
Through the looking glass: can classroom observation and coaching improve teacher performance in Brazil?

BARBARA BRUNS (CENTER FOR GLOBAL DEVELOPMENT, VISITING FELLOW)

LEANDRO COSTA (WORLD BANK, ECONOMIST)

NINA CUNHA (STANFORD UNIVERSITY, PHD CANDIDATE)

Outline

1. Motivation
2. Intervention and experiment design
 - Research questions
3. Instruments and Data
 - The Stallings Instrument
 - Sample and Balance checks
4. Results
 - Descriptive statistics
 - Intention to treat effects
 - Partial Compliance
5. Experiment Threats and Robustness checks
 - Attrition
 - Spillover
 - Treatment Contamination
 - Evaluation-Driven Effects
6. Conclusions

Motivation

- How to improve teachers' classroom effectiveness?
- “Observable” teacher characteristics do not explain differences in individual teachers’ ability to produce classroom level learning gains
- Large variations in teachers’ practice and classroom-level results (Hanushek and Rivkin, 2010; Rockoff, 2004; Araujo et al, 2016; Bruns and Luque 2014)
- Quality of teachers’ classroom practice, as measured through classroom observations, is important for student learning and students’ socio-emotional skills (MET, Kane and Staiger, 2012; Araujo et al, 2016; Howes et al, 2008; Grossman et al, 2010; Chetty et al, 2014; Jackson et al, 2014; Jennings and DiPrete, 2010)
- Differences in teachers’ classroom practice are not explained by teacher background and status (Kane and Staiger, 2012; Araujo et al, 2016)

Intervention and experiment design

- The intervention in the northeast Brazilian state of Ceará was designed to improve the effectiveness of teachers in service
- Partnership between Seduc/CE, Lemann Foundation and WB/SIEF
- Program design:
 - Information “shock” (benchmarked feedback) and
 - Expert coaching to promote increased professional interaction among teachers in the same school
 - Self-help materials
- The program design benefits from the large variation in teacher quality *within* schools
- The exchange of practice among teachers (Japan’s lesson study (Easton, 2008; Lewis et al 2004), Singapore (OECD, 2013) and Shanghai (Liang, 2016))
- “Professional learning community” within the school (OECD, 2011; Mourshed et al. 2010; Fullan, 2013)

Experiment Design - RCT

Baseline						Endline and final Analysis	
Stallings training	Baseline observations	Analysis	Toolkit development	Randomize	Treatments Implemented	Stallings observations	Analysis
Oct. 2014	November 2014	Dec. 2014	Sept. 2014- Feb. 2015	Feb. 2015	March-October 2015	Nov. 2015	Jan-Mar 2016

RCT – Randomized Control Trial

Stratified Random Sample - A representative sample of **350** schools randomly chosen from among 573 state schools, stratified by school size, students learning, and schools regional. Each school has at least 6 teachers and 20 students at 10th , 11th and 12th grades.

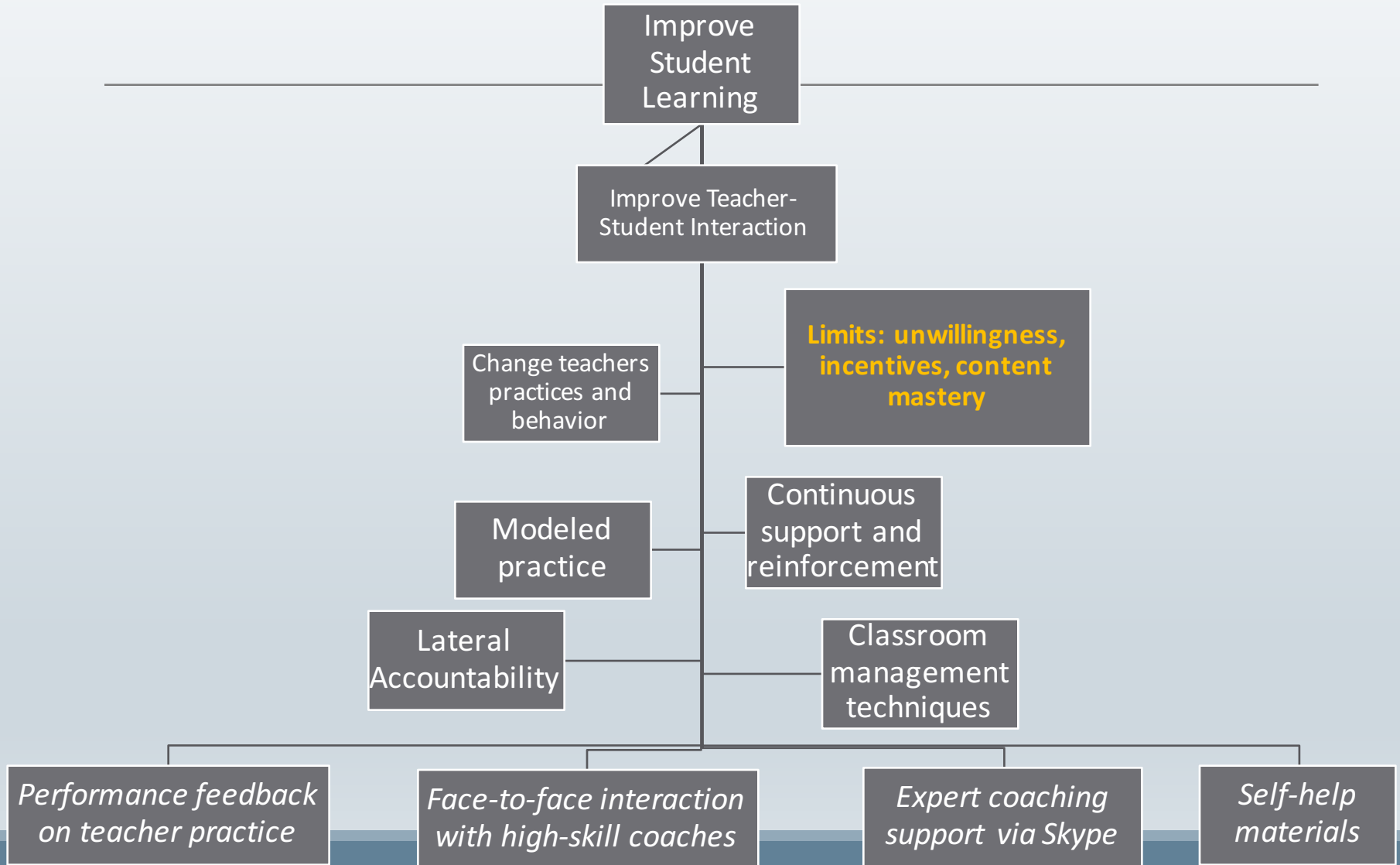
Treatment group - 175 schools were assigned to the treatment group through simple randomization.

Intervention - i) Performance feedback on teacher practice, ii) Self-help materials, iii) Face-to-face interaction with high-skill coaches, and iv) Expert coaching support via Skype.

Control group – 175 schools were assigned to the control group through simple randomization.

Intervention - no treatment during 2015.

Theory of Change



Logical Chain

Inputs	Activities	Outcomes	Impacts
<ul style="list-style-type: none"> - 175 Bulletins with results from classroom observations - 4680 books “Teach Like A Champion” for 175 schools - Lemann Foundation face-to-face training (ELOS trainers) - Lemann Foundation skype experts coaching (ELOS coaches) - SEDUC/CE logistic support - SEDUC/CE coordination - SEDUC/CE teachers training policy - Pedagogical coordinators time 	<ul style="list-style-type: none"> - School-level professional interaction among teachers - <i>Face-to-face interaction with high-skill coaches</i> - <i>Expert coaching support via Skype</i> 	<ul style="list-style-type: none"> - Improve teachers practices and behavior - Improve teachers use of time for instruction - Improve teachers use of pedagogical materials 	<ul style="list-style-type: none"> - Improve teachers-student interaction - Improve students engagement - Improve students learning

Cost and Resources

COST ELEMENT	RS	US\$	RS/STUDENT	US\$/STUDENT
Program Costs				
Classroom observations in 165 treatment schools – Nov 2014	536,000.00	134,000.00	4.4	1.1
Transport, lodging, subsistence for 400 participants at 4 face-to-face training sessions	152,000.00	38,000.00	1.2	0.3
Aula Nota 10 book for 175 schools	117,000.00	29,250.00	1.0	0.2
ELOS training team (xx hours of coaching support)	468,000.00	117,000.00	3.8	1.0
SUBTOTAL	1,273,000.00	318,250.00	10.3	2.6
Evaluation Costs				
Classroom observations in 175 control schools in Nov 2014 and 292 schools in Nov 2015	456,000.00	114,000.00	3.7	0.9
SUB TOTAL	456,000.00	114,000.00	3.7	0.9
GRAND TOTAL	1,729,000.00	432,250.00	14.1	3.5
RESOURCES	RS	US\$		
SEDUC/CE	197,518.10	49,379.53		
Lemann Foundation	624,858.39	156,214.60		
World Bank	906,623.51	226,655.88		

Research Questions

- Can providing schools with individualized feedback based on classroom observations plus support materials and coaching stimulate measurable changes in teacher practice in a relatively short period (a single school year)?
- Can providing classroom observation feedback and coaching for pedagogical coordinators reduce variation in teacher practices within a school?
- Can providing classroom observation feedback and coaching for pedagogical coordinators improve student test performance? Is the combined program developed in Ceará (classroom observation feedback and school-level coaching) cost-effective in producing learning results when compared with alternative teacher training programs?

Instruments and Data

- Stalling classroom observation sheet, demographic sheet, and principal questionnaire
 - Stallings observations are coded at ten different moments in every class, at exact intervals whose spacing depends on the length of the class; every 5 minutes in a 50-minute class, etc.
 - Each observation consists of a 15 second scan of the classroom, starting with the teacher and proceeding clockwise around the room. Observers code what the teacher is doing; what materials s/he is using and what the students are doing.
1. **Instruction:** Reading Aloud; Demonstration/Lecture; Discussion/Debate/Question and Answer; Practice & Drill; Assignment/Class Work; Copying
 2. **Classroom Management:** Verbal Instruction; Disciplining students; Classroom Management with Students; Classroom Management Alone
 3. **Teacher Off-Task:** Teacher in Social Interaction with Students; Teacher in Social Interaction with Outsiders or Teacher Uninvolved; Teacher out of the classroom
 4. **Students Off-Task:** Students being disciplined; students in Social Interaction; Student(s) Uninvolved

CLASSROOM OBSERVATION SNAPSHOT

ACTIVITY	MATERIAL												
	NO MATERIAL	TEXTBOOK	NOTEBOOK	BLACK BOARD	LEARNING AIDES/MANIPULATIVES	ICT	COOPERATIVE						
1. READING ALOUD	T	1 S L E	1 S L E	1 S L E	1 S L E	1 S L E	1 S L E	S L E					
	S	1 S L	1 S L	1 S L	1 S L	1 S L	1 S L	S L					
2. DEMONSTRATION / LECTURE	T	1 S L E	1 S L E	1 S L E	1 S L E	1 S L E	1 S L E	S L E					
	S	1 S L	1 S L	1 S L	1 S L	1 S L	1 S L	S L					
3. DISCUSSION/ DEBATE/QUESTION & ANSWERS	T	1 S L E	1 S L E	1 S L E	1 S L E	1 S L E	1 S L E	S L E					
	S	1 S L	1 S L	1 S L	1 S L	1 S L	1 S L	S L					
4. PRACTICE & DRILL	T	1 S L E	1 S L E	1 S L E	1 S L E	1 S L E	1 S L E	S L E					
	S	1 S L	1 S L	1 S L	1 S L	1 S L	1 S L	S L					
5. ASSIGNMENT/ CLASS WORK	T	1 S L E	1 S L E	1 S L E	1 S L E	1 S L E	1 S L E	S L E					
	S	1 S L	1 S L	1 S L	1 S L	1 S L	1 S L	S L					
6. COPYING	T	1 S L E	1 S L E	1 S L E	1 S L E	1 S L E	1 S L E	S L E					
	S	1 S L	1 S L	1 S L	1 S L	1 S L	1 S L	S L					
7. VERBAL INSTRUCTION	T	1 S L E	1 S L E	1 S L E	1 S L E	1 S L E	1 S L E	S L E					
	S	1 S L	1 S L	1 S L	1 S L	1 S L	1 S L	S L					
8. SOCIAL INTERACTION	T	1 S L E	16. What subject is being taught? <input type="radio"/> English <input type="radio"/> Mathematics <input type="radio"/> Science <input type="radio"/> Other _____ Comments: 										
	S	1 S L											
9. STUDENT(S) UNINVOLVED	S	1 S L											
10. DISCIPLINE	T	1 S L E											
	S	1 S L											
11. CLASSROOM MANAGEMENT	T	1 S L E											
	S	1 S L											
12. CLASSROOM MANAGEMENT ALONE									T				
13. TEACHER SOCIAL INTERACTION OR TEACHER UNINVOLVED									T				
14. TEACHER OUT OF THE ROOM									T				

Balance Check

	Random Sample (350 Schools)			Baseline Data (292 Schools)		
	Control Means	Treatment Means	Difference	Control Means	Treatment Means	Difference
2013 Covariates						
Portuguese proficiency	257.4 [19.73]	260.8 [22.39]	-3.245 [2.259]	256.9 [18.69]	261.4 [23.08]	-4.454 [2.481]
Mathematical proficiency	267.4 [23.81]	272.2 [29.77]	-4.679 [2.882]	267.7 [22.67]	273.3 [30.72]	-5.562 [3.199]
High School enrollment	641.4 [368.2]	588.9 [330.3]	55.15 [37.44]	676.3 [349.3]	575.3 [321.5]	101.0* [39.27]
High school enrollment - vocational	46.63 [132.6]	68.21 [154.1]	-21.18 [15.35]	47.11 [136.0]	76.08 [160.9]	-28.97 [17.58]
Rural Area	0.0286 [0.167]	0.0517 [0.222]	-0.0229 [0.0210]	0.0368 [0.189]	0.0577 [0.234]	-0.0209 [0.0251]
Pass rate	83.33 [10.33]	84.56 [10.74]	-1.248 [1.125]	84.46 [10.07]	85.57 [10.50]	-1.115 [1.208]
Failure rate	6.938 [5.614]	6.311 [5.283]	0.649 [0.582]	6.398 [5.620]	6.051 [5.227]	0.347 [0.635]
Dropout rate	9.731 [7.179]	9.129 [7.002]	0.600 [0.757]	9.144 [6.896]	8.375 [6.637]	0.769 [0.793]
Students per class	34.06 [4.939]	34.00 [5.198]	0.0734 [0.541]	34.38 [4.941]	34.03 [5.317]	0.349 [0.604]
Female principals	0.520 [0.501]	0.511 [0.501]	0.00571 [0.0536]	0.485 [0.502]	0.519 [0.501]	-0.0339 [0.0588]
Experience as a principal (> 10 years)	0.543 [0.500]	0.517 [0.501]	0.0229 [0.0535]	0.507 [0.502]	0.500 [0.502]	0.00735 [0.0589]
Principal with graduate degree	0.994 [0.0756]	0.994 [0.0758]	0 [0.00808]	0.993 [0.0857]	0.994 [0.0801]	-0.000943 [0.00971]
Female teachers	0.551 [0.180]	0.515 [0.181]	0.0341 [0.0193]	0.562 [0.184]	0.515 [0.183]	0.0476* [0.0216]
Temporary teachers	0.995 [0.0148]	0.994 [0.0188]	0.00114 [0.00181]	0.995 [0.0155]	0.994 [0.0193]	0.000713 [0.00207]
Teacher's age	35.00 [27.09]	30.34 [63.98]	4.609 [5.239]	35.34 [25.52]	30.15 [67.22]	5.197 [6.117]
Experience as a teacher (>10 years)	0.816 [0.0871]	0.814 [0.0850]	0.00194 [0.00919]	0.819 [0.0858]	0.812 [0.0873]	0.00749 [0.0102]
Low salary (< 2m.w.)	0.185 [0.141]	0.184 [0.152]	0.000229 [0.0157]	0.194 [0.146]	0.183 [0.155]	0.0109 [0.0177]
High Salary (> 5 m.w.)	0.225 [0.179]	0.200 [0.183]	0.0253 [0.0194]	0.219 [0.183]	0.187 [0.179]	0.0327 [0.0212]
Mother's education < middle school	0.472 [0.104]	0.485 [0.108]	-0.0115 [0.0114]	0.490 [0.0966]	0.488 [0.109]	0.00159 [0.0122]
Mothers with graduate degree	0.0507 [0.0301]	0.0523 [0.0302]	-0.00143 [0.00322]	0.0548 [0.0282]	0.0546 [0.0305]	0.000228 [0.00345]
2014 Covariates						
Portuguese proficiency	252.8 [17.72]	256.5 [20.53]	-3.675 [2.053]	252.3 [17.76]	257.1 [21.24]	-4.764* [2.311]
Mathematical proficiency	252.8 [21.58]	258.8 [27.66]	-5.972* [2.655]	253.1 [21.79]	260.2 [28.59]	-7.082* [3.009]
Age-Grade distortion	33.72 [15.21]	32.06 [15.47]	1.662 [1.642]	31.63 [14.04]	30.66 [15.18]	0.964 [1.720]
Proportion of students per teacher	0.0588 [0.0214]	0.0593 [0.0215]	-0.000576 [0.00230]	0.0534 [0.0142]	0.0586 [0.0208]	-0.00526* [0.00212]
Proportion of black and brown teachers	0.298 [0.232]	0.302 [0.228]	-0.00400 [0.0246]	0.281 [0.238]	0.302 [0.231]	-0.0209 [0.0275]
Proportion of black and brown students	0.606 [0.216]	0.606 [0.230]	0.000215 [0.0239]	0.595 [0.220]	0.607 [0.229]	-0.0115 [0.0264]
Joint test (p-value) - All Variables			0.620			0.18
Joint test (p-value) - Only proficiency variables			0.120			0.13
Joint test (p-value) - Other variables excluding proficiency			0.850			0.31
Number of schools	175	175		136	156	
Response Rate				78%	89%	0.11
p-value of the response rate difference						0.00

Note: Numbers in parentheses are standard deviations in the column of means and standard errors in columns of differences. * p<0.05

** p<0.01 *** p<0.001

Balance Check

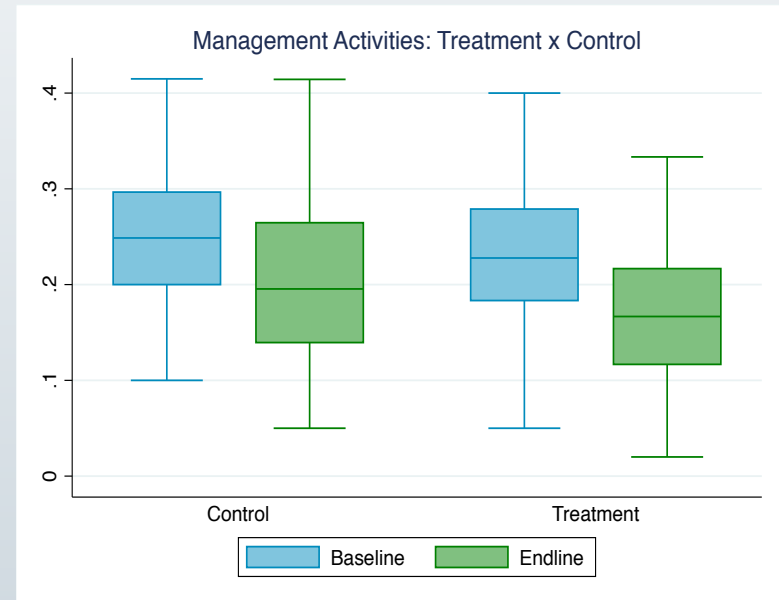
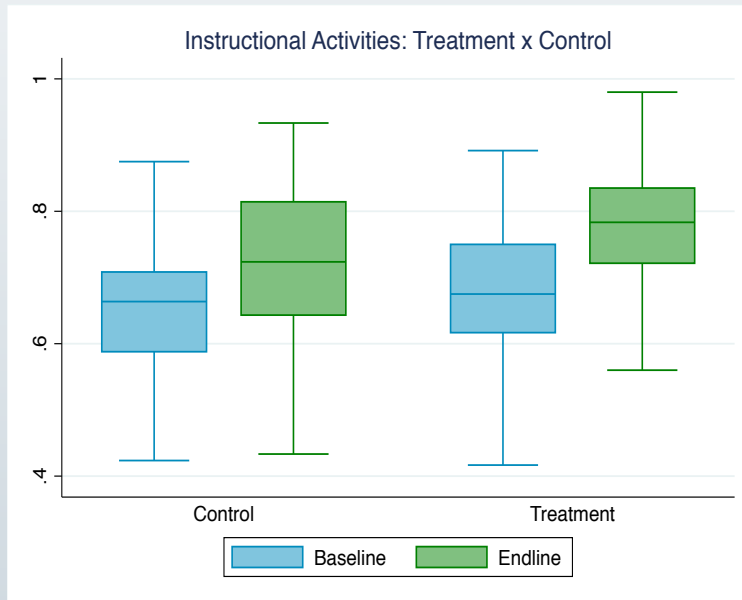
	No Weights		
	Control Means	Treatment Means	Difference
Instructional activities	0.656 [0.101]	0.674 [0.102]	-0.0184 [0.0119]
Classroom management activities	0.250 [0.0724]	0.228 [0.0812]	0.0220* [0.00906]
Off-task activities	0.0940 [0.0618]	0.0976 [0.0654]	-0.00361 [0.00748]
Student off-task	0.227 [0.146]	0.189 [0.136]	0.0383* [0.0165]
Instructional activities with all students engaged	0.194 [0.144]	0.236 [0.153]	-0.0424* [0.0174]
Reading aloud	0.0430 [0.0363]	0.0432 [0.0351]	-0.000226 [0.00418]
Demonstration/Lecture	0.326 [0.112]	0.334 [0.110]	-0.00807 [0.0130]
Discussion/Debate/Q&A	0.0972 [0.0590]	0.0990 [0.0726]	-0.00182 [0.00781]
Practice & Drill	0.00431 [0.00874]	0.00442 [0.0128]	-0.000119 [0.00131]
Assignment/Class work	0.122 [0.0801]	0.132 [0.0994]	-0.00984 [0.0107]
Copying	0.0629 [0.0431]	0.0613 [0.0484]	0.00167 [0.00540]
Verbal Instruction	0.0604 [0.0351]	0.0569 [0.0347]	0.00352 [0.00409]
Discipline	0.0205 [0.0190]	0.0167 [0.0166]	0.00387 [0.00209]
Classroom management	0.0807 [0.0421]	0.0767 [0.0450]	0.00395 [0.00512]
Classroom management alone	0.0886 [0.0573]	0.0779 [0.0525]	0.0107 [0.00643]
Social interaction	0.0156 [0.0229]	0.0175 [0.0283]	-0.00185 [0.00305]
Teacher out of the room	0.0572 [0.0397]	0.0581 [0.0478]	-0.000815 [0.00518]
Teacher uninvolved	0.0211 [0.0307]	0.0221 [0.0274]	-0.000941 [0.00340]
No material	0.128 [0.0777]	0.131 [0.0667]	-0.00240 [0.00845]
Textbook	0.101 [0.0820]	0.0938 [0.0811]	0.00731 [0.00956]
Notebook	0.119 [0.0738]	0.137 [0.117]	-0.0186 [0.0116]
Blackboard	0.271 [0.108]	0.270 [0.112]	0.000989 [0.0130]
Learning aides	0.0255 [0.0476]	0.0216 [0.0354]	0.00386 [0.00487]
TIC	0.0632 [0.0813]	0.0686 [0.0813]	-0.00543 [0.00954]
Cooperative	0.00795 [0.0188]	0.00859 [0.0234]	-0.000640 [0.00251]
Joint test (p-value)			0.81
Number of schools	136	156	

Note: Numbers in parentheses are standard deviations in the column of means and standard errors in columns of differences. * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

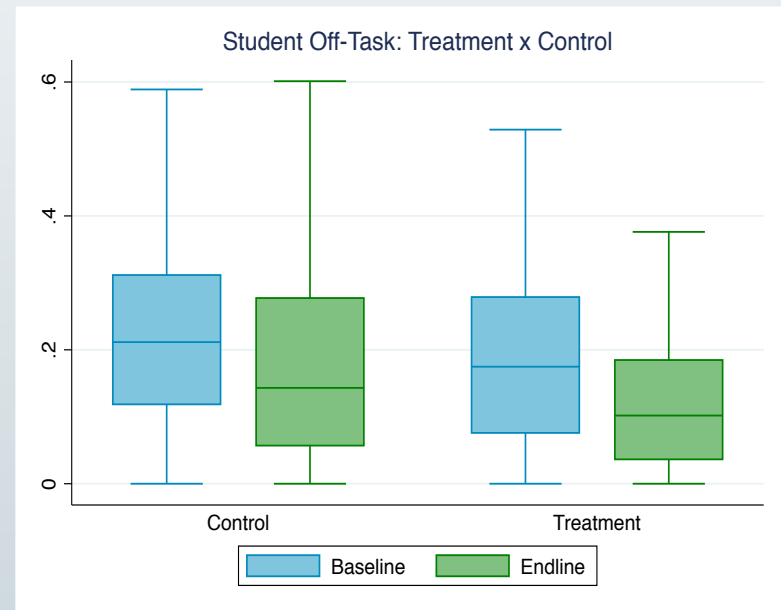
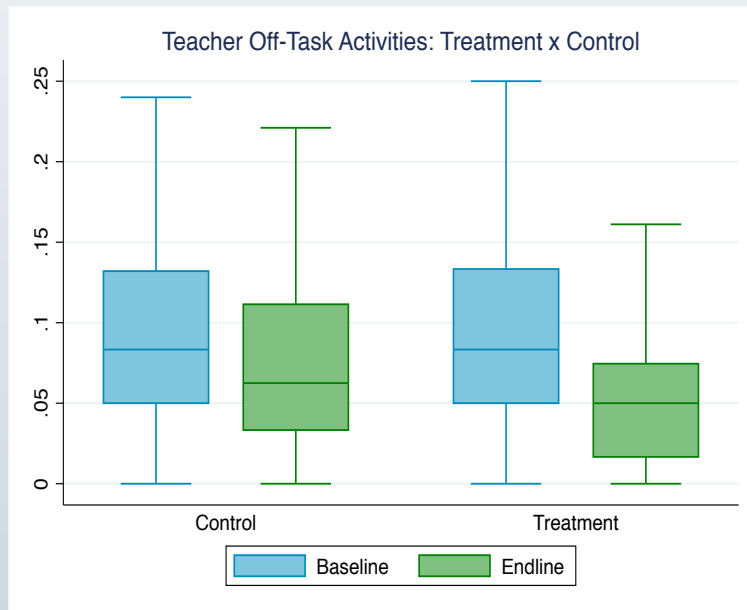
Descriptive statistics

	Baseline Means and Std			Endline Means and Std		
	All Sample	Control	Treatment	All Sample	Control	Treatment
Instructional activities	0.655 [0.212]	0.646 [0.211]	0.665 [0.212]	0.735 [0.199]	0.704 [0.209]	0.766 [0.183]
Classroom management activities	0.244 [0.176]	0.255 [0.176]	0.233 [0.176]	0.194 [0.157]	0.211 [0.166]	0.176 [0.145]
Off-task activities	0.101 0.0608	0.0992 0.0611	0.102 0.0605	0.0718 0.0402	0.0848 0.0498	0.0587 0.0306
o/w Teacher out of the room	[0.0996] [0.132]	[0.0998] [0.132]	[0.0995] [0.133]	[0.0766] [0.118]	[0.0872] [0.128]	[0.0629] [0.105]
Instructional activities with all students engaged	0.200 [0.263]	0.183 [0.251]	0.217 [0.273]	0.267 [0.302]	0.265 [0.302]	0.269 [0.303]
Student off-task	0.223 [0.284]	0.242 [0.296]	0.203 [0.271]	0.166 [0.265]	0.187 [0.280]	0.144 [0.246]
Sample Size	3121	1560	1561	3121	1560	1561

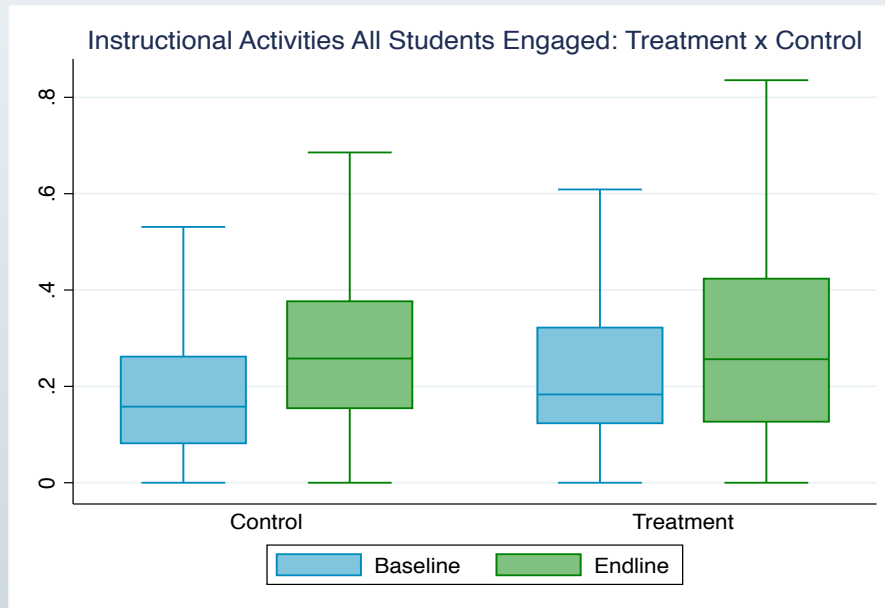
Figures



Figures



Figures



Impact estimate

- Intent-to-treat effects (ITT) - Differences between treatment and control group means for each treatment arm. In other words, ITT provides an estimate of the impact of being offered a chance to participate in a given arm of the experiment

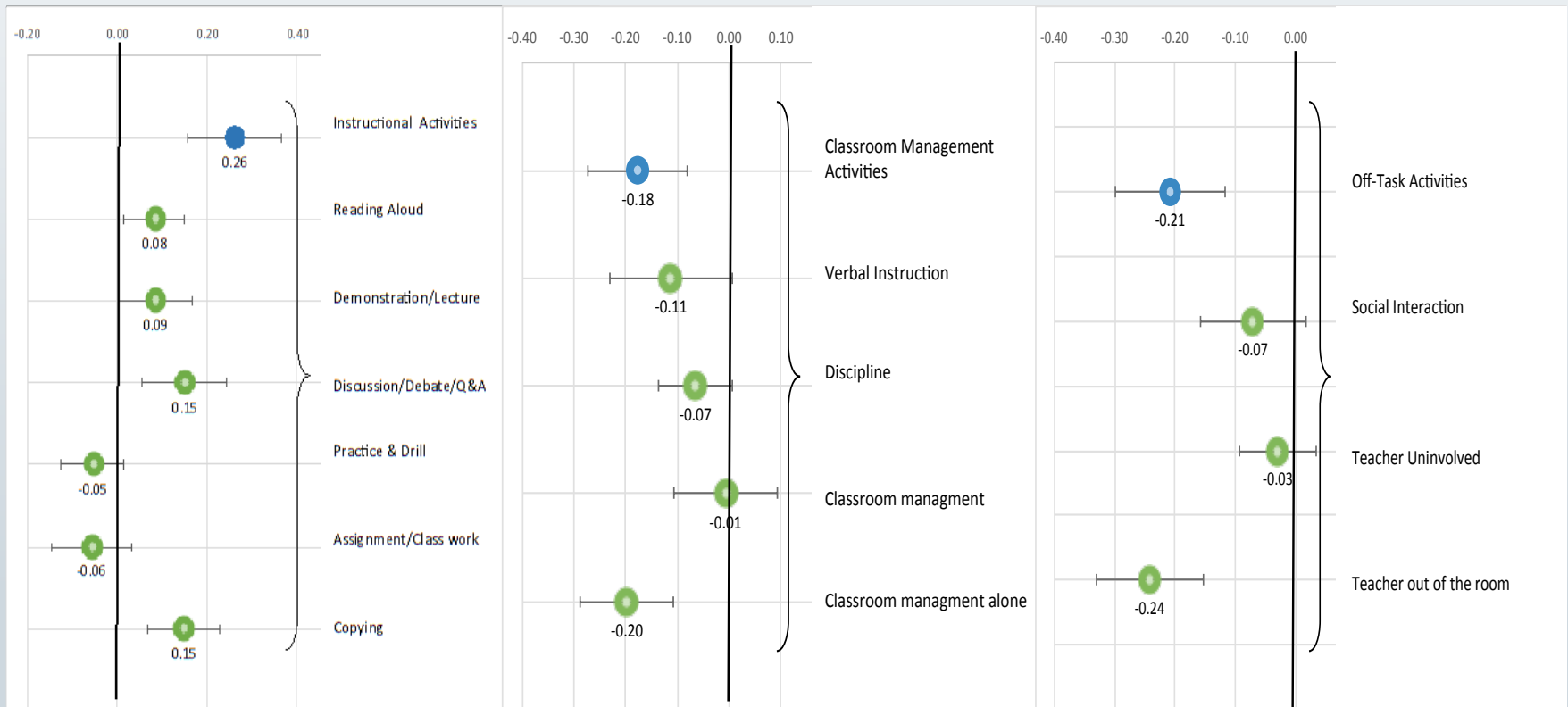
$$y_i = \beta_0 + \beta_1 y_{i,t-1} + \mathbf{X}'_i \beta_2 + \alpha_0 Z_i + \varepsilon_i \quad (1)$$

Results

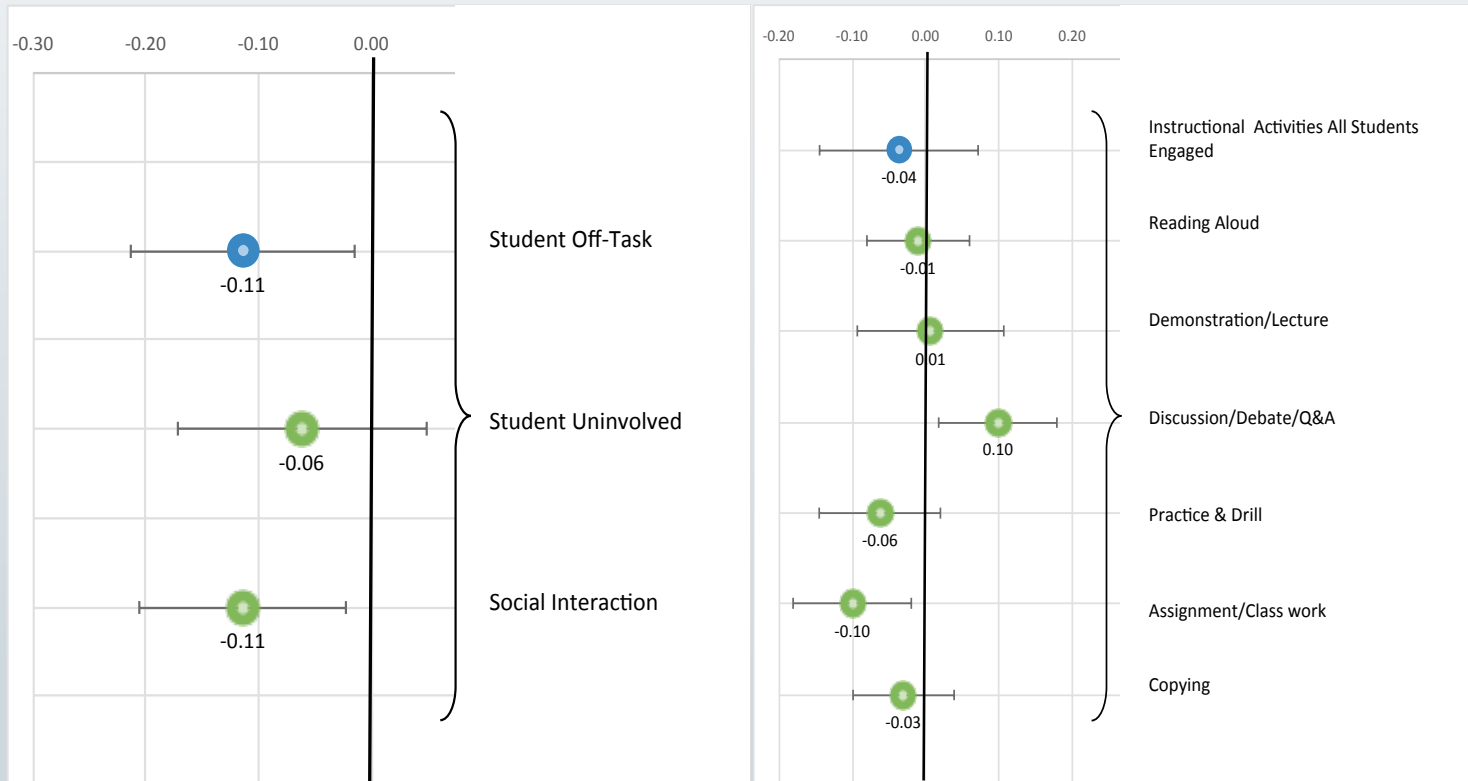
Dependent variable	OLS results (1)	OLS results with baseline (2)	OLS results with baseline, student, teacher and class covariates (3)	OLS results with baseline and all covariates (4)	Sample size (5)
A. Instructional activities	0.311*** (0.0656)	0.307*** (0.0651)	0.289*** (0.0653)	0.261*** (0.0635)	3121
B. Classroom management activities	-0.227*** (0.0575)	-0.226*** (0.0576)	-0.207*** (0.0601)	-0.177*** (0.0586)	3121
C. Off-task activities	-0.221*** (0.0606)	-0.223*** (0.0590)	-0.216*** (0.0567)	-0.208*** (0.0552)	3121
D. Instructional activities all students engaged	0.00675 (0.0667)	-0.0103 (0.0643)	-0.0180 (0.0674)	-0.0360 (0.0660)	3085
E. Big group (>6) of student off-task	-0.158** (0.0632)	-0.137** (0.0594)	-0.137** (0.0606)	-0.114* (0.0605)	3085

Note: Standardized dependent variables (z-scores). Robust standard errors in parentheses, clustered at the school level. Variables D and E only consider the time teacher was instructing. These variables assumes missing values if the teacher did not spend any time instructing. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Figures



Figures



Impact Estimates

- Intent to treat effect – intra-school variation: we calculate the standard deviation of each of the main summary variables at the school level and use it as a dependent variable

$$\mu_i = \beta_0 + \beta_1\mu_{i,t-1} + \mathbf{X}'_i\beta_2 + \alpha_0Z_i + \varepsilon_i \quad (2)$$

Results

Dependent variable	OLS results (1)	OLS results with baseline (2)	OLS results with baseline, student, teacher and class covariates (3)	OLS results with baseline and all covariates (4)	Sample size (5)
A. Instructional activities	-0.342*** (0.116)	-0.344*** (0.115)	-0.243** (0.118)	-0.222* (0.119)	292
B. Classroom management activities	-0.301** (0.116)	-0.299** (0.116)	-0.211* (0.121)	-0.196 (0.121)	292
C. Off-task activities	-0.342*** (0.116)	-0.342*** (0.112)	-0.326*** (0.120)	-0.294** (0.119)	292
D. Instructional activities all students engaged	-0.0923 (0.117)	-0.168 (0.115)	-0.0530 (0.119)	-0.0371 (0.119)	292
E. Big group (>6) of student off-task	-0.293** (0.116)	-0.191* (0.110)	-0.181 (0.114)	-0.158 (0.114)	292

Note: Standardized dependent variables (z-scores). Robust standard errors in brackets, clustered at the school level. * p<0.10 ** p<0.05 *** p<0.01

Partial Compliance

Certification by ELOS

Certifield	Control	Treatment	Attrition	Total
No	0	18	3	21
(%)	0	5.14	0.86	6
Yes	0	138	15	153
(%)	0	39.43	4.29	43.71
Not Evaluated	136	0	40	176
(%)	38.86	0	11.43	50.29
Total	136	156	58	350
(%)	38.86	44.57	16.57	100

Grade for Certification by ELOS

Grade	Not Certified	Certified	Total
Bad	18	0	18
(%)	12.18	0	12.18
Regular	0	30	30
(%)	0	18.59	18.59
Good	0	59	59
(%)	0	37.82	37.82
Excelent	0	49	49
(%)	0	31.41	31.41
Total	18	138	156
	12.18	87.82	100

IV estimate

2SLS estimation:

$$y_i = \beta_0 + \beta_1 y_{i,t-1} + \mathbf{X}'_i \beta_2 + \alpha_0 c_i + \tau_i \quad (5)$$

$$c_i = \mathbf{X}'_i \gamma_1 + \pi Z_i + \mu_i \quad (6)$$

Results

Dependent variable	2SLS results (1)	2SLS results with baseline (2)	2SLS results with baseline, student, teacher and class covariates (3)	2SLS results with baseline and all covariates (4)	Sample size (5)
A. Instructional activities	0.355*** (0.0746)	0.350*** (0.0740)	0.328*** (0.0736)	0.293*** (0.0705)	3121
B. Classroom management activities	-0.260*** (0.0654)	-0.258*** (0.0655)	-0.236*** (0.0677)	-0.200*** (0.0650)	3121
C. Off-task activities	-0.253*** (0.0690)	-0.248*** (0.0681)	-0.239*** (0.0640)	-0.229*** (0.0617)	3121
D. Instructional activities all students engaged	0.00770 (0.0760)	-0.0117 (0.0732)	-0.0204 (0.0761)	-0.0404 (0.0738)	3085
E. Big group (>6) of student off-task	-0.181** (0.0718)	-0.156** (0.0675)	-0.155** (0.0678)	-0.129* (0.0670)	3085

Note: Standardized dependent variables (z-scores). Robust standard errors in parentheses, clustered at the school level. Variables D and E only consider the time teacher was instructing. These variables assumes missing values if the teacher did not spend any time instructing. * p<0.10 ** p<0.05 *** p<0.01

More Experiments treats

- Attrition
- Spillover
- Treatment Contamination
- Evaluation-Driven Effects

Thank you!

barbara.brunsgmail.com

lcosta@worldbank.org

ninamcunhagmail.com