

# Scores, Camera, Action?

## Incentivizing Teachers in Remote Areas

**Arya Gaduh**  
*U. of Arkansas*

**Menno Pradhan**  
*U. of Amsterdam,  
VU Amsterdam*

**Jan Priebe**  
*GIGA Institute  
Hamburg,  
U. of Gottingen*

**Dewi Susanti**  
*World Bank, Jakarta*

**CIES conference**  
**28 February 2021**

# Education Delivery Challenges in Remote Areas

- Ensuring quality (public) service is a daunting task in **remote areas**
  - **Undesirable** locations for service providers (teachers, nurses, doctors)
  - National standards may be too high for **local conditions**
  - Difficult to **monitor and enforce** service standards
- Incentivize remote assignments via **hardship allowance...** but is ineffective
  - **Gambia**: 30-40% hardship premium → no effect on learning (Pugatch and Schroeder 2018)
  - **Indonesia**: Remote-area allowance recipients absent more than non-recipients in same location (31.5 v. 23.6) (SMERU, 2010)

## Policy experiment in education:

Bottom-up monitoring against local standards

Three variations of teacher incentives,  
including incentivized Government remote area allowance

# KIAT Guru Experiment - Indonesia

*Teacher Performance and Accountability*



## High teacher absenteeism in Indonesia's remote areas

19% in remote areas v. 9% nationally (ACDP 2014)

- **Collaborate** with education ministry, district governments
- **Remote** schools...
  - at least one-hour drive from disadvantaged-district capital
- Work with **government-paid remote-area-allowance (TSA)**
  - “**hardship** allowance = base salary” or *Teacher Special Allowance (TSA)*
  - received by selected **permanent, government-contracted** teachers...
  - ... and registered **private school** teachers
- Core components:
  - **social accountability only**
  - social accountability plus **two** types of **performance pay (PP)**

# 5 Districts, 270 primary schools

baseline: late 2016/early 2017 | **endline:** early 2018 | **follow-up:** early 2019



**2**  
Provinces



**270**  
Primary schools



**1778**  
Primary school teachers



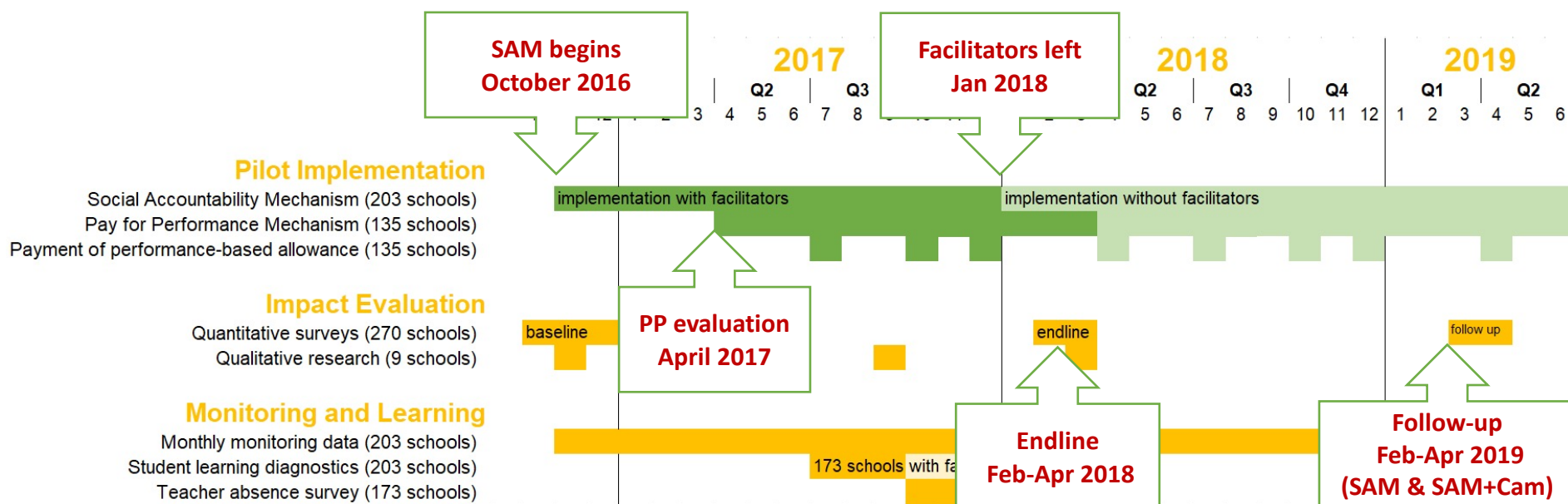
**1827**  
User committee members



**26062**  
Primary school students

# Implementation Timeline

One-Year (all) and Two-Year Impacts (SAM & SAM+Cam)





# Intervention Design



# Social Accountability

- **Local standard: Service Agreement**
  - Assign roles for teachers and parents/community
  - Service agreement can be revised at least every semester
- **Teacher component → Teacher-specific scorecard**



# Teacher-Specific Score Cards

No	Teacher service indicator	Max weight	Service description (Put mark on corresponding condition)	Score	Actual score	Total Indicator Score	The reason for the value of the score
1	Teacher arrives in class from Monday to Thursday at 07:30 - 10:00 hours to take picture with KIAT to return home from work	25	Teacher arrives on time for 24 days in a month	15	13	23	Teacher went to Sintony for three days to fake his salary
			Teacher arrives on time for 20 days in a month	5	5		
			Teacher arrives on time for 15 days in a month	5	5		
			Teacher arrives on time for 10 days in a month	0	0		
4	Teacher does not commit any violent action in school areas	5	a Teacher does not commit any violent action in school areas	5	5	5	
			b Teacher commits any violent action in school areas	0	0		
6	While teaching, teacher uses props (varied methods) 1 time minimum in 1 week (or 4 times in minimum in a month)	10	a Teacher uses props (varied methods) 1 time minimum in 1 week (or 4 times in minimum in a month)	2,5	2,5		According to agreement
			b Teacher uses props (varied methods) 2 times minimum in 1 week (or 8 times in minimum in a month)	2			
			c Teacher uses props (varied methods) minimum in the third week of the month	2			
			d Teacher uses props (varied methods) 1 time minimum in the third week of the month	2			
3	Every Saturday, students do morning exercise, read library book in class, learn Art and Cultural Skills, (hereafter SBK) accompanied by the teachers. In every 2 weeks, students and teachers will do community service by cleaning school areas.	15	a exercise by teachers	3			According to agreement
			b Student exercise by the teachers in every second Saturday of the month.	3	3		
Total Weight		100					

Acknowledged by,  
Teacher/ School Principal\*

Evaluated by,  
Representative of User Committee

Approved by,  
School Principal/Head of (sub-district) education department\*  
+Stamp



# Social Accountability

- **Local standard: Service Agreement**
  - Assign roles for teachers and parents/community
  - Service agreement can be revised months later
- **Teacher component → Teacher-specific scorecard**
  - List of 5 to 8 measurable **indicators** and **weights**
  - **Required** indicator: Teacher presence in school
  - **Rubric** to define scoring method of all indicators
  - For **all** teachers (not just allowance recipients)
- **User committee (UC)** to monitor
  - Formation facilitated through **consultative processes**
  - Minimum **9 members**
  - Three **community leaders** + **parents** from each grade
  - Meet **monthly** to discuss **scorecard evaluation**
  - **PP**: Signs off on scores that determine pay cuts



# Experimental Design

number of schools

	No Salary incentive	Remote area allowance cut based on	
		absenteeism as recorded by camera (CAM)	on teacher scorecard (Score)
Control	67	x	x
Social accountability mechanism (SAM)	68	68	67

# SAM+Cam

## Presence Indicator + Tamper-Proof Camera

---

- UC monitors **all indicators**, but **only presence** affected **performance pay**
- **Camera evidence** of presence:
  - Teachers take selfies at start and end of day
  - (Excused) absences **verified by UC** every month
- **Quantifying** (cuts in) presence:
  - partial presence = 1.5% cut
  - excused absence = 2% cut
  - unexcused absence = 5% cut
- Remote area allowance (RAO) payment:
  - **Total presence  $\geq 85\%$** : total percentage
  - **Total presence  $< 85\%$** : 0 ([adherence](#))
- **Non-RAO teachers** not financially affected

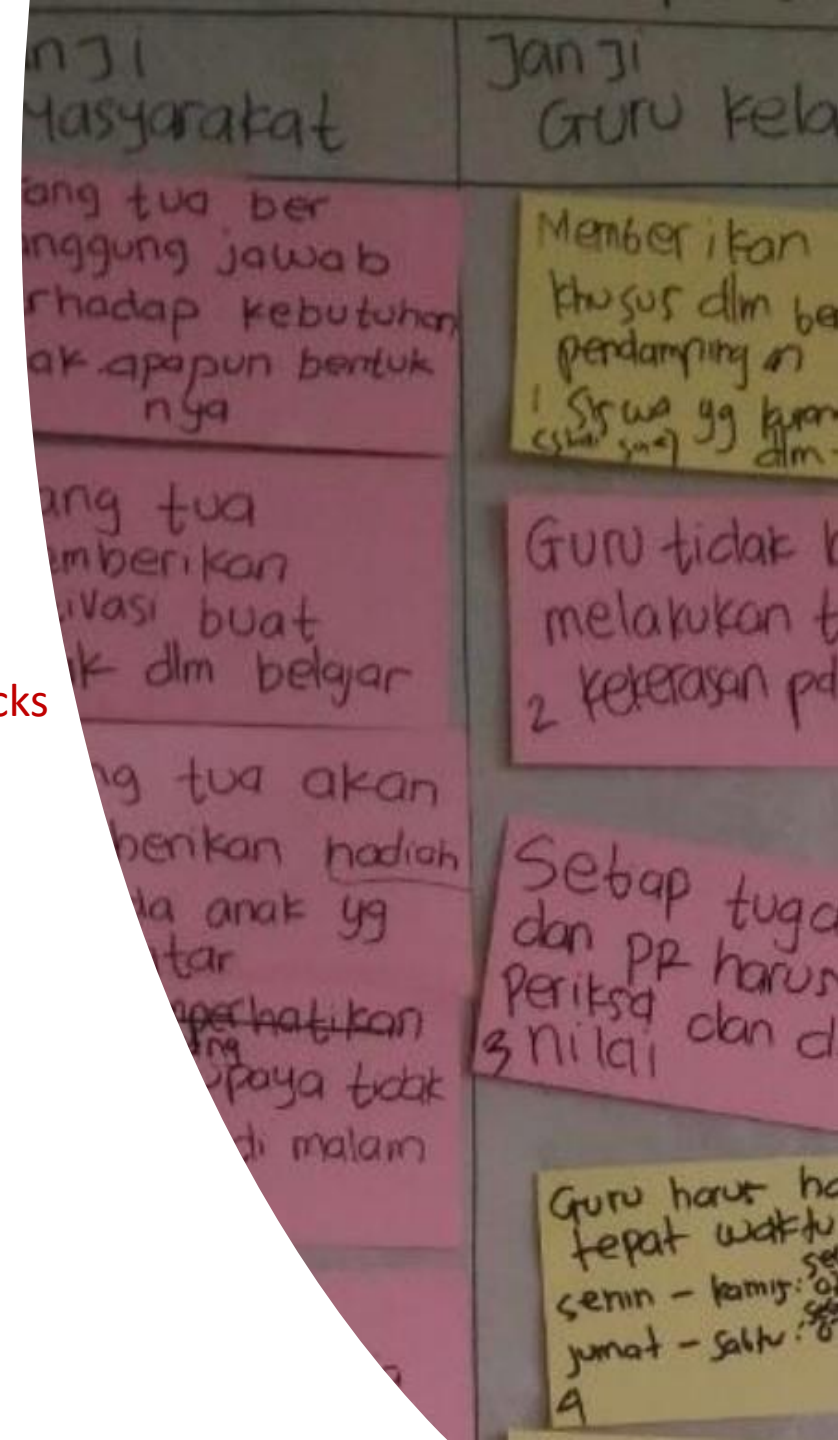


# SAM+Score

## Average Score on All Indicators

---

- UC monitors **all indicators**
  - Must include **presence**
  - Presence monitoring based on UC spot-checks
- **PP** based on **compound** indicator
  - Report score **out of 100**
- **Percentage** TSA paid = score obtained
  - Score 79 = 79% of TSA
- **Non-TSA** teachers not financially affected





# Results



**endline (one-year impact, 2018):** all treatments

**follow-up (a year after facilitators left, 2019):** **SAM** & **SAM+Cam** only



## Empirical Strategy

# Assignment and Estimation

- Stratified-random assignment into groups
- Estimate:

$$Y_{ijt}^k = \alpha_k + \delta Y_{ijt-1}^k + \sum_R \gamma^r T_j^{kr} + X_{ijt}^k \beta + \varepsilon_{ijt}^k$$

- Individual  $i$ , school  $j$ , time  $t$ , strata  $k$
- **Strata FE, cluster** at the **school** level
- For **student learning** outcomes (Indonesian + math), **controls** for:
  - Sex, age, parental education
  - Baseline outcome + mean school-level learning outcomes
- For **individual teacher** behavior, **controls** for:
  - Age, gender, marital status
  - Baseline outcome variables
- For **parental** behavior, **controls** for:
  - Children's sex, age, parental education
  - Baseline outcome variables
- Controls for private/public status

# Student Learning Outcomes

Mean of Grade-Adjusted Standardized Math and Indonesian Scores



**SAM+Cam** yielded the **strongest** and **persistent** impact on learning

[\[fig by subject\]](#) [\[table\]](#)

## Impact heterogeneity:

- Stronger in [lower grades](#)
- Gender **neutral**
- Don't depend on **years with TSA teachers**
- Stronger for **better students...**
- ...but **more persistent** in **weaker schools**

[\[table\]](#)

# Student Learning, Teacher Behavior, and Parent Engagement

- **Learning improvements across all treatments...**
  - One-year effect on Indonesian and mathematics scores
    - Indonesian: 0.08 – 0.15 s.d ↑
    - Mathematics: 0.07 – 0.18 s.d ↑
    - Overall: 0.08-0.17 s.d. ↑
  - Strongest and persistent in SAM+Cam
- **Weak, not persistent improvements in teacher behaviors**
  - Weakly improved attendance, work behavior in SAM+Cam
    - ...concentrated among TSA teachers
    - But negative effects on non-TSA teachers (especially in SAM+Score)
  - Year 2: No effect ([table](#))
- **Increased parental investments in education** (all ↑, but SAM+Cam ↑) ([table](#))
- **Improved school principal's evaluation practices** (all ↑) ([table](#))

**SAM+Cam produced most consistent, persistent improvements**

## Sustainability and Scalability

# Self-Reported Satisfaction

- **Sustainability** and **scale-up potential** affected by **support**
  - Performance pay can lead to teacher dissatisfaction
- **Parents** reported improved assessments of school, teachers
  - Comparable increase in satisfactions across treatments [\[table\]](#)
- **Teachers** reported:
  - feeling **more appreciated** in **all treatments** [\[table\]](#)
  - overall **more satisfied** of their job and salary [\[table\]](#)
    - **Non-TSA teachers more satisfied** (on job satisfaction non-persistent)
    - **TSA teachers more satisfied on salary** (persistent only for **SAM+Cam**)

**No evidence** of widespread teacher dissatisfaction

# Summary Results....and a Puzzle

- **SAM: modest, not persistent learning**
  - Engaged **parents**, improved **evaluation**
- **Learning: SAM+Cam > SAM+Score**
  - Engaged **parents**, improved **teacher**
  - No **negative spillovers** on non-TSA

	Intimidated (1)	Pressure to Increase Score (2)	Threats for Low Score (3)
SAM+Cam	0.021 (0.040)	-0.005 (0.056)	0.071 (0.044)
SAM+Score	0.066 (0.041)	0.119 (0.056)**	0.165 (0.044)***
Constant	0.107 (0.083)	0.035 (0.114)	-0.038 (0.090)
Control group mean	0.030	0.075	0.000
Observations	201	201	201
Strata FE	Yes	Yes	Yes

- **Puzzle: Why isn't SAM+Score performing better?**
  - Performance pay on subjective indicators → room for **negotiation**?
  - Qualitative study reported more **conflict** from SAM+Score
  - **Stronger teacher pressure** (on UC members) to increase score in **SAM+Score**

**PP: Narrow objective (presence) > Comprehensive subjective**





• Thank you!

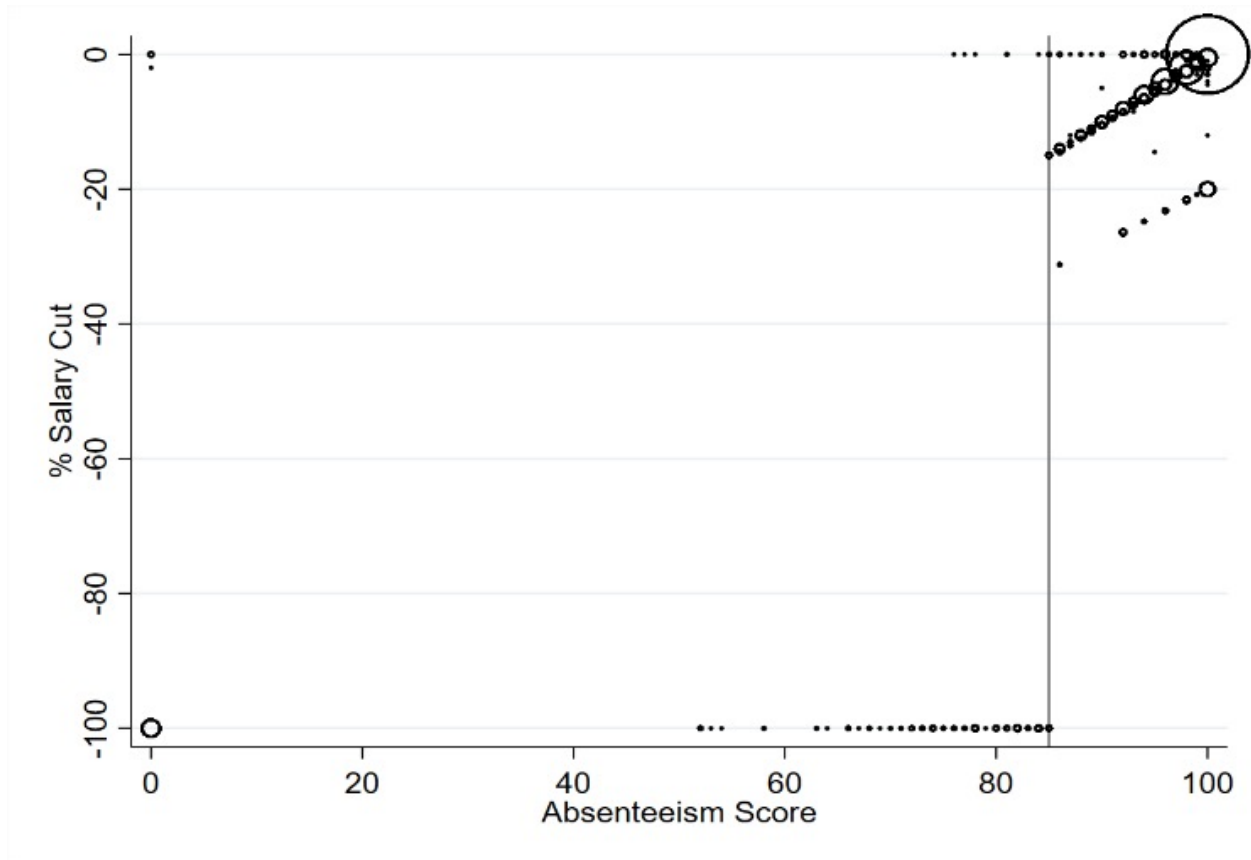


# Appendix

## Slides

# Payment rule in SAM+Cam generally followed

In 90% of cases the payment is the  $\pm 2$  p.p. band of the rule



Impact on (TSA) Teacher Behaviors

**Negative** for non-TSA teachers in **SAM+Score**

**Weak Positive** but **Not Persistent** for **SAM+Cam**

	2018			
	Presence		Working	
	(1)	(2)	(3)	(4)
SAM	0.007 (0.024)	-0.010 (0.038)	0.019 (0.029)	-0.009 (0.041)
SAM+Cam	0.023 (0.025)	-0.019 (0.046)	0.030 (0.030)	-0.058 (0.047)
SAM+Score	-0.063 (0.027)**	-0.133 (0.044)***	-0.076 (0.033)**	-0.162 (0.052)***
<u>Total impacts on TSA receivers</u>				
SAM		0.018 (0.035)		0.038 (0.039)
SAM+Cam		0.049 (0.033)		0.084** (0.038)
SAM+Score		-0.017 (0.036)		-0.020 (0.040)
Control group mean	0.84	0.84	0.80	0.80
Observations	1711	1711	1711	1711
Individual controls	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes

[back](#)

# Impact on Parent Investments

## Persistent Impacts on Parent Investments

### Stronger Impacts for SAM+Cam

	2018		
	Education Expenditure	Hours Accomp. Learning	Meetings w/ Teachers
	(1)	(2)	(3)
SAM	13746.9 (13405.4)	0.235 (0.194)	1.059 (0.213)***
SAM+Cam	27166.2 (14045.6)*	0.290 (0.194)	1.193 (0.219)***
SAM+Score	8312.7 (14199.6)	0.259 (0.196)	1.057 (0.243)***
Control group mean	324580.2	2.464	1.201
Observations	5377	5370	5377
Individual controls	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes



# Impact on Principal's Evaluation Practice

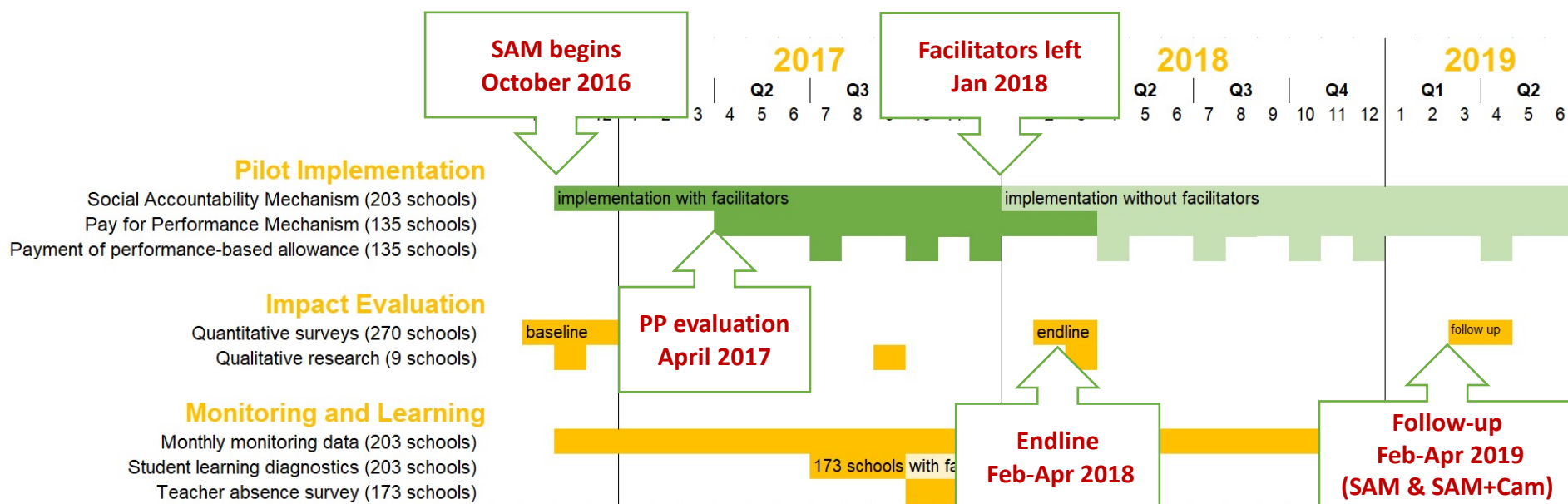
## Positive and Persistent Improvements on Principal's Evaluation Practices

	2018			2019		
	Routine Evaluation (1)	Evaluation Frequency (2)	Observed while Teaching (3)	Routine Evaluation (4)	Evaluation Frequency (5)	Observed while Teaching (6)
SAM	0.121 (0.050)**	1.924 (0.437)***	0.086 (0.029)***	0.093 (0.055)*	1.480 (0.517)***	0.065 (0.029)**
SAM+Cam	0.149 (0.049)***	2.506 (0.439)***	0.091 (0.028)***	0.125 (0.052)**	1.883 (0.517)***	0.069 (0.028)**
SAM+Score	0.146 (0.049)***	2.306 (0.425)***	0.099 (0.029)***			
Control group mean	0.42	2.79	0.67	0.45	3.44	0.70
Observations	270	270	2021	203	203	1430
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes

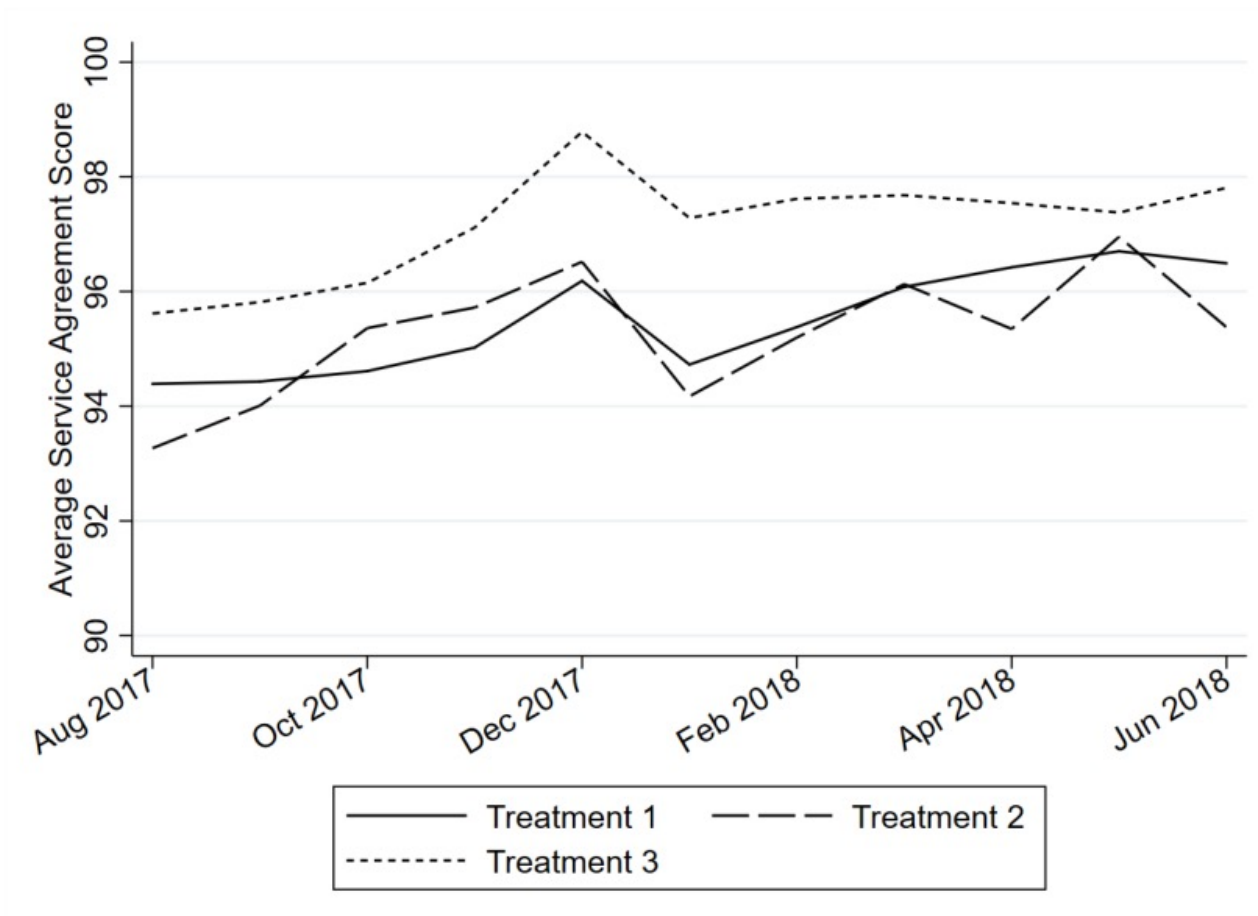
[back](#)

# Implementation Timeline

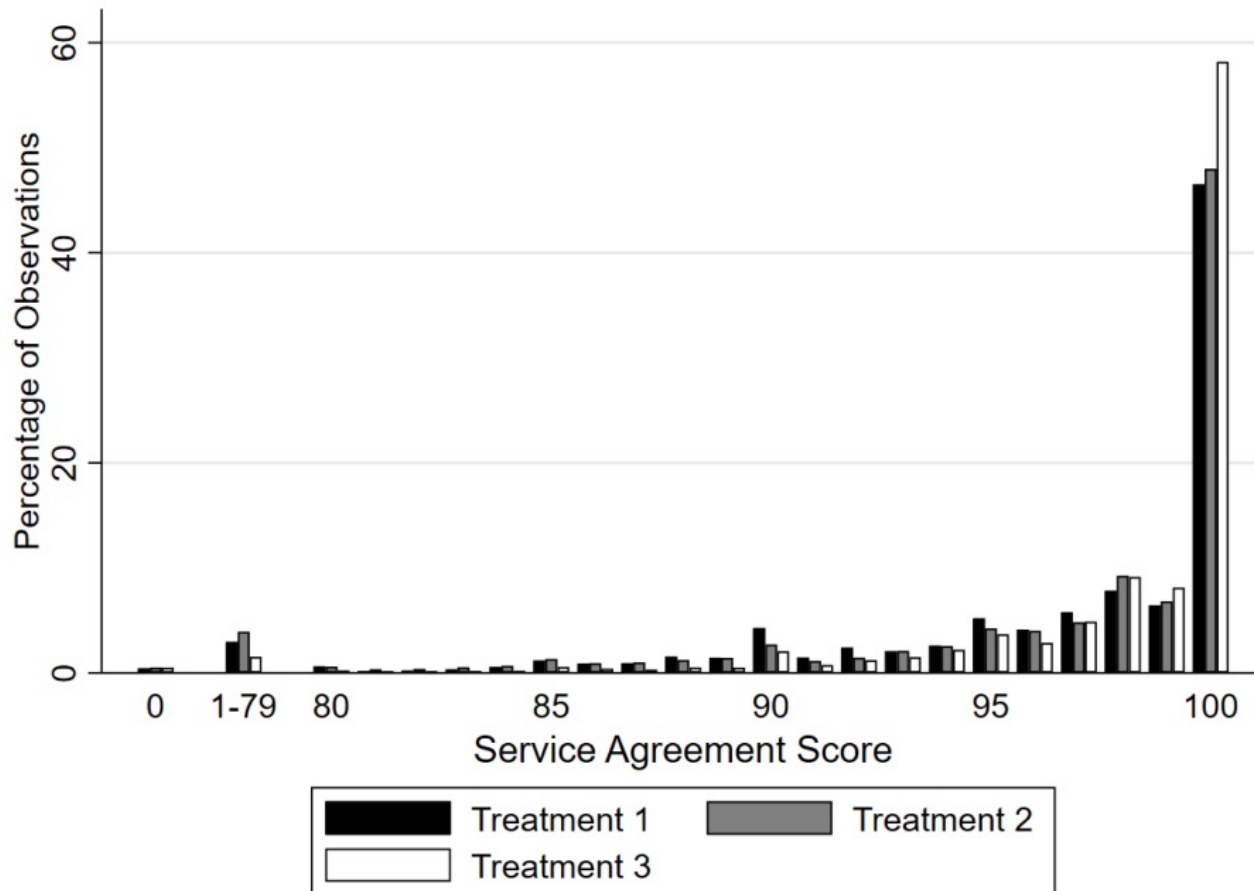
One-Year (all) and Two-Year Impacts (SAM & SAM+Cam)



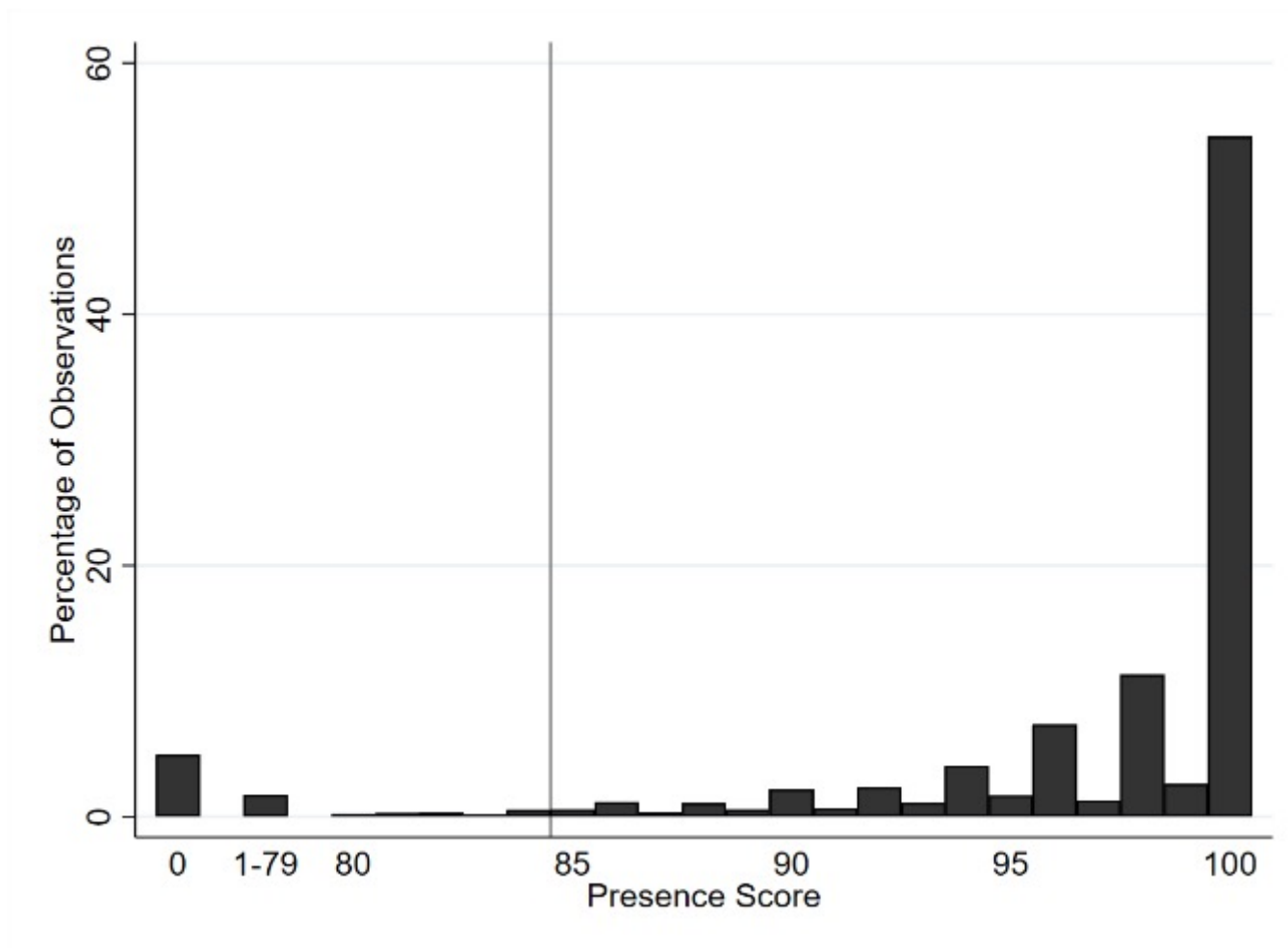
# Trend in Teacher Scorecard's Mean Scores



# Distribution of Mean Scores By Treatment



# Distribution of Presence Scores in SAM+Cam





# Balance Tables

## Student Characteristics

	Mean ( $\mu$ ) (standard errors)				Differences = $\mu_{[\dots]} - \mu_{Control}$ (p-value)			Differences between $\mu_{[\dots]}$ and $\mu_{[\dots]}$ (p-value)		
	Control	Treatment 1	Treatment 2	Treatment 3	Treatment 1	Treatment 2	Treatment 3	Tr. 2 - Tr. 1	Tr. 3 - Tr. 1	Tr. 3 - Tr. 2
Male	0.51 (0.50)	0.54 (0.50)	0.52 (0.50)	0.54 (0.50)	0.02** (0.01)	0.01 (0.38)	0.02** (0.01)	-0.02* (0.08)	-0.00 (0.85)	0.02* (0.08)
Age	10.76 (2.03)	10.63 (2.05)	10.69 (1.99)	10.65 (1.98)	-0.13 (0.12)	-0.07 (0.38)	-0.11 (0.15)	0.06 (0.47)	0.02 (0.82)	-0.04 (0.58)
Share having mothers with:										
...no education	0.09 (0.29)	0.07 (0.25)	0.11 (0.32)	0.09 (0.29)	-0.02 (0.19)	0.02 (0.51)	-0.00 (0.93)	0.05 (0.13)	0.02 (0.28)	-0.02 (0.49)
...primary education	0.75 (0.43)	0.74 (0.44)	0.71 (0.45)	0.73 (0.44)	-0.01 (0.85)	-0.04 (0.26)	-0.02 (0.45)	-0.03 (0.37)	-0.02 (0.60)	0.02 (0.65)
...more than primary education	0.16 (0.36)	0.19 (0.39)	0.18 (0.38)	0.18 (0.39)	0.03 (0.22)	0.02 (0.46)	0.02 (0.28)	-0.01 (0.64)	-0.01 (0.78)	0.01 (0.82)
Share having fathers with:										
...no education	0.08 (0.26)	0.05 (0.22)	0.09 (0.29)	0.08 (0.27)	-0.03* (0.09)	0.02 (0.48)	0.00 (0.96)	0.04* (0.08)	0.03 (0.13)	-0.02 (0.53)
...primary education	0.71 (0.45)	0.70 (0.46)	0.67 (0.47)	0.69 (0.46)	-0.02 (0.59)	-0.05 (0.13)	-0.03 (0.30)	-0.03 (0.34)	-0.01 (0.67)	0.02 (0.55)
...more than primary education	0.21 (0.41)	0.25 (0.43)	0.24 (0.43)	0.24 (0.43)	0.04 (0.13)	0.03 (0.26)	0.03 (0.24)	-0.01 (0.72)	-0.01 (0.59)	-0.00 (0.91)
Baseline learning assessment scores:										
Indonesian	37.83 (21.26)	36.94 (20.24)	38.46 (20.74)	36.56 (20.66)	-0.89 (0.65)	0.63 (0.74)	-1.27 (0.54)	1.52 (0.40)	-0.38 (0.85)	-1.91 (0.33)
Mathematics	38.63 (22.45)	37.14 (21.32)	37.93 (21.16)	36.82 (21.50)	-1.48 (0.49)	-0.69 (0.72)	-1.81 (0.43)	0.79 (0.70)	-0.33 (0.89)	-1.12 (0.61)

Notes: Standard errors clustered at the school level. \*/\*\*/\*\* denotes 10/5/1 percent significance levels

# Balance Tables

## Teacher Characteristics

	Mean ( $\mu$ ) (standard errors)				Differences = $\mu_{[\dots]} - \mu_{Control}$ (p-value)			Differences between $\mu_{[\dots]}$ and $\mu_{[\dots]}$ (p-value)		
	Control	Treatment 1	Treatment 2	Treatment 3	Treatment 1	Treatment 2	Treatment 3	Tr. 2 - Tr. 1	Tr. 3 - Tr. 1	Tr. 3 - Tr. 2
Male	0.53 (0.50)	0.52 (0.50)	0.49 (0.50)	0.50 (0.50)	-0.01 (0.63)	-0.04 (0.12)	-0.03 (0.23)	-0.03 (0.25)	-0.02 (0.46)	0.01 (0.68)
Age	37.39 (10.69)	37.65 (10.33)	37.55 (10.31)	37.37 (10.35)	0.26 (0.68)	0.16 (0.81)	-0.02 (0.97)	-0.10 (0.88)	-0.28 (0.66)	-0.18 (0.78)
Married	0.84 (0.36)	0.84 (0.37)	0.83 (0.38)	0.84 (0.37)	-0.00 (0.86)	-0.02 (0.42)	-0.01 (0.77)	-0.01 (0.53)	-0.00 (0.91)	0.01 (0.59)
Share with [...] education:										
... less than high school	0.01 (0.11)	0.01 (0.07)	0.00 (0.06)	0.01 (0.07)	-0.01 (0.39)	-0.01 (0.27)	-0.01 (0.37)	-0.00 (0.65)	-0.00 (0.96)	0.00 (0.68)
... high school	0.30 (0.46)	0.29 (0.46)	0.26 (0.44)	0.27 (0.44)	-0.01 (0.84)	-0.04 (0.25)	-0.03 (0.38)	-0.03 (0.31)	-0.02 (0.47)	0.01 (0.76)
... more than high school	0.69 (0.46)	0.70 (0.46)	0.74 (0.44)	0.73 (0.45)	0.01 (0.69)	0.05 (0.16)	0.04 (0.28)	0.03 (0.29)	0.02 (0.48)	-0.01 (0.73)
Share with [...] status:										
... civil servant	0.49 (0.50)	0.49 (0.50)	0.49 (0.50)	0.51 (0.50)	-0.01 (0.78)	0.00 (0.97)	0.02 (0.55)	0.01 (0.75)	0.03 (0.40)	0.02 (0.57)
... certified	0.20 (0.40)	0.19 (0.39)	0.21 (0.41)	0.23 (0.42)	-0.01 (0.64)	0.01 (0.74)	0.03 (0.30)	0.02 (0.44)	0.04 (0.15)	0.02 (0.48)
... TSA-receiving	0.16 (0.37)	0.19 (0.39)	0.19 (0.39)	0.18 (0.38)	0.03 (0.35)	0.03 (0.36)	0.02 (0.56)	0.00 (0.98)	-0.01 (0.73)	-0.01 (0.73)
Share of teachers observed to be:										
... present in school	0.79 (0.41)	0.78 (0.41)	0.81 (0.39)	0.84 (0.37)	-0.01 (0.74)	0.02 (0.61)	0.04 (0.19)	0.03 (0.39)	0.05* (0.09)	0.03 (0.41)
... working when in school	0.74 (0.44)	0.73 (0.44)	0.75 (0.43)	0.74 (0.44)	-0.01 (0.81)	0.02 (0.63)	0.00 (0.97)	0.02 (0.46)	0.01 (0.80)	-0.02 (0.69)
... teaching when in class	0.61 (0.49)	0.62 (0.49)	0.62 (0.49)	0.61 (0.49)	0.01 (0.81)	0.01 (0.69)	0.00 (0.99)	0.01 (0.87)	-0.01 (0.82)	-0.01 (0.70)
(Self-reported) hours spent monthly:										
... preparing lessons	17.83 (18.32)	16.42 (16.21)	17.47 (16.09)	18.26 (15.94)	-1.40 (0.38)	-0.36 (0.82)	0.43 (0.78)	1.05 (0.46)	1.84 (0.18)	0.79 (0.55)
... teaching curricular materials	62.54 (22.64)	65.27 (23.11)	67.10 (21.54)	64.41 (20.74)	2.74 (0.21)	4.57** (0.04)	1.87 (0.35)	1.83 (0.37)	-0.86 (0.64)	-2.69 (0.15)
... assessing student work	13.90 (13.25)	12.26 (10.08)	11.88 (10.90)	13.35 (11.91)	-1.64* (0.09)	-2.02** (0.04)	-0.55 (0.61)	-0.38 (0.62)	1.08 (0.23)	1.47 (0.10)
... teaching extra-curricular materials	4.73 (7.08)	3.78 (5.63)	4.13 (5.42)	4.25 (5.88)	-0.96* (0.09)	-0.60 (0.27)	-0.48 (0.40)	0.36 (0.47)	0.48 (0.37)	0.12 (0.82)
... on off-own-school employment	19.66 (35.84)	15.78 (30.78)	17.30 (35.67)	15.81 (27.70)	-3.87 (0.20)	-2.36 (0.48)	-3.84 (0.17)	1.51 (0.61)	0.03 (0.99)	-1.48 (0.59)

Notes: Standard errors clustered at the school level. \*/\*\*/\*\* denotes 10/5/1 percent significance levels

# Balance Tables

## Parent Characteristics

	Mean ( $\mu$ ) (standard errors)				Differences = $\mu_{[...]}$ - $\mu_{Control}$ (p-value)			Differences between $\mu_{[...]}$ and $\mu_{[...]}$ (p-value)		
	Control	Treatment 1	Treatment 2	Treatment 3	Treatment 1	Treatment 2	Treatment 3	Tr. 2 - Tr. 1	Tr. 3 - Tr. 1	Tr. 3 - Tr. 2
Mother is the respondent	0.46 (0.50)	0.47 (0.50)	0.46 (0.50)	0.51 (0.50)	0.01 (0.70)	0.01 (0.85)	0.06** (0.04)	-0.01 (0.85)	0.05* (0.09)	0.05* (0.06)
Respondent's age	39.68 (8.98)	39.14 (8.48)	39.37 (8.62)	39.12 (8.75)	-0.54 (0.25)	-0.30 (0.50)	-0.55 (0.22)	0.23 (0.59)	-0.02 (0.97)	-0.25 (0.55)
Education expenditure (Rp.)	301,890 (250,895)	311,114 (252,715)	298,330 (239,781)	326,076 (264,421)	9,224 (0.60)	-3,561 (0.84)	24,186 (0.17)	-12,785 (0.45)	14,962 (0.40)	27,746 (0.11)
Accompanied learning hours in the previous week	2.46 (2.95)	2.83 (3.26)	2.49 (2.75)	2.76 (3.15)	0.37** (0.02)	0.03 (0.84)	0.31** (0.05)	-0.34** (0.04)	-0.06 (0.71)	0.28 (0.10)
Paid tutor	0.00 (0.05)	0.00 (0.06)	0.01 (0.08)	0.00 (0.05)	0.00 (0.60)	0.00 (0.19)	0.00 (1.00)	0.00 (0.46)	-0.00 (0.60)	-0.00 (0.19)
Number of meetings with teacher on:										
... learning	1.88 (12.49)	1.88 (4.51)	1.75 (3.89)	1.82 (5.27)	0.00 (1.00)	-0.13 (0.78)	-0.06 (0.91)	-0.13 (0.62)	-0.06 (0.84)	0.07 (0.81)
... other issues	0.87 (3.10)	1.10 (2.98)	1.04 (2.74)	1.19 (3.16)	0.23 (0.28)	0.17 (0.44)	0.32 (0.18)	-0.06 (0.79)	0.09 (0.72)	0.15 (0.56)
Share of parents who believe:										
School quality is good or very good	0.89 (0.32)	0.88 (0.33)	0.92 (0.27)	0.91 (0.29)	-0.01 (0.69)	0.04 (0.13)	0.02 (0.37)	0.05** (0.02)	0.03 (0.12)	-0.01 (0.41)
Teacher absence is a main problem	0.23 (0.42)	0.22 (0.41)	0.24 (0.42)	0.20 (0.40)	-0.01 (0.66)	0.01 (0.89)	-0.03 (0.40)	0.02 (0.51)	-0.01 (0.60)	-0.03 (0.27)

Notes: Standard errors clustered at the school level. \*/\*\*/\*\* denotes 10/5/1 percent significance levels

# Selective Attrition and Entry Students

	Attrition (1)	Entry (2)
Treatment 1	-0.023 (0.024)	0.013 (0.054)
... × Above-median student	0.002 (0.008)	
... × Male	-0.003 (0.008)	0.001 (0.012)
... × Age	0.001 (0.002)	-0.000 (0.005)
... × Mother has post-primary education	0.002 (0.010)	-0.003 (0.024)
... × Father has post-primary education	0.013 (0.010)	-0.021 (0.022)
Treatment 2	-0.043 (0.026)*	0.050 (0.053)
... × Above-median student	0.006 (0.008)	
... × Male	0.000 (0.009)	-0.022 (0.013)*
... × Age	0.002 (0.002)	-0.005 (0.005)
... × Mother has post-primary education	-0.003 (0.010)	0.015 (0.023)
... × Father has post-primary education	0.015 (0.009)*	-0.006 (0.020)

Treatment 3	-0.024 (0.023)	0.088 (0.048)*
... × Above-median student	0.005 (0.007)	
... × Male	-0.006 (0.008)	-0.017 (0.013)
... × Age	0.001 (0.002)	-0.008 (0.004)*
... × Mother has post-primary education	0.001 (0.009)	-0.006 (0.025)
... × Father has post-primary education	0.006 (0.008)	0.016 (0.021)
Control group mean	0.08	0.20
R2	0.050	0.379
Observations	25483	30576
Strata FE	Yes	Yes
Individual Controls	Yes	Yes

Notes: Individual control variables are sex, age, both parents education, and dummy variables for individuals with missing controls. Above-median students are those whose average standardized scores of both subjects are above their class median. Standard errors are clustered at the school level. \*/\*\*/\*\* denotes 10/5/1 percent significance levels

...T3 continues ►

# Selective Attrition and Entry Teachers

	Attrition (1)	Entry (2)
Treatment 1	0.112 (0.358)	-0.296 (0.401)
... × Male	0.057 (0.040)	-0.063 (0.045)
... × Age	-0.013 (0.020)	0.014 (0.021)
... × Age <sup>2</sup>	0.000 (0.000)	-0.000 (0.000)
... × Married	0.167 (0.076)**	0.034 (0.074)
... × Civil servant	0.007 (0.052)	-0.020 (0.064)
... × Certified	-0.003 (0.061)	-0.120 (0.067)*
Treatment 2	0.400 (0.359)	-0.118 (0.414)
... × Male	0.015 (0.038)	-0.102 (0.043)**
... × Age	-0.030 (0.019)	0.007 (0.022)
... × Age <sup>2</sup>	0.000 (0.000)	-0.000 (0.000)
... × Married	0.255 (0.075)***	-0.013 (0.080)
... × Civil servant	-0.058 (0.050)	0.056 (0.066)
... × Certified	-0.003 (0.067)	-0.142 (0.076)*

Treatment 3	-0.130 (0.415)	0.419 (0.425)
... × Male	0.012 (0.041)	-0.085 (0.048)*
... × Age	-0.001 (0.023)	-0.021 (0.022)
... × Age <sup>2</sup>	-0.000 (0.000)	0.000 (0.000)
... × Married	0.232 (0.073)***	0.093 (0.076)
... × Civil servant	-0.001 (0.051)	0.006 (0.069)
... × Certified	0.024 (0.063)	-0.089 (0.073)
Control group mean	0.13	0.16
R2	0.209	0.326
Observations	2292	2331
Strata FE	Yes	Yes
Individual Controls	Yes	Yes

Notes: Individual control variables are sex, age, both parents education, and dummy variables for individuals with missing controls. Above-median students are those whose average standardized scores of both subjects are above their class median. Standard errors are clustered at the school level. \*/\*\*/\*\* denotes 10/5/1 percent significance levels

...T3 continues ►



## Impact on Student Learning Outcome

# SAM+Cam: strong and persistent impacts on learning

	Indonesian		Mathematics		Average Score		Grade Repetition	
	2018 (1)	2019 (2)	2018 (3)	2019 (4)	2018 (5)	2019 (6)	2018 (7)	2019 (8)
SAM	0.094 (0.036)***	0.009 (0.027)	0.073 (0.040)*	0.038 (0.044)	0.084 (0.035)**	0.024 (0.032)	0.010 (0.010)	-0.000 (0.008)
SAM+Cam	0.186 (0.035)***	0.089 (0.028)***	0.203 (0.041)***	0.171 (0.046)***	0.195 (0.035)***	0.128 (0.034)***	0.004 (0.010)	0.014 (0.008)
SAM+Score	0.118 (0.033)***		0.095 (0.038)**		0.109 (0.033)***		0.009 (0.010)	
Control group mean							0.08	0.04
Control group raw-score mean	47.13	38.64	47.03	44.34	47.08	41.49		
Test of equality (P-val)								
SAM v. SAM+Cam	0.015	0.005	0.003	0.003	0.003	0.002	0.565	0.117
SAM+Cam v. SAM+Score	0.058		0.013		0.017		0.614	
SAM v. SAM+Score	0.499		0.602		0.501		0.963	
Observations	31022	15942	31022	15942	31022	15942	24719	13257
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

*Notes:* Standardized scores are grade adjusted. Control variables include sex, age dummies, both parents' education, baseline outcome, dummy variables for missing controls (one for each control variable), school-level mean scores, and dummy variables for whether the school is a private school and whether it was among the three control schools who became TSA-ineligible due to the change in the government's definition of remoteness. Standard errors are clustered at the school level.

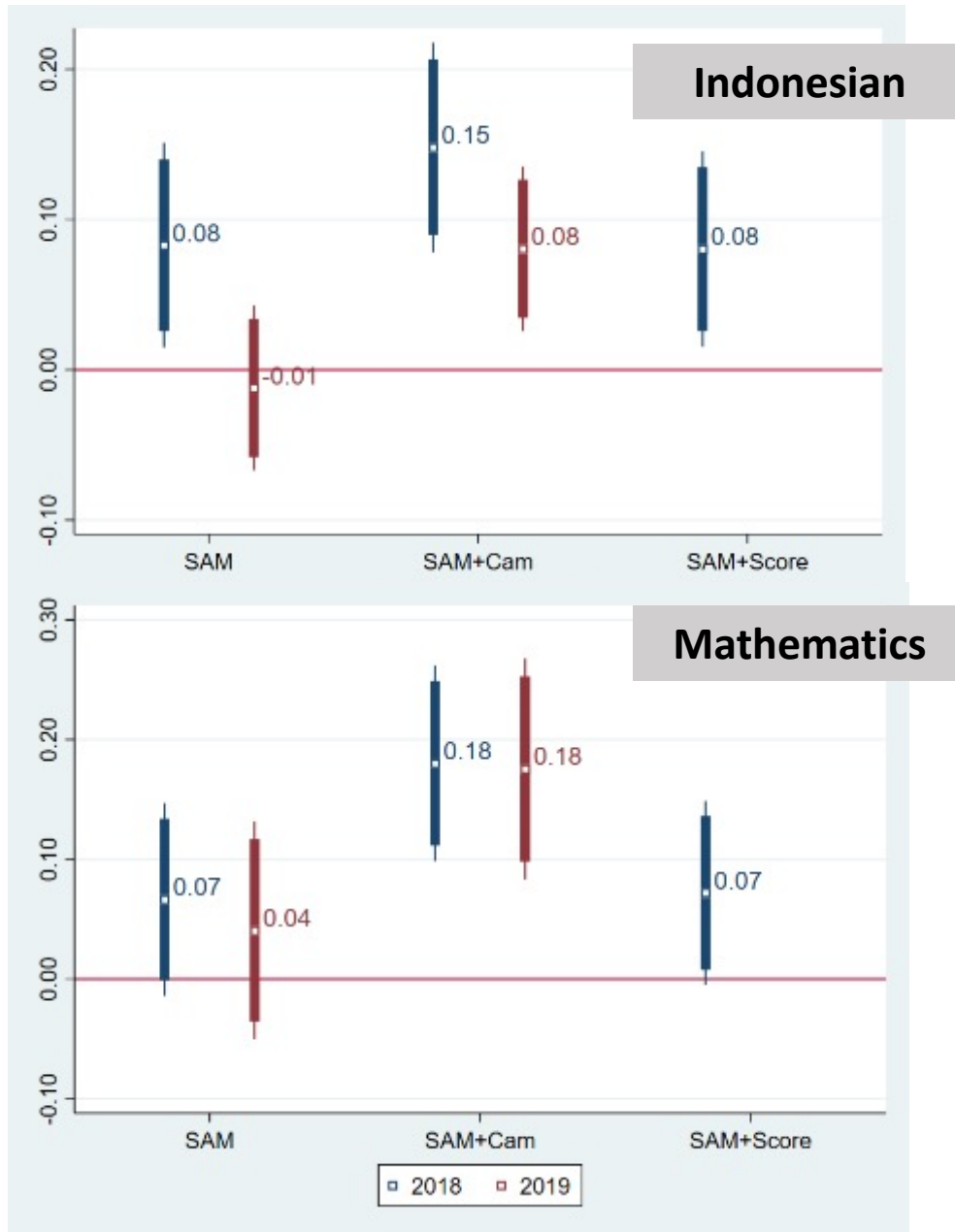
\*/\*\*/\*\*\* denotes 10/5/1 percent significance levels



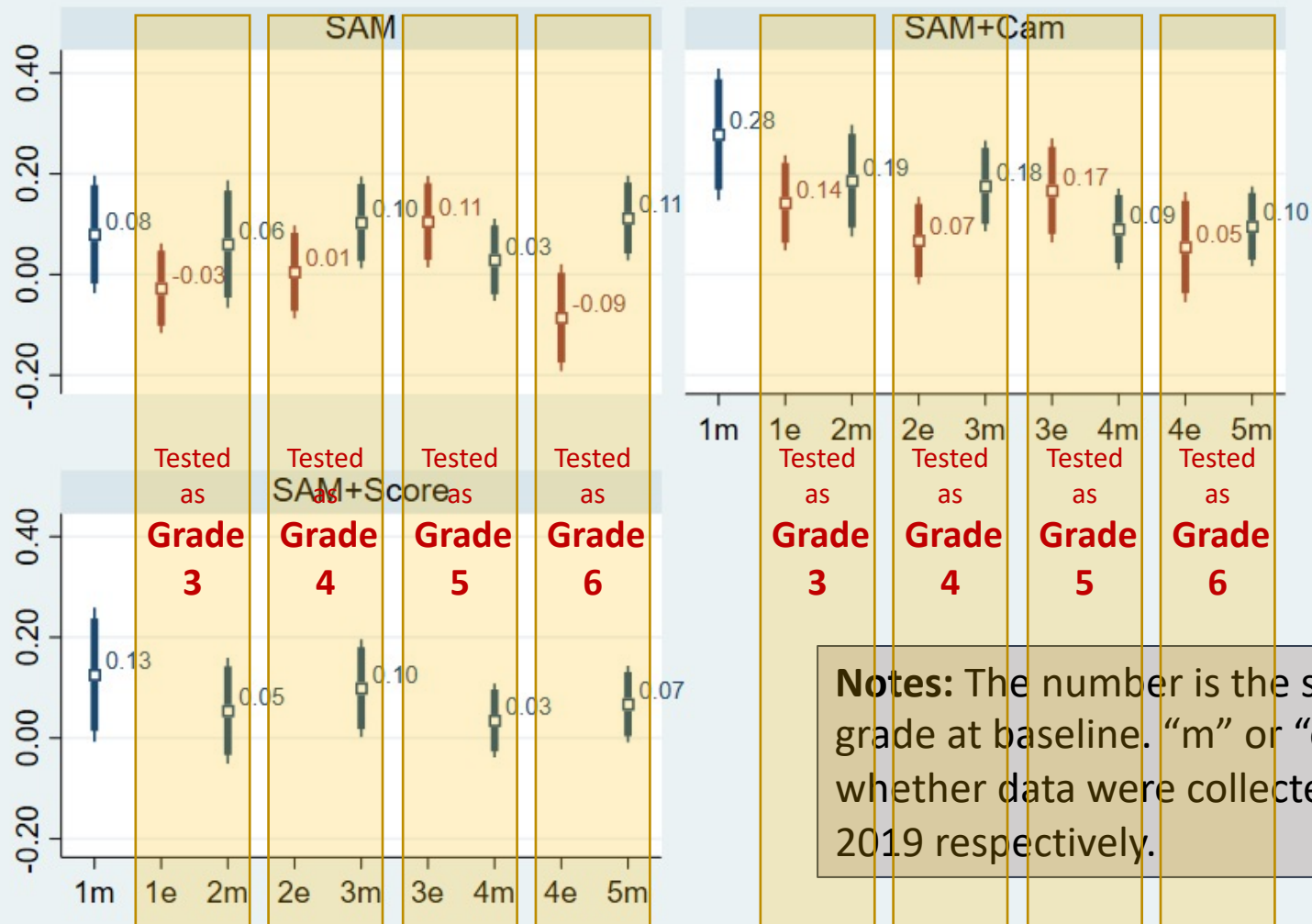
# Student Learning Outcomes by Subject

Impacts on mathematics are more persistent

[\[student learning\]](#)



# Student Learning Outcomes: Mean Scores By Grade



**Notes:** The number is the student's grade at baseline. "m" or "e" indicates whether data were collected in 2018 or 2019 respectively.

# Heterogenous Impact on Student Learning Outcome

## Gender, Student Ability, School Quality

	Male		Years with TSA teachers		Above-median student				Above-median school	
					in school		across all schools			
	2018 (1)	2019 (2)	2018 (3)	2019 (4)	2018 (5)	2019 (6)	2018 (7)	2019 (8)	2018 (9)	2019 (10)
SAM	0.055 (0.039)	0.019 (0.039)	0.051 (0.062)	-0.015 (0.070)	0.070 (0.040)*	-0.001 (0.038)	0.060 (0.040)	-0.006 (0.039)	0.109 (0.051)**	0.076 (0.047)
SAM+Cam	0.148 (0.039)***	0.123 (0.040)***	0.135 (0.064)**	0.109 (0.069)	0.130 (0.041)***	0.104 (0.040)***	0.130 (0.044)***	0.107 (0.044)**	0.161 (0.057)***	0.193 (0.057)***
SAM+Score	0.063 (0.038)*		0.020 (0.058)		0.066 (0.038)*		0.079 (0.042)*		0.110 (0.051)**	
Covariate: [...]	-0.149 (0.023)***	-0.216 (0.030)***	-0.013 (0.029)	-0.021 (0.019)	0.109 (0.027)***	0.108 (0.031)***	0.155 (0.033)***	0.095 (0.032)***	0.071 (0.068)	0.199 (0.067)***
... × SAM	0.039 (0.029)	0.006 (0.040)	0.021 (0.041)	0.022 (0.035)	0.012 (0.035)	0.050 (0.035)	0.031 (0.047)	0.055 (0.044)	-0.070 (0.072)	-0.119 (0.062)*
... × SAM+Cam	0.031 (0.030)	0.012 (0.039)	0.024 (0.041)	0.013 (0.032)	0.071 (0.034)**	0.056 (0.039)	0.064 (0.044)	0.045 (0.042)	-0.000 (0.080)	-0.133 (0.076)*
... × SAM+Score	0.026 (0.032)		0.046 (0.038)		0.023 (0.035)		-0.008 (0.043)		-0.071 (0.072)	
Observations	24719	13668	24719	13668	24700	13655	24700	13655	24719	13668
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Control variables include sex, age dummies, both parents' education, baseline outcome, dummy variables for missing controls (one for each control variable), school-level mean scores, and dummy variables for whether the school is a private school and whether it was among the three control schools who became TSA-ineligible due to the change in the government's definition of remoteness. Standard errors are clustered at the school level. \*/\*\*/\*\* denotes 10/5/1 percent significance levels

## Impact on Parent Satisfactions, Aspirations Improved Satisfactions and Aspirations

	Satisfaction with child learning in				Considers school to be good/very good		Prefers child pursues university over working	
	Indonesian		Mathematics		2018 (5)	2019 (6)	2018 (7)	2019 (8)
	2018 (1)	2019 (2)	2018 (3)	2019 (4)				
SAM	-0.0171 (0.074)	0.0356 (0.070)	-0.0489 (0.079)	0.0575 (0.069)	0.0474 (0.019)**	0.0496 (0.019)***	0.0946 (0.027)***	0.0547 (0.033)*
SAM+Cam	0.0305 (0.073)	0.314 (0.073)***	-0.0178 (0.076)	0.309 (0.070)***	0.0522 (0.019)***	0.0546 (0.017)***	0.0877 (0.028)***	0.0677 (0.031)**
SAM+Score	0.00339 (0.081)		0.0186 (0.085)		0.0510 (0.019)***		0.0765 (0.026)***	
Control group mean	4.745	4.973	4.580	4.784	0.913	0.904	3.512	3.483
Test of equality (P-val)								
SAM v. SAM+Cam	0.515	0.000	0.699	0.001	0.662	0.679	0.801	0.690
SAM+Cam v. SAM+Score	0.733		0.679		0.908		0.663	
SAM v. SAM+Score	0.792		0.446		0.748		0.475	
R2	0.992	0.996	0.989	0.995	0.999	0.999	0.977	0.977
Observations	5377	3875	5377	3875	5291	3874	5377	3875
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

*Notes:* Student-level control variables include sex, age dummies, both parents' education, whether the respondent is the child's mother, and the baseline outcome. School-level control variables include dummy variables for whether the school is a private school and whether it was among the three control schools who became TSA-ineligible due to the change in the government's definition of remoteness. Controls also include dummy variables for missing controls (one for each control variable). Standard errors are clustered at the school level. \*/\*\*/\*\* denotes 10/5/1 percent significance levels

	2018		2019	
	Appreciated by [...]		Appreciated by [...]	
	district (1)	village (2)	district (3)	village (4)
SAM	0.002 (0.171)	0.189 (0.149)	0.420 (0.177)**	0.488 (0.161)***
SAM+Cam	0.398 (0.173)**	0.239 (0.150)	0.273 (0.178)	0.557 (0.161)***
SAM+Score	0.506 (0.170)***	0.166 (0.147)		
TSA-receiving teacher	0.402 (0.179)**	-0.049 (0.155)	0.521 (0.191)***	0.414 (0.173)**
... × SAM	0.202 (0.226)	0.097 (0.196)	0.063 (0.235)	-0.248 (0.213)
... × SAM+Cam	-0.089 (0.228)	0.141 (0.197)	0.229 (0.236)	-0.088 (0.214)
... × SAM+Score	-0.073 (0.226)	0.361 (0.196)*		
Total impacts on TSA receivers				
SAM	0.204 (0.150)	0.286** (0.130)	0.484*** (0.157)	0.240* (0.142)
SAM+Cam	0.309** (0.148)	0.380*** (0.129)	0.502*** (0.156)	0.469*** (0.141)
SAM+Score	0.433*** (0.151)	0.527*** (0.131)		
Control group mean	4.35	4.97	4.50	4.94
Observations	1773	1773	1254	1254
Controls	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes

Impact on Teacher  
Satisfactions

Teachers in  
Treatment  
Schools Feel  
More  
Appreciated

Notes: Individual controls include sex, age, education, and the baseline outcome. School-level controls include school-level mean scores for the outcome, the total number of teachers and civil-servant teachers, the total number of students, and dummy variables for whether the school is a private school and whether it was among the three control schools who became TSA-eligible due to the change in the government's definition of remoteness. Controls also include dummy variables for missing controls (one for each control variable). Standard errors are clustered at the school level. \*/\*\*/\*\* denotes 10/5/1 percent significance levels



	2018		2019	
	Satisfaction of [...]		Satisfaction of [...]	
	salary (1)	job (2)	salary (3)	job (4)
SAM	0.464 (0.173)***	0.087 (0.061)	0.548 (0.179)***	-0.019 (0.068)
SAM+Cam	0.686 (0.174)***	0.236 (0.062)***	0.687 (0.180)***	0.002 (0.068)
SAM+Score	0.824 (0.171)***	0.150 (0.061)**		
TSA-receiving teacher	1.092 (0.180)***	0.145 (0.064)**	1.104 (0.193)***	0.099 (0.073)
... × SAM	-0.440 (0.228)*	-0.037 (0.081)	-0.493 (0.237)**	0.001 (0.089)
... × SAM+Cam	-0.484 (0.229)**	-0.187 (0.082)**	-0.367 (0.238)	-0.042 (0.090)
... × SAM+Score	-0.285 (0.228)	-0.098 (0.081)		
Total impacts on TSA receivers				
SAM	0.024 (0.151)	0.049 (0.054)	0.056 (0.158)	-0.019 (0.060)
SAM+Cam	0.202 (0.149)	0.049 (0.053)	0.321** (0.157)	-0.040 (0.059)
SAM+Score	0.538*** (0.152)	0.052 (0.054)		
Control group mean	3.96	3.00	4.20	3.05
Observations	1773	1773	1254	1255
Controls	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes

Impact on Teacher  
Satisfactions

Non-TSA  
Teachers More  
Satisfied of  
Salary, Job [Y1]

TSA Teachers  
More Satisfied  
of Salary  
[SAM+Score-Y1,  
SAM+Cam],  
Not Job

Notes: Individual controls include sex, age, education, and the baseline outcome. School-level controls include school-level mean scores for the outcome, the total number of teachers and civil-servant teachers, the total number of students, and dummy variables for whether the school is a private school and whether it was among the three control schools who became TSA-ineligible due to the change in the government's definition of remoteness. Controls also include dummy variables for missing controls (one for each control variable). Standard errors are clustered at the school level. \*/\*\*/\*\* denotes 10/5/1 percent significance levels



# Norms and Credible Threats

## Performance Pay More Effective in Punishing Societies

- **Lab-in-the field** at baseline: public good games with **punishment**
  - Conducted in **180** out of **270 schools**
  - Estimate  $\beta$  = school-specific *punishment gradient* for below-mean contributors
  - Group schools as above v. below-median  $\beta$

	Learning Outcomes		Teacher Presence	
	2018 (1)	2019 (2)	2018 (3)	2019 (4)
SAM	-0.020 (0.066)	-0.064 (0.071)	-0.003 (0.046)	-0.032 (0.089)
SAM+Cam	0.066 (0.059)	-0.007 (0.070)	-0.056 (0.049)	-0.008 (0.085)
SAM+Score	-0.045 (0.061)		-0.028 (0.059)	
Above-Median Punishment	-0.231 (0.067)***	-0.143 (0.072)**	-0.108 (0.064)*	0.010 (0.097)
... $\times$ SAM	0.190 (0.098)*	0.096 (0.102)	0.143 (0.073)**	0.036 (0.134)
... $\times$ SAM+Cam	0.227 (0.095)**	0.216 (0.103)**	0.323 (0.089)***	-0.065 (0.124)
... $\times$ SAM+Score	0.277 (0.095)***		0.049 (0.080)	
Observations	16801	9114	667	432
Controls	Yes	Yes	Yes	Yes
Strata FE	Yes	Yes	Yes	Yes

### Schools with a stronger punishing norm:

- drove impact on **teacher attendance** in **SAM+Cam**
- experienced **greater improvement in learning** in **PP treatments**

*Notes:* Individual controls include sex, age, education, and the baseline outcome. School-level controls include school-level mean scores for the outcome, the total number of teachers and civil-servant teachers, the total number of students, and dummy variables for whether the school is a private school and whether it was among the three control schools who became TSA-ineligible due to the change in the government's definition of remoteness. Controls also include dummy variables for missing controls (one for each control variable). Standard errors are clustered at the school level. \*/\*\*/\*\* denotes 10/5/1 percent significance levels