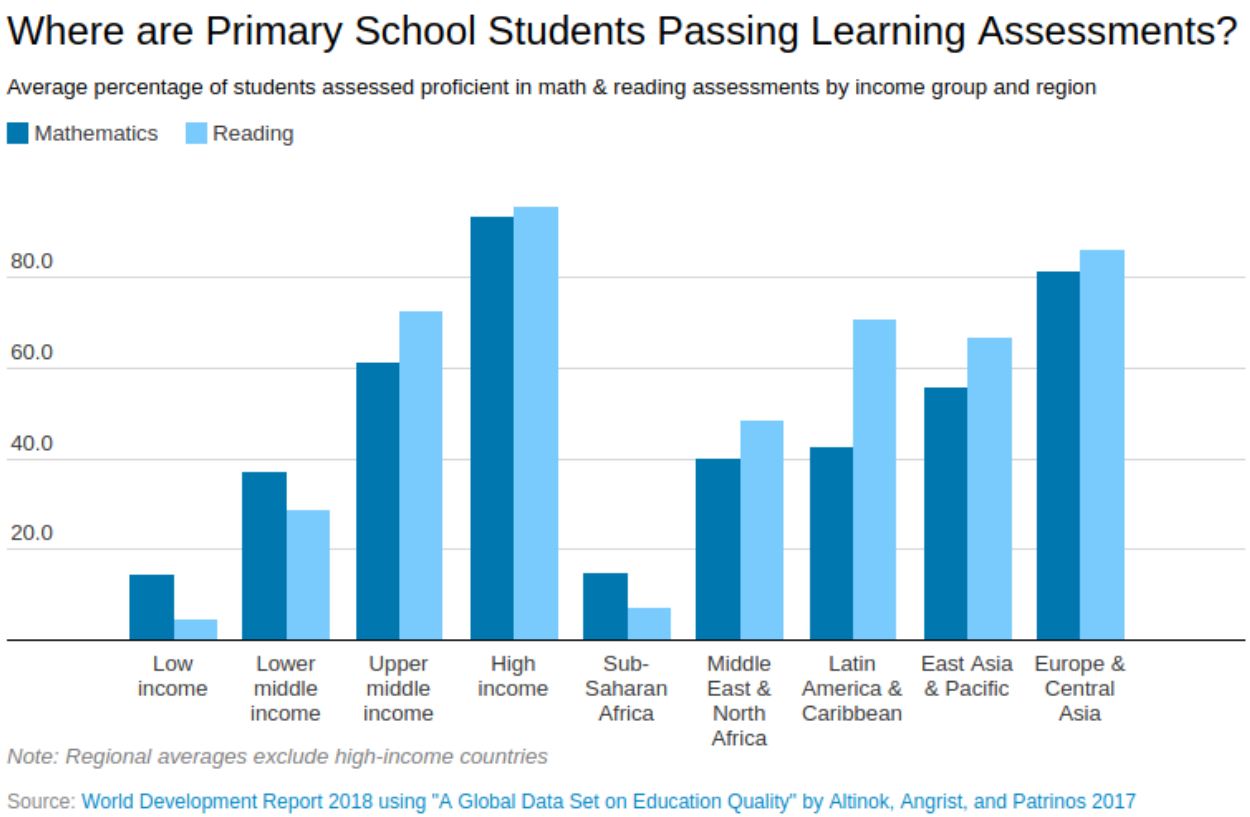
Motivation: Improved Universal Primary Education

- MDGs (until 2015):
- Impressive progress in the primary completion rate!



Motivation: Global Learning Crisis

- SDGs (post 2015): Inclusive and quality education for all
  - Global learning crisis: More than 60 percent of primary school children in low- and middle- income countries fail to achieve a minimum proficiency in mathematics and reading (UNESCO 2013, WDR 2018).



Research Questions

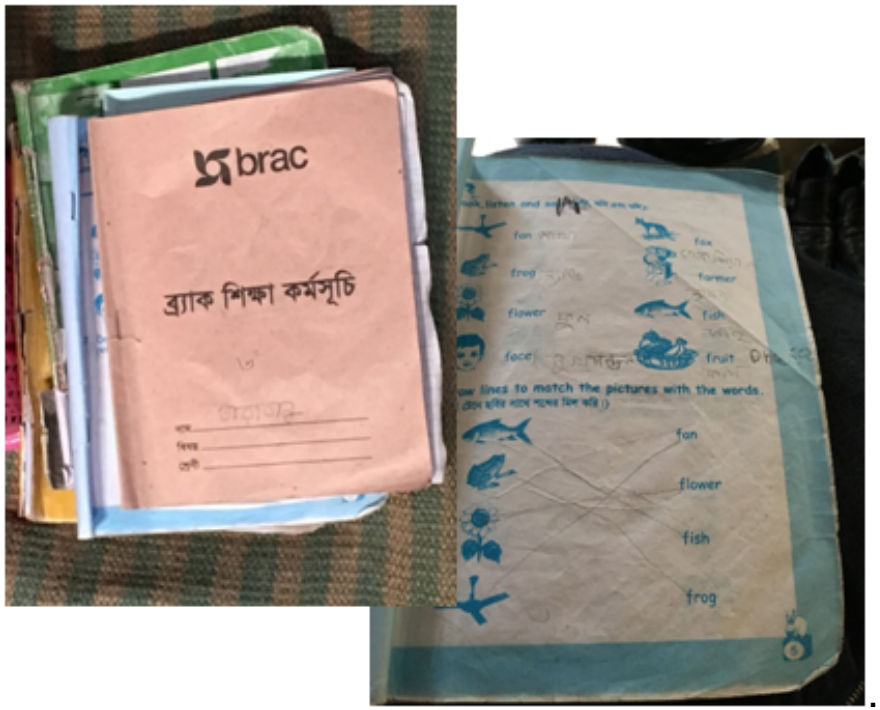
What happens to learning if we introduce one of globally successful non-formal education methods to disadvantaged schools in very low income communities?

- To target poor students studying at BRAC primary schools (BPS)
- To examine the KUMON Method of Learning (Kumon), a globally popular non-formal education program (4.35 mill enrollments in 50 countries and regions in March 2017)
- To investigate both cognitive and non-cognitive abilities
- To adopt an RCT-based experiment

Preview

- Overall impact:
  - Robust and better learning ability.
- Impact on teachers:
  - Better assessment capacities of students' performance, mitigating teachers' stereotyping behaviors.
- Relevance to the literature:
  - Pedagogical interventions that match teaching to students' learning "works" (Kremer et al. 2011; USAID 2011; Banerjee et al. 2016, NBER No. 22746; Banerjee et al. 2007; 3ie 2016)
- Policy contribution:
  - Complementary to lecture-style education for solving learning crisis

BRAC Primary School



↑ BEP uses National Curriculum and Textbook Board (NCTB) textbooks from grade four to five and grade one to three have their own textbooks followed by the National Curriculum.

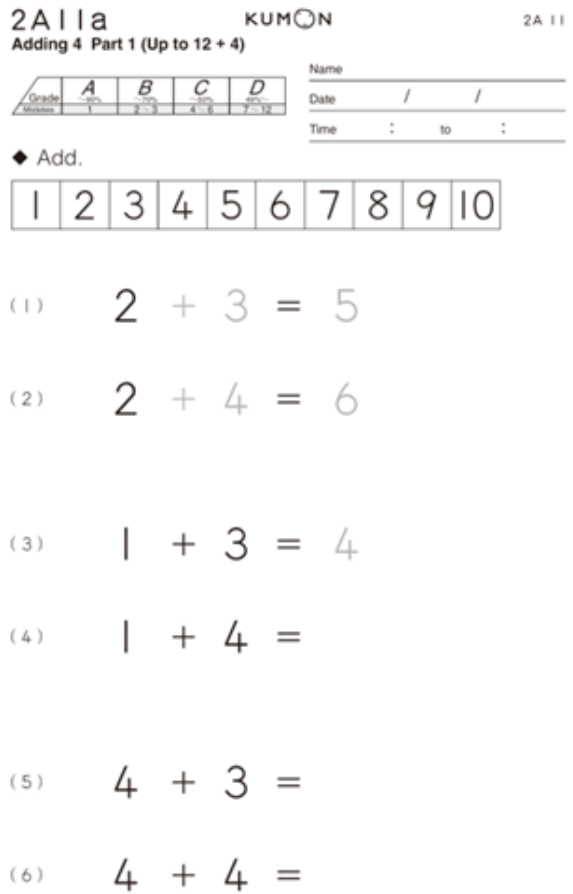
A BPS in Bangladesh during a Kumon session →

Kumon

- Diagnosis test at the very begining to identify each level
- Small-steps self-learning at the right level

↑ A Kumon center in India

An example of a worksheet →



2A II a KUMON 2A II

Adding 4 Part 1 (Up to 12 + 4)

Name

Grade A B C D

Date / /

Time : to :

◆ Add.

1	2	3	4	5	6	7	8	9	10

(1) 2 + 3 = 5

(2) 2 + 4 = 6

(3) 1 + 3 = 4

(4) 1 + 4 =

(5) 4 + 3 =

(6) 4 + 4 =

3A I a

Numbers up to 100 Part 1

KUMON

Grade	A	B	C	D
Minutes	1	2~3	4~5	6~11

Name

Date / /

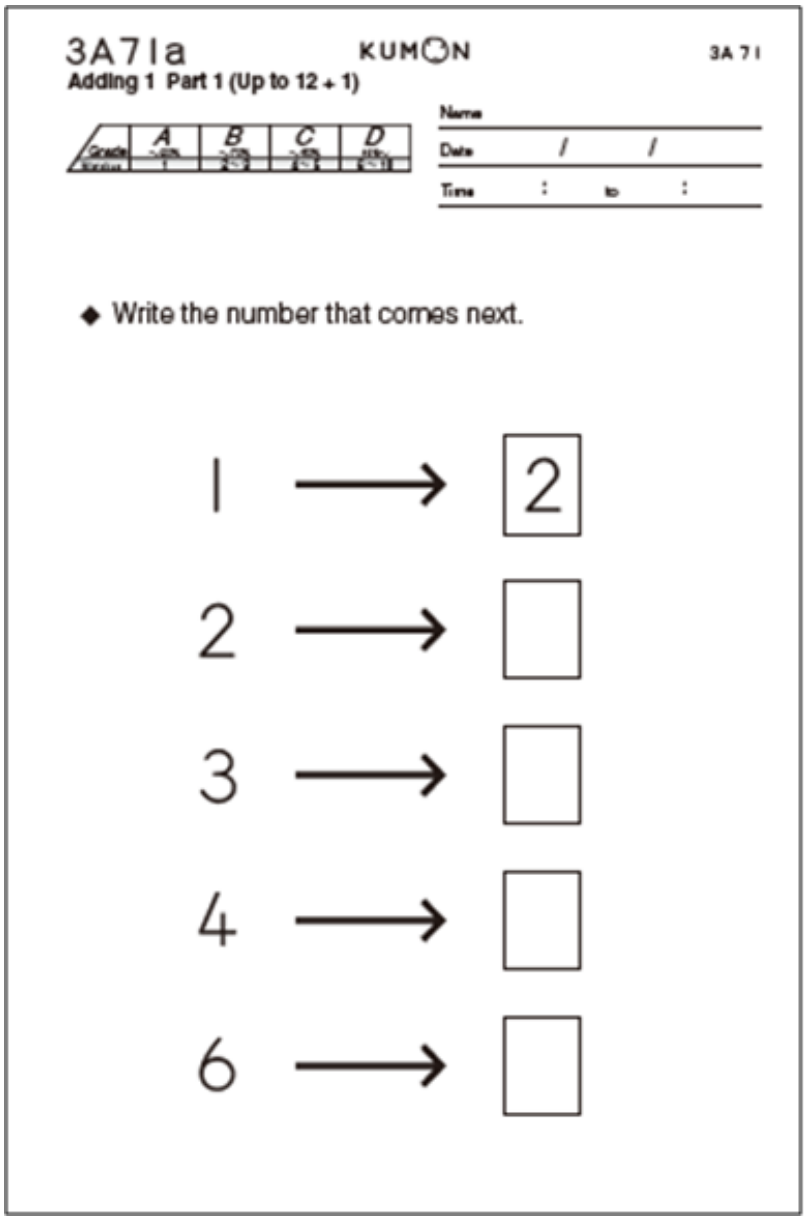
Time : to :

3A I

Write the numbers.

1	2	3	4	5	6	7	8	9	10

1	2	3	4	5					



3A71a KUMON 3A 71

Adding 1 Part 1 (Up to 12 + 1)

Grade A B C D

Name

Date / /

Time : to :

◆ Write the number that comes next.

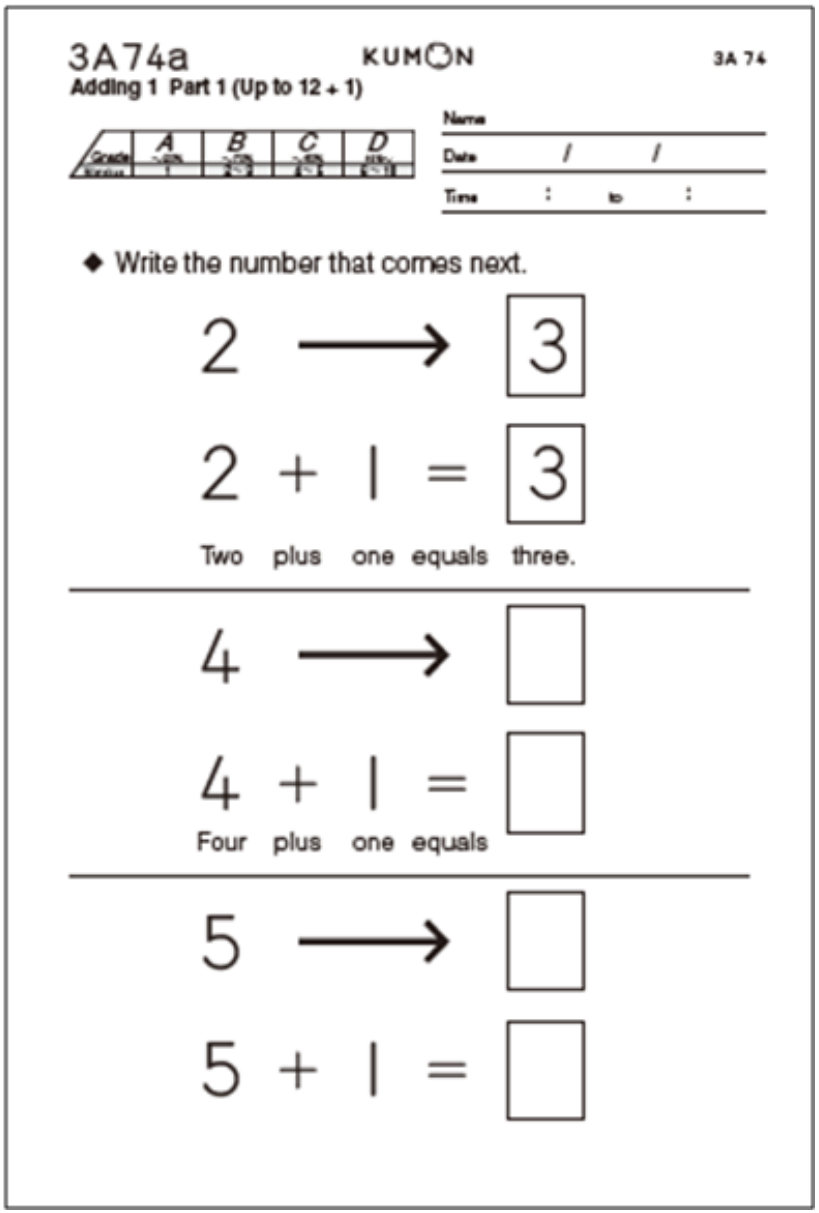
1 ⟶ 2

2 ⟶ ☐

3 ⟶ ☐

4 ⟶ ☐

6 ⟶ ☐



3A74a KUMON 3A 74

Adding 1 Part 1 (Up to 12 + 1)

Grade A B C D

Name

Date / /

Time : to :

◆ Write the number that comes next.

2 ⟶ 3

2 + 1 = 3

Two plus one equals three.

4 ⟶ ☐

4 + 1 = ☐

Four plus one equals

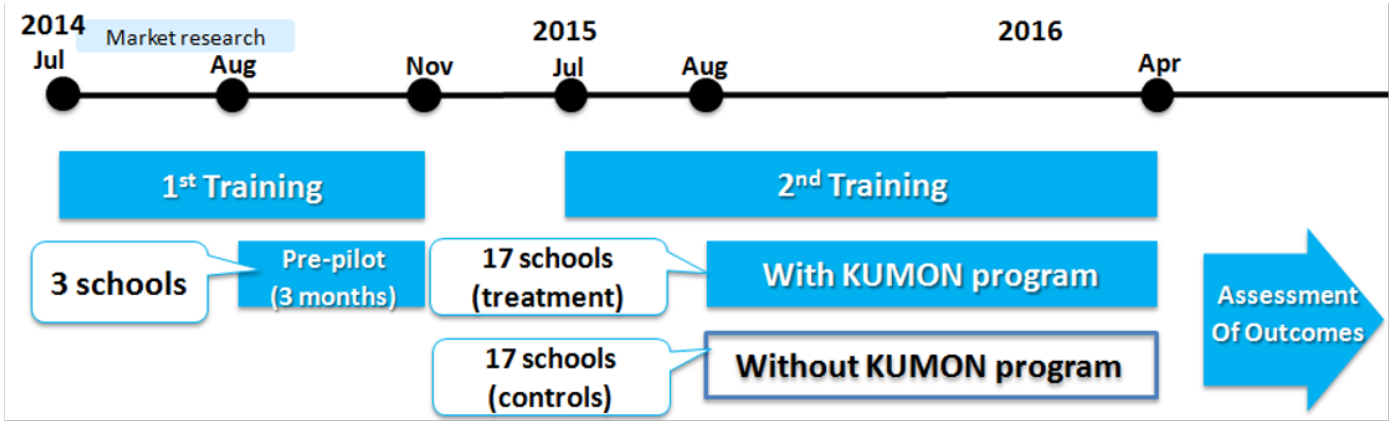
5 ⟶ ☐

5 + 1 = ☐

Implementation of Kumon in BPS and Research Design

- RCT to evaluate the applicability of the Kumon Method of Learning in strengthening the cognitive and non-cognitive abilities of the disadvantaged students studying at BPS:
  - 34 schools (30 students per school) are randomly selected out of 179 schools in 4 branches of BEP with grade specific strata.
  - RCT (17 treatment & 17 control)
  - Around 1,000 students and 34 teachers in the 3rd or 4th grades
  - 8 months tracking data
- Kumon×BEP interventions:
  - 30 min Kumon program before regular classes everyday
  - 2 Marking Assistants were employed per school.
  - Field Staff were assigned.
  - 3 days of preparatory training for BPS Teachers and Field Staff

Study Timeline



<Tests, Questionnaire, and Hearing>

↔ and ● : Timing or period of data collection ○ : Planned time to collect data of some schools

No.	Name of survey	To whom	The numbers of people*	In charge of the data	2015					2016									2017	
					Aug.	Sep.	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.	Apr.	May	Jun.	...	Nov.	...	...	Nov.
1	DT test	Students	About 1,000	KUMON	●								●							
2	PTS II	Students	About 1,000	KUMON	●								●							
3	Quizzes	Students	About 1,000	U-Tokyo				●		●										
4	Teachers' survey	Teachers	34 + 7**	KUMON		●									●					
5	parent guardian survey	Parents	About 1,000 Households	KUMON												●				
6	PSC	Students	226***	U-Tokyo													●			○

<Study Report>

7	Monthly Reports	Students	About 500	KUMON	↔								↔							
8	Grading books	Students	About 500	U-Tokyo	↔								↔							

*For 1-6, the number includes people in the control group.
**Some teachers were replaced at the time of endline.
***Excluding the number of samples planned to collect in 2017.

Impacts on Cognitive and Non-Cognitive Abilities

- ANCOVA:

$$Y_{it} = \alpha + \beta Y_{it-1} + \delta d_i + \varepsilon_{it} \tag{1}$$

$$Y_{it} = \alpha + \beta Y_{it-1} + \delta_0 d_i + \sum_j \delta_j d_i \cdot X_{ij} + \varepsilon_{it} \tag{2}$$

where

- Y_{it} = Cognitive or non-cognitive outcome of student i at t
- X = Student's baseline cognitive and/or non-cognitive score

Teacher's Assessment Ability

- Association b/w DT test/PTSII test (Y) and evaluation of students in teacher's survey (X):

$$\begin{aligned} Y_{it} = & (\alpha_{00} + \beta_{00} X_{it})(1 - d_i)(1 - T_t) \\ & + (\alpha_{10} + \beta_{10} X_{it}) d_i (1 - T_t) \\ & + (\alpha_{01} + \beta_{01} X_{it})(1 - d_i) T_t \\ & + (\alpha_{11} + \beta_{11} X_{it}) d_i T_t \\ & + u_i + \varepsilon_{it} \end{aligned} \tag{3}$$

- No difference in baseline: $\beta_{00} = \beta_{10}$
- Mitigation of teacher's stereotyping behaviors: $\beta_{11} > \beta_{01}$

Balancing Test

Balancing Test Results

Dependent Variables	DT per min (1)	DT Score (2)	DT Time (3)	PTSII Score (4)	RSES (5)	CPCS (6)	GRIT (7)	DT per min (8)	DT Score (9)	DT Time (10)	PTSII Score (11)	RSES (12)	CPCS (13)	GRIT (14)
Panel A: Grade 3														
Treatment vs. Control	**0.540****	**0.631***	**-0.149**	**-0.226**	**0.162**	**0.268**	**0.165**	**0.417**	**0.627**	**0.120**	**-0.291**	**0.277**	**0.313**	**0.0862**
	(0.164)	(0.245)	(0.0688)	(0.239)	(0.202)	(0.194)	(0.133)	(0.221)	(0.253)	(0.131)	(0.278)	(0.155)	(0.172)	(0.161)
Constant	0.0614	0.289	0.490	0.0759	-0.279	-0.323***	-0.119	-0.627**	-0.555	0.0175	-0.171	-0.382	-0.0147	0.436
	(0.378)	(0.326)	(0.349)	(0.424)	(0.115)	(0.0900)	(0.370)	(0.389)	(0.425)	(0.506)	(0.506)	(0.507)	(0.442)	(0.680)
Other Control Variables	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	582	582	582	582	582	582	582	582	582	582	582	582	582	582
R-squared	0.068	0.068	0.252	0.369	0.062	0.040	0.060	0.165	0.177	0.351	0.442	0.155	0.136	0.176
Panel B: Grade 4														
Treatment vs. Control	**-0.235**	**0.102**	**-0.246**	**-0.241**	**-0.178**	**-0.0921**	**0.213**	**-0.308**	**-0.0373**	**-0.380**	**-0.375**	**0.156**	**0.155**	**0.392**
	(0.185)	(0.0835)	(0.188)	(0.197)	(0.158)	(0.158)	(0.121)	(0.267)	(0.138)	(0.277)	(0.205)	(0.225)	(0.248)	(0.179)
Constant	0.256**	-0.0596	0.278	0.0732	0.421***	0.420***	0.152	-0.880	0.217	-0.884	-0.310	1.565*	1.467**	-0.577
	(0.222)	(0.113)	(0.210)	(0.178)	(0.150)	(0.149)	(0.100)	(0.850)	(0.403)	(0.873)	(0.660)	(0.506)	(0.456)	(0.551)
Other Control Variables	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	422	422	422	422	422	422	422	422	422	422	422	422	422	422
R-squared	0.086	0.034	0.070	0.299	0.060	0.070	0.059	0.232	0.141	0.201	0.391	0.172	0.173	0.185

Asymptotic standard errors are shown in parentheses and are clustered at school level (34 clusters). Asterisks reflect significance level obtained by a cluster wild bootstrap-t procedure.
Control variables: branch dummies, number of days the head of household get sick, number of members in the household, number of adults in the household, number of members in the household finishing primary education, last income, number of male in the household, frequency discussing subjects with child, frequency discussing lessons with child in the past two weeks, frequency of reading story to child, do you agree that parents should help with a child's study, frequency of eating meat of fish, frequency of drinking milk, main source of household income, average income, availability of electricity, availability of gas connection, type of toilet, source of water, house ownership

1-1. Impacts on Cognitive Abilities

ANCOVA

Impacts of KUMON Program on Students' Outcomes

Dependent Variables	Grade 3				Grade 4				All Grade
	DT Score per Minute	DT Score	DT time	PTS Score	DT Score per Minute	DT Score	DT Time	PTS Score	PTS Score
Treatment	**1.993*** **	**0.515*** **	**-1.931*** **	**0.785** **	**2.447** **	**0.320**	**-2.693*** **	**1.213*** **	**0.907*** **
	(0.281)	(0.136)	(0.459)	(0.303)	(0.840)	(0.230)	(0.855)	(0.207)	(0.203)
Baseline Score	0.330***	0.139**	0.296***	0.342**	0.418**	0.127*	-0.0420	0.339***	0.329***
	(0.0803)	(0.0543)	(0.0944)	(0.124)	(0.166)	(0.0702)	(0.142)	(0.0816)	(0.0778)
Constant	-0.445	1.281**	0.0350	0.859**	3.149	0.0796	-3.235	-0.340	0.0209
	(1.110)	(0.593)	(0.967)	(0.332)	(2.103)	(0.402)	(2.068)	(0.276)	(0.264)
Other Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Num of Obs.	473	473	473	501	338	338	338	336	837
R-squared	0.344	0.220	0.401	0.231	0.369	0.271	0.371	0.375	0.301

Asymptotic standard errors are shown in parentheses and are clustered at school level (34 clusters).
Asterisks reflect significance level obtained by a cluster wild bootstrap-t procedure.
Control variables: branch dummies, number of days the head of household get sick, number of members in the household, number of adults in the household, number of members in the household finishing primary education, last income, number of male in the household, frequency discussing subjects with child, frequency discussing lessons with child in the past two weeks, frequency of reading story to child, do you agree that parents should help with a child's study, frequency of eating meat of fish, frequency of drinking milk, main source of household income, average income, availability of electricity, availability of gas connection, type of toilet, source of water, house ownership

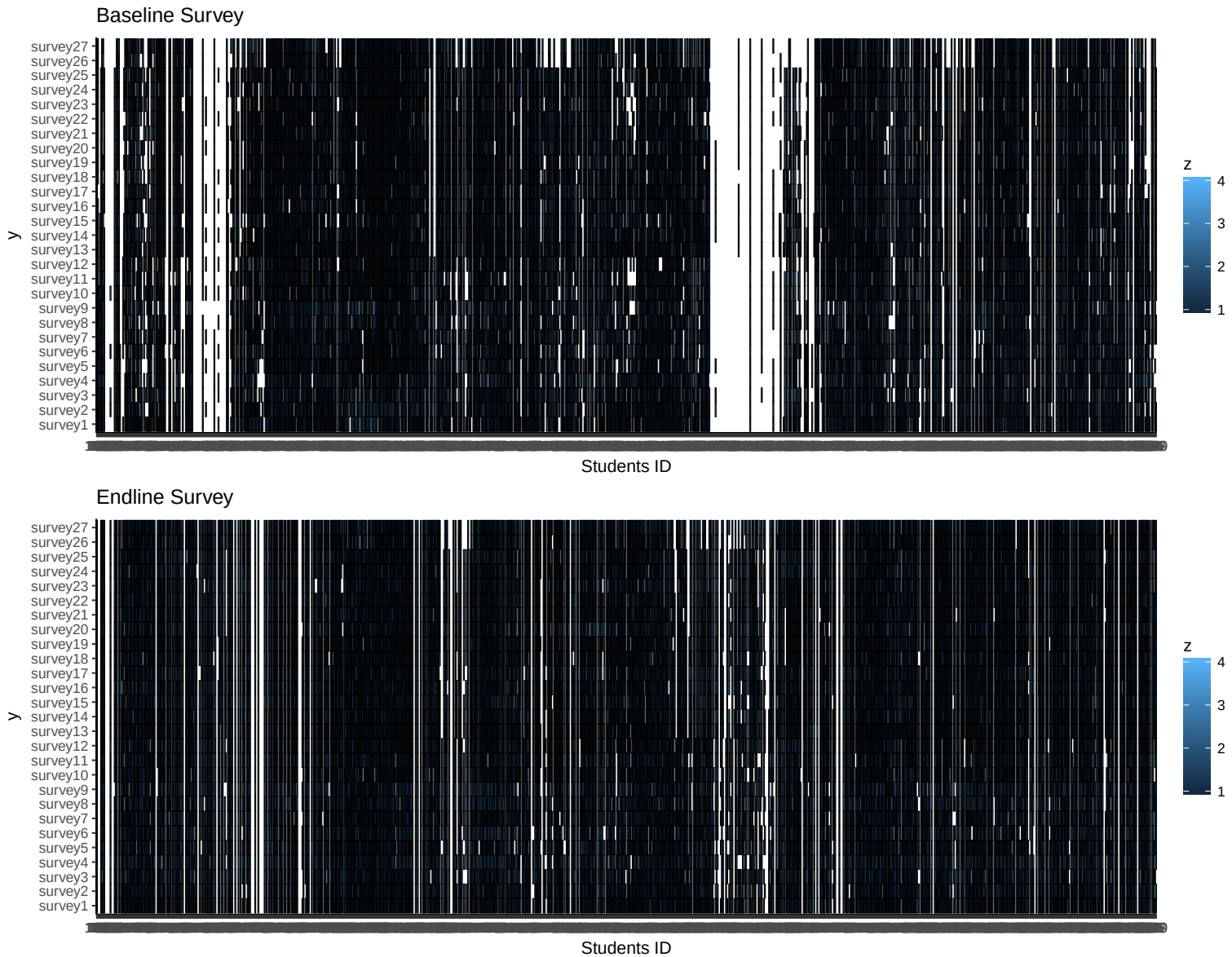
2-1. Impacts on Non-Cognitive Abilities
ANCOVA

Impacts of KUMON Program on Students' Outcomes

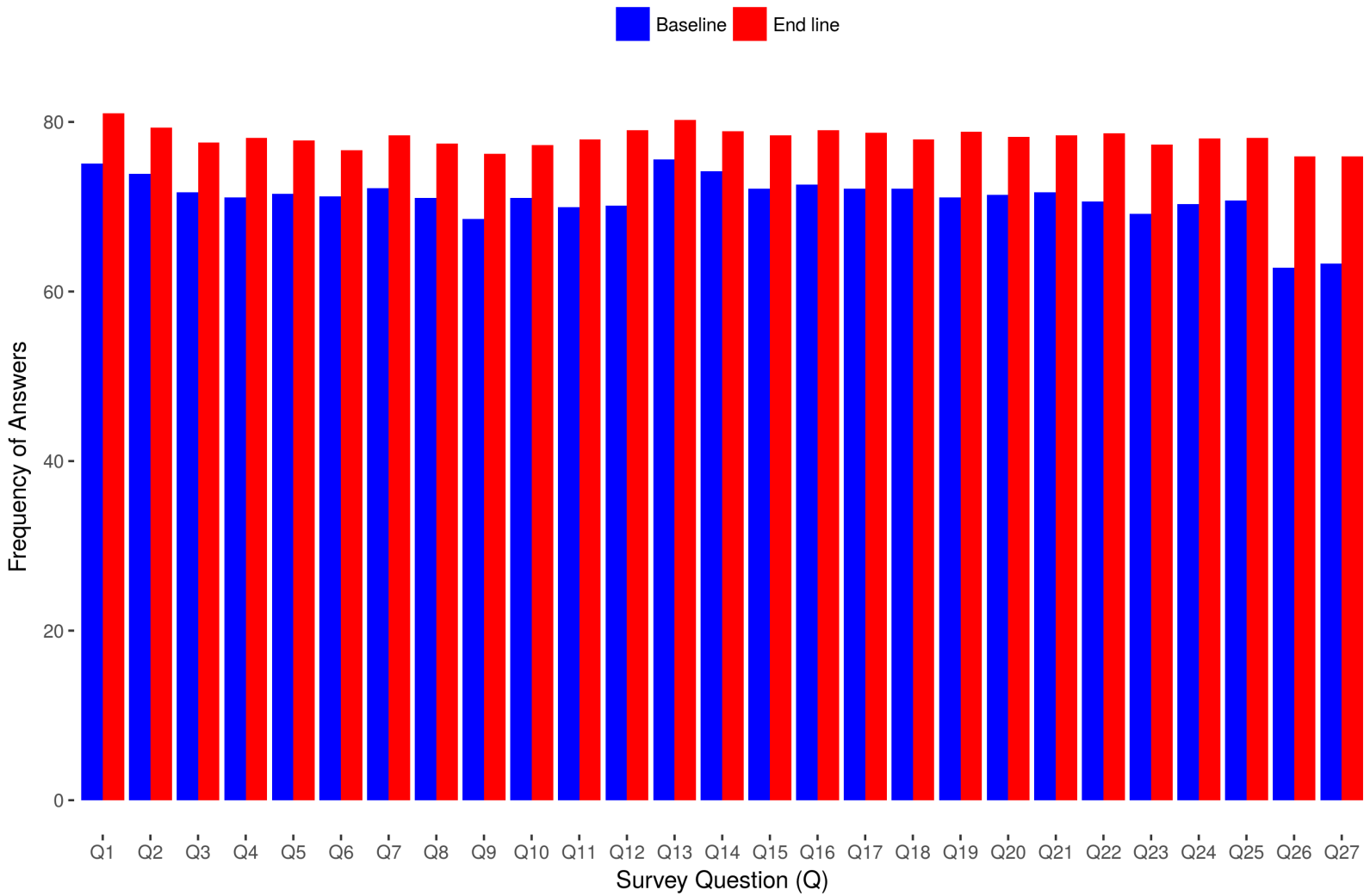
Dependent Variables	Grade 3			Grade 4			All grades		
	Rosenberg	CPCS	GRIT	Rosenberg	CPCS	GRIT	Rosenberg	CPCS	GRIT
Treatment	**-0.0742**	**0.0118**	**0.0231**	**0.218**	**0.279**	**-0.176**	**-0.0258**	**0.0600**	**-0.00912**
	(0.161)	(0.154)	(0.134)	(0.190)	(0.188)	(0.153)	(0.132)	(0.126)	(0.0904)
Baseline Score	0.0529	0.0697	0.0544*	0.159**	0.140**	0.102*	0.105**	0.104**	0.0684**
	(0.0685)	(0.0471)	(0.0312)	(0.0667)	(0.0620)	(0.0545)	(0.0473)	(0.0385)	(0.0273)
Constant	0.779	0.596	-0.213	-0.535	-0.376	0.379	-0.0272	0.118	0.559
	(0.462)	(0.376)	(0.380)	(0.554)	(0.587)	(0.528)	(0.360)	(0.351)	(0.361)
Other Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Num of Obs.	496	496	493	336	336	324	832	832	817
R-squared	0.122	0.119	0.134	0.169	0.165	0.129	0.088	0.090	0.092

Asymptotic standard errors are shown in parentheses and are clustered at school level (34 clusters).
Asterisks reflect significance level obtained by a cluster wild bootstrap-t procedure.
Control variables: branch dummies, number of days the head of household get sick, number of members in the household, number of adults in the household, number of members in the household finishing primary education, last income, number of male in the household, frequency discussing subjects with child, frequency discussing lessons with child in the past two weeks, frequency of reading story to child, do you agree that parents should help with a child's study, frequency of eating meat of fish, frequency of drinking milk, main source of household income, average income, availability of electricity, availability of gas connection, type of toilet, source of water, house ownership

Missing Value Pattern in Non-Cognitive Survey Questions



Frequency of Answers to Non-Cognitive Survey Questions



2-2. Individual Non-Cognitive Survey Questions

Impacts of KUMON Program Individual Survey Questions

Dependent Variables	Self-confidence (Question 1)[a]				Self-confidence (Question 7)[b]			
	End line (1)	End line (2)	ANCOVA (3)	ANCOVA (4)	End line (5)	End line (6)	End line (7)	ANCOVA (8)
Treatment vs. Control	**0.336****	**0.331****	**0.333****	**0.358****	**0.301****	**0.373****	**0.291****	**0.393*****
	(0.134)	(0.146)	(0.132)	(0.153)	(0.118)	(0.142)	(0.119)	(0.138)
Baseline Score			0.0187	0.0200			0.0263	0.0556
			(0.0509)	(0.0516)			(0.0493)	(0.0513)
Treatment x Initial Cognitive Score (=1 if above median)		-0.024		-0.0597		0.0702		0.0541
		(0.132)		(0.142)		(0.113)		(0.115)
Treatment x Initial Noncognitive Score (=1 if above median)		0.0312		-0.00128		-0.208*		-0.275**
		(0.0932)		(0.0983)		(0.114)		(0.122)
Constant	-0.171	-0.171	-0.138	-0.136	-0.153	-0.153	-0.135	-0.117
	(0.103)	(0.103)	(0.0968)	(0.0967)	(0.0911)	(0.0912)	(0.0990)	(0.100)
Num of Obs.	819	819	819	819	793	793	793	793
R-squared	0.028	0.028	0.032	0.033	0.023	0.028	0.024	0.032

Asymptotic standard errors are shown in parentheses and are clustered at school level (34 clusters).
Asterisks reflect significance level obtained by a cluster wild bootstrap-t procedure.
Note [a]: The level of agreement to the statement "I did well in the test", where responses were recorded on four-point scale:
Strongly Agree, Somewhat Agree, Somewhat Disagree, Strongly Disagree.
Note[b]: The level of agreement to the statement "I can confidently express my opinion", where responses were recorded on four-point scale:
Strongly Agree, Somewhat Agree, Somewhat Disagree, Strongly Disagree.

3-1. Control for Longer Session

ANCOVA

Impacts of KUMON Program on Students' Cognitive Outcomes – Cross-sectional Heterogenous Specification & Longer Session

Dependent Variables	Grade 3					Grade 4					All grades
	DT_per_min	PTS Score	Rosenberg	CPCS	GRIT	DT_per_min	PTS Score	Rosenberg	CPCS	GRIT	PTS Score
	(1)	(4)	(5)	(6)	(7)	(8)	(11)	(12)	(13)	(14)	(15)
Treatment vs. Control	**2.301*** **	**0.868** **	**0.0210**	**0.166**	**0.277***	**3.079** **	**0.990*** **	**0.305**	**0.409**	**0.0769**	**0.891*** **
	(0.445)	(0.339)	(0.224)	(0.198)	(0.141)	(1.200)	(0.297)	(0.291)	(0.292)	(0.157)	(0.241)
Treatment x Longer session	**-0.301**	**-0.347**	**-0.113**	**-0.188**	**-0.404**	**-1.763**	**0.132**	**-0.209**	**-0.362**	**-0.442** **	**-0.110**
	(0.549)	(0.450)	(0.325)	(0.306)	(0.309)	(1.222)	(0.285)	(0.333)	(0.308)	(0.156)	(0.310)
Constant	0.723***	1.146***	0.00852	-0.0561	-0.0670	0.979***	0.505***	-0.126	-0.139	0.0490	0.865***
	(0.138)	(0.164)	(0.119)	(0.130)	(0.116)	(0.245)	(0.107)	(0.119)	(0.101)	(0.0954)	(0.119)
Num of Obs.	473	501	496	496	493	338	336	336	336	324	837
R-squared	0.251	0.110	0.002	0.006	0.027	0.219	0.222	0.017	0.033	0.021	0.140

Asymptotic standard errors are shown in parentheses and are clustered at school level (34 clusters).
Asterisks reflect significance level obtained by a cluster wild bootstrap-t procedure.

4-1. Grade 4 Primary School Certificate Math Results

Summary Statistics

	Treatment	Control	Difference
Probability of taking PSC	0.542	0.529	0.014
			(0.095)
Baseline Characteristics			
DT Score	-0.067	0.355	-0.422*
			(0.223)
PTS Score	-0.167	0.549	-0.708
			(0.433)
RSES	0.055	0.064	-0.011
			(0.231)
CSPS	0.100	-0.014	0.114
			(0.236)
GRIT	0.166	-0.128	0.294
			(0.158)
Male	0.322	0.324	-0.002
			(0.075)
Wealth Index	-0.039	0.131	-0.170
			(0.253)
Number of Observations	115	111	

4-2. Grade 4 PSC Mathematics Results

Propensity Score Matching

Math PSC	Treatment	Control	Difference	PSM
Probability of achieving at least A	0.496	0.838	-0.342	0.0435
			(0.211)	(0.0789)
Probability of achieving at least B	0.591	0.874	-0.283	0.113
			(0.186)	(0.0891)
Probability of achieving at least C	0.748	0.919	-0.171	0.217**
			(0.156)	(0.0853)
Probability of passing (at least D)	0.887	0.990	-0.104	-0.0696
			(0.071)	(0.100)
Number of Observation	115	111		

5-1. Impacts on Teachers' Assessent Abilities

Association between Teachers' assessment and students' performance			
Dependent Variables	All Sample	Grade 3	Grade 4
DT Score per Minute			
Teacher_evaluation*(1-Treatment)*(1-Endline)	0.386***	0.362***	0.373***
	(0.0550)	(0.0666)	(0.0874)
Teacher_evaluation*Treatment*(1-Endline)	0.177**	0.0717	0.400*
	(0.0854)	(0.0534)	(0.211)
Teacher_evaluation*(1-Treatment)*Endline	0.479***	0.567***	0.290***
	(0.0640)	(0.0711)	(0.0731)
Teacher_evaluation*Treatment*Endline	0.623***	0.547**	1.399**
	(0.222)	(0.242)	(0.565)
Control_Baseline = Treatment_Baseline	**1.22**	**4.40***	**0.53**
Control_Endline = Treatment_Endline	**3.53***	**3.68***	**2.75**$^{+}$
Num of Obs.	1,268	792	476
R-squared	0.531	0.595	0.510
PTS Score			
Teacher_evaluation*(1-Treatment)*(1-Endline)	0.124	-0.0345	0.351***
	(0.201)	(0.271)	(0.108)
Teacher_evaluation*Treatment*(1-Endline)	0.155	0.169	0.294***
	(0.122)	(0.132)	(0.0926)
Teacher_evaluation*(1-Treatment)*Endline	0.306***	0.321***	0.270***
	(0.0490)	(0.0626)	(0.0538)
Teacher_evaluation*Treatment*Endline	0.545***	0.488***	0.611***
	(0.0829)	(0.0969)	(0.0989)
Control_Baseline = Treatment_Baseline	**0.77**	**1.64**	**0.45**
Control_Endline = Treatment_Endline	**6.94****	**3.79***	**5.48****
Num of Obs.	1,292	822	470
R-squared	0.532	0.585	0.469

$+$: p-value = 0.1213

Conclusion

- Non-formal education has potential to contribute to better learning outcomes.
- This study finds robust improvements in students' cognitive by the Kumon Method of Learning among the students of BRAC Primary Schools.
- Through this intervention, teachers become better able to predict students' performance. This could be one of the potential paths in improving student learning outcomes against teachers' stereotyping behaviors (Ferguson 1998, Urban Education; Lavy 2008, JPubE; Hanna et al. 2012, AEJ; Botleho et al. 2015, AEJ).
- This evidence contributes to the literature that finds pedagogical interventions that match teaching to students' learning "works" (Kremer et al. 2011, AER; Banerjee et al. 2016, NBER No. 22746; Banerjee et al. 2007, QJE; 3ie 2016: 27–28).

PTSII Survey Questions

Non-Cognitive Ability Measurements

Number	Question in English	CPCS	RSES	GritScale
1	I did well in this test.			
2	I can do most things better than other people.	x	x	
3	There are many things about myself I can be proud of.	x	x	
4	I feel that I cannot do anything well no matter what I do.	x	x	
5	I believe I can be someone great.	x		
6	I don't think I am a helpful person.	x	x	
7	I can confidently express my opinion.	x		
8	I don't think I have that many good qualities.	x	x	
9	I am always worried that I might fail.	x	x	
10	I am confident about myself.	x	x	
11	I am satisfied with myself.	x	x	
12	Even if I fail, I think I can get better and better at things if I keep trying			
13	I like to do calculations.			x
14	I can calculate in my head when I go shopping.			x
15	I think speed is important when solving problems.			x
16	When studying, I believe everything will go well if I correctly follow instruction			
17	I am more motivated when people praise me.			
18	I always volunteer in class.			
19	I enjoy studying.			
20	School is fun.			
21	I do things better when I have a goal.			
22	There are many things I want to learn more about.			
23	a. I have a role model around me. b. There is someone around me who I want to be like.			
24	I always have someone who I can go to for advice when I am having trouble with my studies.			
25	a. There is someone around me who I don't want to lose against. b. There is someone around me who I am always competing with.			
26	I always try to do something when things don't go as expected.			
27	It doesn't matter whether I fail in the beginning because I believe that things will eventually work out.			

Children's Perceived Competence Scale: CPCS

- Developmental psychologist, Harter invented it in 1979
  - Japanese version was made and used widely in the field of social psychology
- Competence (White, 1959)
  - Realising their own active work as "I can do it!"
- 28 contents, 4 choices. (1.Strongly Agree 2.Agree 3.Disagree 4.Strongly Disagree)

Rosenberg Self-Esteem Scale: RSES

- Rosenberg invented this in 1965; measuring generally and quantitatively the extent of self-esteem
- Used in many countries and also for comparing internationally
- 10 contents, 4 choices. (1.Strongly Agree 2.Agree 3.Disagree 4.Strongly Disagree)
  - The higher the score is, the more positively they think themselves entirely and more highly they evaluate themselves

Grit Scale

- Ability for setting long-term goals and not-giving-up till they are achieved using any methods and trying as hard as possible
  Huge impact on children's growth and future success
- Angela Lee Duckworth, "Grit: The power of passion and perseverance"
- 12 contents, 5 choices. (1 Very much like me 2 Mostly like me 3 Somewhat like me 4 Not much like me 5 Not like me at all)