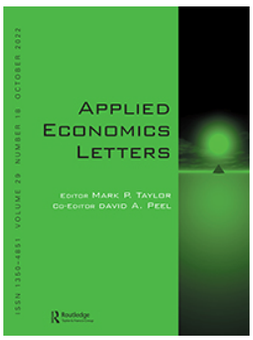

Routledge
Taylor & Francis Group

Applied Economics Letters

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/rael20

Can social recognition for teachers and principals improve student performance? Evidence from India

Puneet Arora, Ishtiaque Fazlul, Tareena Musaddiq & Abhinav Vats

To cite this article: Puneet Arora, Ishtiaque Fazlul, Tareena Musaddiq & Abhinav Vats (2022): Can social recognition for teachers and principals improve student performance? Evidence from India, Applied Economics Letters, DOI: 10.1080/13504851.2022.2099799

To link to this article: https://doi.org/10.1080/13504851.2022.2099799

APPLIED ECONOMICS LETTERS
https://doi.org/10.1080/13504851.2022.2099799

Routledge
Taylor & Francis Group

Can social recognition for teachers and principals improve student performance? Evidence from India

Puneet Arora[a], Ishtiaque Fazlul[b], Tareena Musaddiq[c] and Abhinav Vats[d]

[a]Faculty of Business Administration, Prague University of Economics and Business, Prague; [b]Department of Economics Finance and Quantitative Analysis, Kennesaw State University, U.S.A; [c] Ford School of Public Policy, University of Michigan, U.S.A; [d]Ross School of Business, University of Michigan, U.S.A

ABSTRACT

Empirical evidence on the effectiveness of performance-based rewards for teachers is primarily based on the evaluation of monetary reward schemes. We present results from a randomized evaluation of a teacher and principal incentive programme in India that offered a non-pecuniary recognition reward based on students' test scores on standardized assessments. We find a positive (0.16 SD) yet statistically insignificant effect on student performance when both teachers and principals are incentivized. In schools where only teachers are incentivized, the estimates remain statistically insignificant but are also much smaller in magnitude (0.012 SD). Our findings provide suggestive evidence that recognition rewards may have the potential as a low-cost tool to improve student achievement when both teachers and principals are incentivized. However, further research is required to substantiate the findings and investigate the mechanisms at play.

KEYWORDS

Student achievement; teachers; principals; Non-Monetary incentives; RCT

JEL CLASSIFICATIONS

I21; O15; D91; C93

I. Introduction

Teachers and principals are essential inputs for the education production function. While teacher effectiveness receives much attention, the role of principals is relatively understudied (Grissom, Egalite, and Lindsay 2021). Performance incentive schemes often target teachers but are rarely designed in ways that align the incentives of principals and teachers. Moreover, most teacher incentive schemes use monetary pay-for-performance incentives (De Ree et al. 2018; Barrera-Osorio & Raju 2017; Duflo, Hanna, and Ryan 2012; Muralidharan and Sundararaman 2011; Glewwe, Ilias, and Kremer 2010). In this study, drawing on the literature on social recognition, we consider the potential for a low-cost recognition reward to improve teacher and principal performance.[1]

We partner with the Palwal school district in India and test whether a non-pecuniary reward offered to teachers and principals in secondary schools leads to improvement in student performance, framing the incentive as a recognition award for performance instead of accountability. Student performance on standardized tests improves when both teachers and principals are incentivized (0.16 SD), but the estimates are statistically insignificant at conventional levels and only suggestive. In schools where we incentivized only teachers, the estimates remain statistically insignificant but are much smaller in magnitude (0.012 SD). This study adds to the literature on individual and group level incentives for schools in developing countries. Our work is novel in three respects: (1) we use non-monetary incentives with low out of pocket costs, (2) we use two treatment arms that allow us to study the impacts of incentivizing both teachers and principals (group-level incentives) and only incentivizing teachers (teacher-only incentives), and (3) unlike most RCT studies in developing countries that focus

CONTACT Tareena Musaddiq tareenam@umich.edu Ford School of Public Policy, Joan & Sanford Weill Hall, 735 S State St, Ann Arbor, MI 48109, United States

[1]See Barrera-Osorio et al. (2022) for review of the recognition reward to teachers' literature. It is one of the very few studies from developing countries studying the impact of recognition incentive for teachers on student performance, but they do not involve principals in their incentive design.

on primary schools, we study the effects of this intervention in secondary schools (grades 6–10).[2]

II. Experimental design and implementation

We conducted our experiment within the Palwal block of the Palwal school district of Haryana, India (See Figure A1 for the geographical layout of the district).[3] The Palwal block consists of 84 secondary schools spread across 75 villages. These villages are categorized into 20 clusters for administrative purposes with six villages in each cluster on average. Our unit of treatment is a village. We performed the randomization at the village level to reduce concerns about interactions between principals and teachers across the three intervention conditions and minimize spillover effects. We designed the intervention to cover all 84 secondary schools in the Palwal block ('experimental sample').[4] However, 19 schools could not share data on baseline and endline MAT scores with us and are not part of our final analytic sample. We checked if these schools systematically belong to either the treatment or control groups. Estimates in Table A1 show attrition is not predicted by treatment status.

Our final analytic sample, after accounting for school level attrition, consists of seventh-grade teachers (and principals) in 65 schools in 60 villages of the Palwal block, randomly assigned into three groups (i) Treatment 1: both teachers and principals are incentivized, (ii) Treatment 2: only teachers are incentivized and (iii) Control, through stratified randomization. Table 1 shows the number of schools and villages in each treatment arm. Almost all villages in our sample, with the exception of four villages, have only one secondary school. We, therefore, refer to the school as our unit of treatment in our discussion of results.

The experiment duration was three months (July to September 2018). Teachers and principals in the treatment groups were provided information about the competition named 'Sampoorna' or Complete/Thorough. We conducted the intervention through two information sessions, separately for teachers and principals, in July 2018. The information sessions included: (1) details of the award ceremony to be held later in the year, where local dignitaries would distribute recognition certificates to winners, (2) the judgement criteria for the awards, and (3) the number of awards of each kind. We communicated this information in the presence of the Block Education Officer for each school. An email was later sent to all participating teachers and principals with the program details. The school district reinforced the information through another email in August 2018. Teachers (and principals) in the control schools continued receiving only the business as usual communication from the school district.

Each month the state of Haryana's education department conducts standardized Monthly Assessment Tests (MAT) for all subjects. An employee of a different school marks these tests to remove any potential teacher bias in grading. We used the scores on MAT as a measure for performance. The four types of awards were as follows: (i) 5 Best Teacher Awards (for each subject) awarded to the teachers with the highest average student score in the endline MAT in September, (ii) 5 Highest Growth Teacher Awards (for each subject) awarded to the teachers whose students show highest growth between the baseline (May) and endline (September) tests. In addition, for treatment 1, principals were informed about (i) 5 Best Principal Awards awarded to the principals of the schools where students show highest average score in endline MAT over all subjects combined, and (ii) 5 Highest Growth Principal Awards awarded to the principals of schools where students show highest growth between endline and baseline scores over all subjects combined. Teachers were aware that principals were also part of the competition for this treatment arm and may also receive awards.

[2]Prior interventions in education directly incentivizing both teachers and principals are rare. Muralidharan and Sundararaman (2011), one notable study, evaluated a group-level monetary incentive program for the school as a whole. However, unlike our study, it did not offer direct encouragement for teachers and principals to coordinate.

[3]The Palwal school district in Haryana overall has 251 public schools spread over 4 blocks (Hassanpur, Hathin, Hodal and Palwal). We partnered with the Palwal *block* within the larger Palwal *school district* for our experiment. Palwal block houses 101 public schools across 119 villages. 84 of these 101 schools offer Grade 7.

[4]The experimental sample had 84 schools across 75 villages. The distribution across the treatment arms was as follows: (i) Treatment 1 had 26 schools from 24 villages, (ii) Treatment 2 had 29 schools from 25 villages and, (iii) Control group had 29 schools across 26 villages.

Table 1. Descriptive statistics, by treatment and control schools.

	Full Sample	Control	T1 or T2	T1	T2
Baseline Score (Raw)	14.46	15.20	14.13	14.22	14.06
	(8.74)	(10.07)	(8.08)	(6.96)	(8.93)
Baseline Score	0	0.09	−0.04	−0.03	−0.05
(Standardized)	(1)	(1.16)	(0.92)	(0.80)	(1.02)
Female	0.51	0.58	0.48	0.47	0.48
	(0.50)	(0.49)	(0.50)	(0.50)	(0.50)
Endline Score (Raw)	18.33	18.14	18.41	18.04	18.77
	(7.48)	(7.96)	(7.26)	(7.28)	(7.23)
Endline Score	0	−0.02	0.01	−0.04	0.06
(Standardized)	(1)	(1.07)	(0.97)	(0.97)	(0.96)
Observations	8,405	2,594	5,811	2,670	3,141
No. of schools	65	20	45	20	25
No. of villages	60	20	40	19	21

Baseline score is the average score across all subjects for the May MAT. Endline score is the average score across all subjects for the September MAT. Both May and September MAT for all subjects were marked out of a total of 40 points. Baseline Score (Standardized) and Endline Score (Standardized) are standardized scores with respect to subject specific mean and standard deviation. Female is a binary variable equal to one if the student is female and zero otherwise. Average scores by subject are shown in Table A1 and balance check on baseline variables is shown in Table A2. Standard deviation in parentheses.

All participating teachers and principals were informed that if a student is absent on the test day, their score would be entered as zero in the calculation. This procedure was to curb any strategic manipulation by teachers, such as encouraging low-scoring students not to show up on the test day or incorrectly marking them as absent even if they appear. Additionally, to prevent any student cheating behaviour, whether independently or with the help of a teacher or principal, the education department appoints multiple teams of external evaluators who pay surprise visits to the examination halls.

III. Results

Table 1 shows descriptive statistics for the entire sample and by treatment and control schools (For descriptive statistics by subject, see Table A2). On average, students scored 14.46 points out of 40, across all subjects, on the baseline MAT conducted in May. Balance check in Table A3 shows that the sample is balanced with respect to the baseline May MAT score, with no statistically significant difference between the average score for treatment and control schools. In terms of magnitude, however, the average score for control students is 0.13 SD higher than treatment students at baseline (See Table 1). This may potentially be due to selective attrition at the school level. We control for baseline performance in our main estimations.

The only baseline demographic characteristic made available to us by the school district is student gender. Table 1 shows that control schools have a much higher average proportion of female students (58%) than treatment schools (48%). For reasons explained earlier in our experimental design, we stratified our intervention by cluster and not by the gender that the schools served (i.e. all boys, all girls or co-education), resulting in a lower proportion of girls in treatment than in control schools. We control for student gender in our estimations and argue for conditional independence.

Students in our sample scored 18.33 (out of 40 points) on average in the endline MAT conducted in September (See Table 1).[5] While the average endline scores between control and treatment groups are similar, raw differences in the average scores do not account for differences across clusters, subjects, and gender. We account for these factors and the correlation in standard error in estimating the treatment effects. We estimate the impact of the treatment on the endline standardized test scores through Ordinary Least Squares, controlling for baseline test scores, student gender, and cluster fixed effects.

Table 2 shows the results of our estimations. Averaging across all subjects, students in schools that incentivized both teachers and principals scored 0.16 standard deviations (SD) higher than those in control schools, but the estimate is imprecisely estimated. In comparison, the impact from incentivizing teachers only is much smaller in magnitude (0.012 SD)[6] Barrera-Osorio et al. (2022), in an intervention similar to ours, estimates a much larger effect (0.13 SD) of offering recognition

[5]Our analytic sample consists of students who appear for both the May and September MAT. We check for systematic attrition at the student level. Results are shown in Table A4. We find an overall attrition rate of 5.5%. Attrition rates for T1, T2, and the control group are 6%, 5.79%, and 4.6%, respectively.

[6]For robustness, Table 2 also reports randomization inference p-values (with 1000 repetitions; Young 2019) and the wild bootstrapped p-values (with 999 repetitions; Cameron, Gelbach, and Miller 2008) below each estimate. The results do not change qualitatively. We also do not find any clear distributional effects by student baseline performance or meaningful difference in estimated treatment effects by teacher gender for either treatments. While it would have been interesting to study other heterogeneous effects and the potential mechanisms at play, we are limited by the administrative data restrictions with respect to what information could be availed for the study..

Table 2. Effect of treatment on student test scores.

	(1)	(2)	(3)	(4)	(5)
	All Subjects	Maths	English	Science	Social Studies
T1: Teacher and Principal	0.160	0.227	0.227	0.198*	0.064
	(0.107)	(0.183)	(0.162)	(0.099)	(0.105)
	[0.266]	[0.349]	[0.301]	[0.151]	[0.603]
	{0.260}	{0.318}	{0.282}	{0.131}	{0.591}
T2: Teacher Only	0.012	0.191	−0.051	−0.118	−0.002
	(0.085)	(0.180)	(0.133)	(0.102)	(0.091)
	[0.899]	[0.411]	[0.772]	[0.404]	[0.986]
	{0.902}	{0.409}	{0.759}	{0.367}	{0.988}
Observations	8,405	2,165	1,910	2,161	2,169
R-squared	0.423	0.377	0.496	0.524	0.529

Outcome variable in all columns is the standardized test score for endline MAT. Scores are standardized using subject specific mean and standard deviation. All regressions control for student's standardized baseline test score, student gender and cluster fixed effects. Additionally, Column 1 also controls for subject fixed effects. Fifty five out of 65 schools offered English as a subject due to which the sample size for English is smaller than other subjects. Standard errors are clustered by village. The randomization inference p-values (with 1000 repetitions; Young 2019) and the wild bootstrapped p-values (with 999 repetitions; Cameron, Gelbach, and Miller 2008) are given in square and curly parentheses below each estimate, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

incentives to teachers on student performance in Guinea. Their estimate, however, like ours is statistically insignificant.

Evans and Yuan (2022) studies standardized effect sizes in international education studies and identifies a median effect size of 0.10 SD on learning. For studies with sample sizes similar to ours (~2000 students per subject), the effect sizes range from −0.49 to 0.92 SD. Our estimates are within this range, with the effect sizes for teacher and principal level treatment above the median and for teacher only treatment below the median. We note, however, that our estimates may be impacted by a higher rate of attrition of schools in the control group compared to the treatment group (See Table A1 for details).

IV. Conclusion

In this study, we expand the literature on teacher incentives by providing new evidence on recognition rewards, which are understudied in the literature. The literature on principals or teacher incentives often ignores the possible synergies that may exist in incentivizing both, as opposed to only teachers or principals. Our findings suggest that a low-cost, non-monetary social recognition reward aligning the incentives for teachers and principals may have the potential of being effective in improving student performance. The magnitude of the impact is much smaller when only teachers are offered the incentive. An important shortcoming for our experiment relates to school level attrition- we lose a greater proportion of control schools (0.31) compared to treatment schools (0.18) and though the difference is not statistically significant, the magnitude suggests that this selection may impact the estimates. Given our imprecise estimates and these limitations, we implore future studies to test this idea further – with a larger sample, with addition of a principal only treatment, with long-term effect of repeated incentives, and with more quantitative and qualitative evidence on the underlying mechanism.

Acknowledgement

Puneet Arora is funded by the Excellent Teams Project of the Faculty of Business Administration, Prague University of Economics and Business (IP310031). Tareena Musaddiq's postdoc fellowship is supported by training grant R305B170015 from the U.S. Department of Education's Institute of Education Sciences.

Disclosure statement

Abhinav Vats was formerly Chief Minister's Good Governance Associate for the Palwal District in India. He is currently pursuing his M.B.A at Ross School of Business and no longer works for the Indian government in any capacity.

ORCID

Tareena Musaddiq http://orcid.org/0000-0002-7004-8961

Ethics approval

This study received the ethics approval from Georgia State University: IRB Protocol ID# H18636. Consent requirement was waived by IRB.

Data availability statement

The experiment for this study was conducted with schools of Palwal School District. The study uses administrative data on student, teachers, principals and schools accessed from the district's education department. The data shared by the school district is restricted and was shared with the study team under an MOU for this study. The study team cannot share this data without a formal data sharing request approved from the administrators of Palwal School District.

References

Barrera-Osorio, F., J. Cilliers, M. H. Cloutier, and D. Filmer. 2022. "Heterogenous Teacher Effects of Two Incentive Schemes: Evidence from a low-income Country." *Journal of Development Economics* 156: 102820. doi:10.1016/j.jdeveco.2022.102820.

Barrera-Osorio, F., and D. Raju. 2017. "Teacher Performance Pay: Experimental Evidence from Pakistan." *Journal of Public Economics* 148: 75–91. doi:10.1016/j.jpubeco.2017.02.001.

Cameron, A. C., J. B. Gelbach, and D. L. Miller. 2008. "Bootstrap-based Improvements for Inference with Clustered Errors." *The Review of Economics and Statistics* 90 (3): 414–427. doi:10.1162/rest.90.3.414.

De Ree, J., K. Muralidharan, M. Pradhan, and H. Rogers. 2018. "Double for Nothing? Experimental Evidence on an Unconditional Teacher Salary Increase in Indonesia." *The Quarterly Journal of Economics* 133 (2): 993–1039. doi:10.1093/qje/qjx040.

Duflo, E., R. Hanna, and S. P. Ryan. 2012. "Incentives Work: Getting Teachers to Come to School." *American Economic Review* 102 (4): 1241–1278. doi:10.1257/aer.102.4.1241.

Evans, D., and F. Yuan. 2022. "How Big are Effect Sizes in International Studies." *Educational Evaluation and Policy Analysis, Forthcoming* 016237372210796. doi:10.3102/01623737221079646.

Glewwe, P., N. Ilias, and M. Kremer. 2010. "Teacher Incentives." *American Economic Journal. Applied Economics* 2 (3): 205–227.

Grissom, J. A., A. J. Egalite, and C. A. Lindsay, 2021. "How Principals Affect Students and Schools: A Systematic Synthesis of Two Decades of Research." New York: Wallace Foundation. Available at http://www.wallacefoundation.org/principalsynthesis

Muralidharan, K., and V. Sundararaman. 2011. "Teacher Performance Pay: Experimental Evidence from India." *Journal of Political Economy* 119 (1): 39–77. doi:10.1086/659655.

Young, A. 2019. "Channeling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results." *The Quarterly Journal of Economics* 134 (2): 557–598. doi:10.1093/qje/qjy029.

Appendix

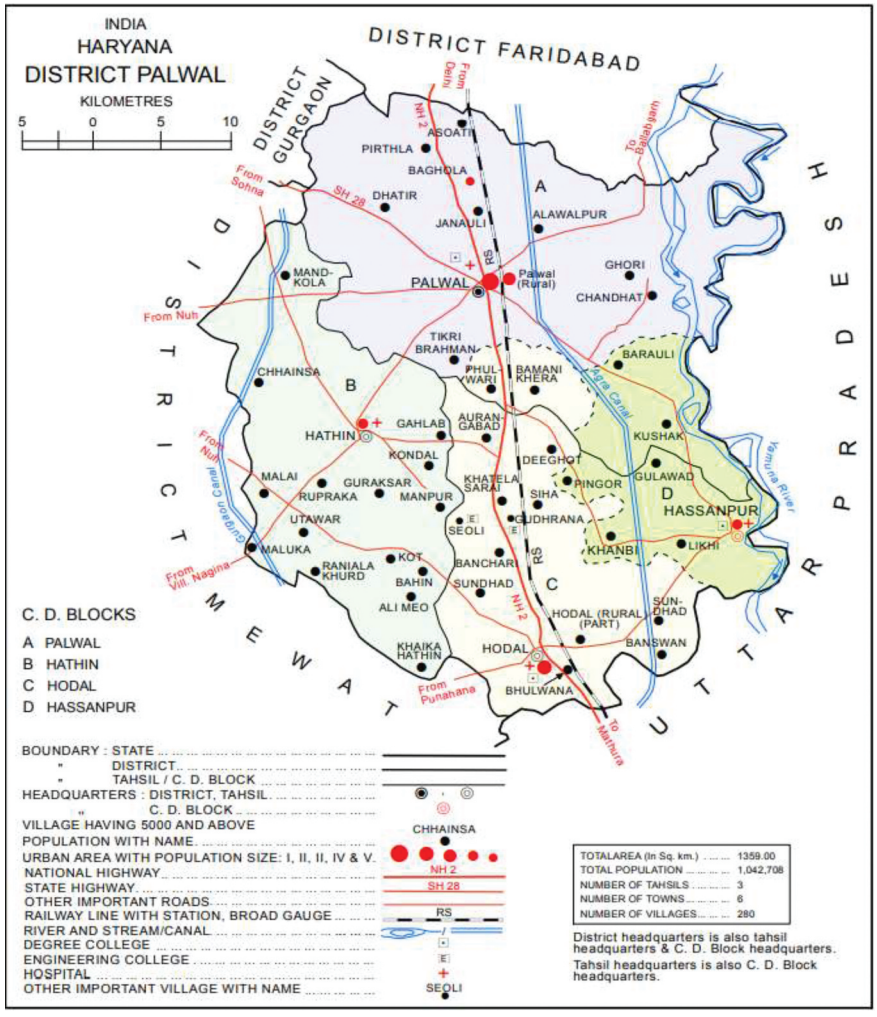

Figure A1. Map of Palwal District, Haryana, India. The figure shows the map of the Palwal district comprising of four blocks: Palwal, Hassanpur, Hathin and Hodal. Our intervention was with schools in the Palwal block.Source: District Census Handbook Palwal, Census of India 2011

Table A1. School Level Attrition from the Experimental Sample

	Attrition
T1: Teacher and Principal	−0.020
	(0.088)
T2: Teacher only	−0.070
	(0.091)
Mean attrition (control group)	0.31
Observations	84
R-squared	0.66

Sample comprises of all treatment and control schools in the experimental sample. Outcome variable is a binary indicator equal to one if the school was part of the experimental sample but did not report data and hence was not part of the analytic sample, and zero if the school did report data. 'T1' is a binary indicator equal to one if both teachers and principals from the school were incentivized and zero otherwise. 'T2' is a binary indicator equal to one if only teachers were incentivized. Attrition by treatment status was as follows: (1) 10 of 29 control schools, (2) 5 out of 26 T1 schools, and (3) 4 out of 29 T2 schools did not share the data. Estimation controls for cluster fixed effects. Standard errors are clustered at the village level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A2. Students Baseline and Endline Test Scores by Subject

Panel A: Raw Scores	Full Sample	Control	T1 or T2	T1	T2
Baseline Maths	14.14	15.87	13.38	13.31	13.44
	(8.64)	(10.22)	(7.72)	(6.37)	(8.76)
Endline Maths	18.43	17.50	18.84	17.98	19.60
	(7.80)	(7.85)	(7.74)	(7.48)	(7.90)
Baseline English	13.86	14.35	13.61	14.35	13.05
	(8.90)	(9.93)	(8.31)	(7.64)	(8.75)
Endline English	17.66	18.05	17.46	17.46	17.76
	(7.26)	(8.19)	(6.72)	(6.72)	(6.66)
Baseline Science	15.35	14.90	15.53	15.37	15.67
	(8.76)	(9.93)	(8.22)	(6.59)	(9.44)
Endline Science	19.43	18.89	19.66	19.43	19.86
	(7.38)	(7.66)	(7.25)	(7.83)	(6.69)
Baseline Social Stud.	14.42	15.64	13.90	13.84	13.94
	(8.60)	(9.91)	(7.92)	(7.19)	(8.51)
Endline Social Stud.	17.72	18.14	17.54	17.27	17.77
	(7.33)	(7.99)	(7.02)	(6.69)	(7.29)
Panel B:Standardized Scores	**Full Sample**	**Control**	**T1or T2**	**T1**	**T2**
Baseline Maths	0	0.20	−0.09	−0.10	−0.08
	(1)	(1.18)	(0.89)	(0.73)	(1.01)
Endline Maths	0	−0.12	0.05	−0.06	0.15
	(1)	(1)	(0.99)	(0.96)	(1.01)
Baseline English	0	0.06	−0.03	0.05	−0.09
	(1)	(1.12)	(0.93)	(0.86)	(0.98)
Endline English	0	0.05	−0.03	−0.08	0.01
	(1)	(1.13)	(0.93)	(0.94)	(0.92)
Baseline Science	0	−0.05	0.02	0	0.04
	(1)	(1.13)	(0.94)	(0.75)	(1.08)
Endline Science	0	−0.07	0.03	0	0.06
	(1)	(1.04)	(0.98)	(1.06)	(0.91)
Baseline Social Stud.	0	0.14	−0.06	−0.07	−0.06
	(1)	(1.15)	(0.92)	(0.84)	(0.99)
Endline Social Stud.	0	0.06	−0.02	−0.06	0.01
	(1)	(1.09)	(0.96)	(0.91)	(1.00)

Statistics shown are sample averages with standard deviation in parentheses. Baseline score and Endline score are average score for the May MAT and September MAT, respectively. In Panel A, both May and September MAT for all subjects are marked out of a total of 40 points. Panel B shows the scores, standardized by mean and standard deviation within subjects. Therefore for the full sample (Panel B, Column 1), all subjects have a mean of zero and standard deviation of 1. T1 is set of students from schools where both teachers and principals were incentivized. T2 is set of students from schools where only teachers were incentivized.

Table A3. Balance Check for Randomization

	(1) Baseline Standardized Score	(2) Female
T1: Teacher & Principal	−0.119	−0.112**
	(0.265)	(0.053)
T2: Teacher Only	−0.134	−0.101**
	(0.250)	(0.049)
Observations	8,405	8,405
R-squared	0.003	0.010

Estimates shown are coefficients from regressing baseline variables on treatment indicators. Outcome in Column 1 is standardized test score at baseline. Outcome in Column 2 is binary indicator equal to one if student is female, zero otherwise. Standard error clustered at village level and shown in parentheses.*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A4. XStudent Level Attrition from the Experimental Sample

	Attrition
T1: Teacher and Principal	−0.002
	(0.012)
T2: Teacher only	0.004
	(0.014)
Observations	8,893
R-squared	0.098

Sample comprises of students in the experimental sample who were enrolled in schools and took the May MAT. Outcome variable is a binary indicator equal to one if the student was not enrolled in September and hence did not take the September (endline) MAT. 'T1' is a binary indicator equal to one if both teachers and principals were incentivized and zero otherwise. 'T2' is a binary indicator equal to one if only teachers were incentivized. Average attrition rate was 5.5%. Estimation controls for cluster fixed effects. Standard errors are clustered at the village level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.