Long Term Effects of Teacher Performance Pay: Experimental Evidence from India

Karthik Muralidharan
UC San Diego, NBER, BREAD, and J-PAL

Conference on the Economics of Public Service Reform

CMPO, University of Bristol,
London, 24 May 2013

Weak State Capacity in Developing Countries

- Improving quality of governance requires both better policies as well as better capacity to *implement* these policies
  - Theoretical literature on role of state capacity in growth and development (Besley and Persson 2009, 2010)
  - Empirical literature highlighting challenges in achieving even basic measures of service delivery such as teacher and health worker attendance (WDR 2004, Chaudhury et al 2006, Muralidharan et al 2013)
  - Disconnect between policy and implementation has led to coinage of the term "flailing state" (Pritchett 2009)
  - "The biggest mystery in skills reform today is how to get something done after everybody who matters in government agrees with you. This is mostly a question of institutional capacity, structure and incentives." (Sabharwal 2013)
- Fundamental determinant of state capacity is the effectiveness of public employees, which can be improved by:
  - Hiring more competent workers (better pay and working conditions)
  - Increasing the effort of existing workers (improve norms of effort, try performance-linked pay?)

Improving Public Sector Worker Effectiveness

- Limited use of performance-pay in the public sector
  - Multi-tasking (Holmstrom & Milgrom 1991)
  - Multiple principals (Dixit 2002)
  - Implementation challenges (Murnane & Cohen 1986)
  - Unions (Ehrenberg & Schwarz 1986; Gregory & Borland 1999)
  - Decision-makers in the public sector are typically not residual claimants of improved productive efficiency (Bandiera, Pratt & Valletti 2009)

- Hence literature on quality of government workers has emphasized:
  - Bureaucratic culture (Wilson 1989) and professionalism (Evans and Rauch 1999)
  - Selecting workers motivated by public interest (Besley & Ghatak 2005)
  - Improving quality/human capital of those who join the public sector (Dolton 2006; Dal Bo, Finan, and Rossi 2011)

- But, there has been a steady increase in the use of performance-linked pay in the private sector (Lemieux, Macleod, & Parent 2009):
  - Growing interest in doing so in the public sector (especially for teachers)

Teacher Performance Pay

- Increased spending on education, but flat trajectories of test scores
  - Hanushek et al (several papers); also true in India
- Strong policy interest in measuring and paying teachers based on measures of performance (based on student learning gains)
  - Teacher salaries are the largest component of spending
  - Several studies show that factors that are rewarded by the status quo (experience, master's degrees in education) are poor predictors of effectiveness
- Performance pay for teachers is being tried in many places
  - Many US states, Teacher Incentive Fund, Race to the Top; Australia; UK; Chile, etc.
  - Small but growing evidence on impact (but almost no long-term evidence)
- Understanding impact is critical for education policy and public employee compensation policy more generally

This Paper

- We present results from a **5-year long** experimental evaluation of both group and individual teacher performance pay in a large representative sample of schools in the Indian state of Andhra Pradesh (AP):
  - Robustness of short-term results (novelty effect, re-optimization, etc)
  - What do outcomes look like for students who **complete their entire primary education** under a system where teachers are rewarded for output?
  - Group and individual performance-pay in the same long-term experiment
  - Longest-running compensation experiment that we know of (in any sector)
  - Can study hysteresis/de-motivation impact of withdrawing incentives

- New estimation techniques for *n*'th year treatment effects – important for cost effectiveness calculations in the presence of test-score decay:
  - Gross vs. Net Treatment Effects
  - Cannot experimentally estimate annual 'gross' treatment effects, though this is probably what matters for long-term outcomes (Chetty et al 2011; Deming 2009)
  - Experimental discontinuation of treatments to demonstrate importance of this
  - Present both parametric and non-parametric estimates of 'gross' effects using all 25 cohorts for whom we have a '1-year' effect

Preview of Results

- Students in schools under the individual teacher incentive program (II) did significantly better than the controls at every duration of program exposure
  - Students in II schools who completed their full primary education (5 years) under this program scored 0.54 SD and 0.35 SD higher in math/language

- They also scored significantly higher on subjects for which there were *no incentives* – scoring 0.52 SD and 0.3 SD higher in science/social studies
  - Also on repeats/non-repeats; MCQ/non-MCQ; mechanical/conceptual questions

- Students in group incentive (GI) schools also do better than controls at all durations of exposure – but individual incentive (II) schools always do better
  - But cannot reject that GI is (1/n) times as effective as II (n = number of teachers)

- Existence of test-score decay means that these are estimates of 'net' treatment effects, which will understate impact relative to discontinuation
  - 'Gross' TE: 0.17 SD/year in math and 0.11 SD/year in language in II schools
  - 0.075 SD/year in math and 0.037 SD/year in language in GI schools
  - '1-year' effect on the discontinued schools is close to zero (and not significant) – suggesting hysteresis/de-motivation were not first-order relative to incentive effects

Related Literature

- Does teacher performance pay improve student learning outcomes?
  - Springer et al (2010) in Tennessee
  - Fryer (2013), Goodman & Turner (2013) in New York City
  - Lavy (2002) and (2008) in Israel
  - Glewwe, Ilias, Kremer (2010) in Kenya
  - Muralidharan and Sundararaman (2011) in India
  - Rau & Contreras (2013) in Chile

- Do bad things happen?
  - Teaching basic as opposed to higher-order skills (Holmstrom, Milgrom 1991)
  - Test preparation instead of longer-term learning (Glewwe et al 2003)
  - Manipulating test-taking population (Jacob 2005; Cullen & Reback 2006)
  - Short-term boosting of caloric content (Figlio & Winicki 2005)
  - Gaming to threshold effects (Neal & Schanzenbach 2010)
  - 'Cheating to the test' (Jacob & Levitt, 2003)

- Suggests that good design is key (and not just evaluation): Neal (2011)
  - Optimal design will be context specific - Lazear (2006), Dixit (2002)
  - Design and implement as well as possible, and test for adverse outcomes

Outline

- Experimental Design
- Results
- Fade out and Cost Effectiveness
- Discussion

Potential concerns were addressed pro-actively in the program design

Potential concern	How addressed
Reduction of intrinsic motivation	• Recognize that framing matters • Program framed in terms of recognition and reward for outstanding teaching as opposed to accountability
Teaching to the test	• Less of a concern given extremely low levels of learning • Research shows that the process of taking a test can enhance learning • Test design is such that you cannot do well without deeper knowledge / understanding
Threshold effects/ Neglecting weak kids	• Minimized by making bonus a function of average improvement of all students, so teachers are not incentivized to focus only on students near some target; • Drop outs assigned low scores
Cheating / paper leaks	• Testing done by independent teams from Azim Premji Foundation, with no connection to the school

Experimental Design

Figure 1: Experiment Design over 5 Years

Treatment	Year 1	Year 2	Year 3	Year 4	Year 5
Control	100	100	100	100	100
Individual Incentive	100	100	100	50	50
Group Incentive	100	100	100	50	50
Individual Incentive Discontinued	0	0	0	50	50
Group Incentive Discontinued	0	0	0	50	50

Treatment Exposure by Cohort

Figure 2 : Nine Distinct Cohorts Exposed to the Interventions

		Year 1	Year 2	Year 3	Year 4	Year 5
One Cohort exposed for five years : 5	Grade 1	5	6	7	8	9
Two Cohorts exposed for four years : 4 , 6	Grade 2	4	5	6	7	8
Two Cohorts exposed for three years : 3 , 7	Grade 3	3	4	5	6	7
Two Cohorts exposed for two years : 2 , 8	Grade 4	2	3	4	5	6
Two Cohorts exposed for one year : 1 , 9	Grade 5	1	2	3	4	5

Specification

$$T_{ijkm}(Y_n) = \alpha + \gamma_{j(Y_0)} \cdot T_{ijkm}(Y_0) + \delta_{II} \cdot II + \delta_{GI} \cdot GI + \beta \cdot Z_m + \varepsilon_k + \varepsilon_{jk} + \varepsilon_{ijk}$$

$$i = Child,\ j = Class, k = School, m = Mandal(Sub - District)$$

Impact of Incentive Programs (Full Sample: Table 5)

	Panel B : Maths				
	One Year	**Two Years**	**Three Years**	**Four Years**	**Five Years**
	(1)	(2)	(3)	(4)	(5)
Individual Incentive	0.175	0.229	0.227	0.425	0.538
	(0.051)***	(0.055)***	(0.062)***	(0.089)***	(0.129)***
Group Incentive	0.127	0.098	0.109	0.137	0.119
	(0.048)***	(0.055)*	(0.055)**	(0.077)*	(0.106)
Observations	34796	21014	12349	5465	1728
R-squared	0.177	0.192	0.213	0.28	0.37
Pvalue II = GI	0.35	0.05	0.08	0.01	0.00

	Panel C : Telugu				
	One Year	**Two Years**	**Three Years**	**Four Years**	**Five Years**
	(1)	(2)	(3)	(4)	(5)
Individual Incentive	0.133	0.180	0.155	0.237	0.350
	(0.043)***	(0.047)***	(0.053)***	(0.062)***	(0.087)***
Group Incentive	0.085	0.024	0.069	0.108	0.139
	(0.044)*	(0.048)	(0.052)	(0.063)*	(0.080)*
Observations	35234	21187	12425	5496	1728
R-squared	0.20	0.21	0.22	0.23	0.30
Pvalue II = GI	0.26	0.00	0.13	0.07	0.02
Cohort/Year/Grade (CYG) Indicator	115 214 313 412 511 621 731 841 951	225 324 423 522 632 742 852	335 434 533 643 753	445 544 654	555

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%

Robustness of Results

- Breakdown results by question type
  - Repeat/non-repeat (~16% repeats in math; ~10% in language)
  - MCQ/non-MCQ (Table 6)
- Mechanical versus conceptual questions
- Re-estimate treatment effects without repeats/MCQ
- Inverse probability weighting to adjust for student attrition (again – no differential attrition, but TE may be different across the distribution)
- Re-estimate with teachers who were always in the program

Impact of Incentive Programs on non-incentive subjects (Table 7)

Table 7: Impact of Performance Pay on Non-Incentive Subjects

	Science					**Social Science**				
	One year	Two Year	Three Year	Four Year	**Five Year**	One year	Two Year	Three Year	Four Year	**Five Year**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Individual Incentives	0.108	0.186	0.114	0.232	**0.520**	0.126	0.223	0.159	0.198	**0.299**
	(0.063)*	(0.057)***	(0.056)**	(0.068)***	**(0.125)*** **	(0.057)**	(0.061)***	(0.057)***	(0.066)***	**(0.113)*** **
Group Incentives	0.114	0.035	0.076	0.168	**0.156**	0.155	0.131	0.085	0.139	**0.086**
	(0.061)*	(0.055)	(0.054)	(0.067)**	**(0.099)**	(0.059)***	(0.061)**	(0.057)	(0.065)**	**(0.095)**
Observations	11765	9081	11133	4997	1592	11765	9081	11133	4997	1592
R-squared	0.259	0.189	0.127	0.160	0.306	0.308	0.181	0.134	0.148	0.211
Pvalue II = GI	0.93	0.03	0.48	0.41	0.01	0.67	0.20	0.19	0.44	0.08
Cohort/Year/Grade (CYG) Indicator	115 214 313	225 324 423	335 434 533 643 753	445 544 654	555	115 214 313	225 324 423	335 434 533 643 753	445 544 654	555

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%

Heterogeneity by Teacher Characteristics

Table 8B: Heterogenous Treatment Effects by Teacher Characteristics

	Dependent Variable : Teacher Value Added (using all cohorts and years)				
Covariates	**Teacher Education**	**Teacher Training**	**Teacher Experience**	**Teacher Salary**	**Teacher Absence**
II	-0.022	-0.120	0.221	0.082	0.132
	(0.134)	(0.129)	(0.113)*	(0.482)	(0.037)***
GI	-0.065	-0.211	0.225	0.573	0.064
	(0.136)	(0.137)	(0.093)**	(0.518)	(0.035)*
Covariate	-0.006	-0.052	-0.027	-0.036	-0.119
	(0.025)	(0.029)*	(0.020)	(0.029)	(0.044)***
II * Covariate	0.049	0.091	-0.035	0.005	0.019
	(0.041)	(0.046)**	(0.044)	(0.052)	(0.078)
GI * Covariate	0.038	0.098	-0.070	-0.056	-0.020
	(0.044)	(0.050)**	(0.037)*	(0.056)	(0.066)
Observations	108560	108560	106592	106674	138594
R-squared	0.057	0.057	0.059	0.058	0.052

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%

Exploring Group Vs. Individual Incentives Further

	Panel B : Maths				
	One Year	**Two Years**	**Three Years**	**Four Years**	**Five Years**
	(1)	(2)	(3)	(4)	(5)
Individual Incentive	0.160***	0.216***	0.210***	0.423***	0.511***
	(0.049)	(0.057)	(0.062)	(0.089)	(0.129)
Group Incentive (Fraction of all teachers)	0.325***	0.231*	0.260*	0.444**	0.147
	(0.123)	(0.134)	(0.151)	(0.220)	(0.274)
Observations	34796	21014	12349	5465	1728
R-squared	0.176	0.190	0.212	0.281	0.369
Pvalue II = GI	**0.14**	**0.91**	**0.73**	**0.92**	**0.18**
	Panel C : Telugu				
	One Year	**Two Years**	**Three Years**	**Four Years**	**Five Years**
	(1)	(2)	(3)	(4)	(5)
Individual Incentive	0.126***	0.180***	0.141***	0.226***	0.324***
	(0.041)	(0.047)	(0.053)	(0.061)	(0.085)
Group Incentive (Fraction of all teachers)	0.228*	0.076	0.132	0.262	0.224
	(0.119)	(0.125)	(0.141)	(0.179)	(0.217)
Observations	35234	21187	12425	5496	1728
R-squared	0.194	0.211	0.214	0.226	0.296
Pvalue II = GI	**0.34**	**0.38**	**0.95**	**0.84**	**0.63**
Cohort/Year/Grade (CYG) Indicator	115 214 313 412 511 621 731 841 951	225 324 423 522 632 742 852	335 434 533 643 753	445 544 654	555

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%

How did Teacher Behavior Change?

Table 9: Teacher Behavior (Observation and Interviews)

Teacher Behavior	Incentive versus Control Schools (All figures in %)						
	Control Schools	Individual Incentive Schools	Group Incentive Schols	P-value (H0: II = Control)	P-value (H0: GI = control)	P-value (H0: II = GI)	Correlation with student test score gains
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Based on School Observation							
Teacher Absence (%)	0.28	0.27	0.28	0.15	0.55	0.47	-0.109***
Actively Teaching at Point of Observation (%)	0.39	0.42	0.40	0.18	0.42	0.58	0.114***
Based on Teacher Interviews							
Did you do any special preparation for the end of	0.22	0.61	0.56	0.00	0.00	0.06	0.108***
What kind of preparation did you do?							
Extra Homework	0.12	0.35	0.32	0.00	0.00	0.12	0.066**
Extra Classwork	0.15	0.39	0.34	0.00	0.00	0.04	0.108***
Extra Classes/Teaching Beyond School Hours	0.03	0.12	0.11	0.00	0.00	0.65	0.153***
Gave Practice Tests	0.10	0.29	0.25	0.00	0.00	0.04	0.118***
Paid Special Attention to Weaker Children	0.06	0.18	0.15	0.00	0.00	0.20	-0.004

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%

Changes in Household Inputs

- Long-term effects of school interventions can be attenuated/amplified by household responses (see Das et al 2011, Pop-Eleches & Urquoila 2011)
- We also collect data on household spending, child time allocation, and perceptions of school quality from parents of children in cohort 5 after 5 years
- Households in II and GI schools report spending less on education by almost 20% (not significant though)
- Households in II and GI schools also report slightly higher perceived academic ability of their children and measures of satisfaction with teachers (again not sig.)
- Overall, it seems like the improvements in school quality from higher teacher effort are not salient enough for parents to adjust behavior much – but we do see some downward adjustments (so estimated effects may be lower bounds of 'production function' effect)

Fade Out and Cost Effectiveness

- Well-established fact that there is fade-out in test score effects of education interventions (Andrabi et al 2011; Rothstein 2010; Jacob, Lefgren & Sims 2009)
- However, there is also a growing literature document significant long-term benefits of education interventions even though the test score gains fade out pretty quickly after the program ends (Deming 2009; Chetty et al 2011)
- So the n-year treatment effects we are estimating are 'net' treatment effects that are the sum of the 'gross' treatment effects and depreciation
- Arguably, it's the 'gross' treatment effects that matter relative to the counterfactual of stopping the treatment (Chetty, Friedman, Rockoff 2011)
    - Medical analog
- But, the n'th year gross treatment effect cannot be estimated experimentally

Specification

$$T_{ijkm}(Y_n) = \alpha + \gamma_{j(Y_0)} \cdot T_{ijkm}(Y_0) + \delta_{II} \cdot II + \delta_{GI} \cdot GI + \beta \cdot Z_m + \varepsilon_k + \varepsilon_{jk} + \varepsilon_{ijk}$$

is okay, but…

$$T_{ijkm}(Y_n) = \alpha + \gamma_{j(Y_{n-1})} \cdot T_{ijkm}(Y_{n-1}) + \delta_{II} \cdot II + \delta_{GI} \cdot GI + \beta \cdot Z_m + \varepsilon_k + \varepsilon_{jk} + \varepsilon_{ijk}$$

…is not!

- But, we have an experimental discontinuation of a sub-set of treated schools
- Allows us to see the treatment effect in later years relative to the counterfactual of discontinuation (which is what we probably need for cost effectiveness)

Continued Vs. Discontinued Cohorts

Table 10 : Long-Term Impact on Continued and Discontinued Cohorts

	Y3 on Y0 Combined	Y4 on Y0 Combined	Y5 on Y0 Combined
GI * discontinued	0.133	0.158	0.132
	(0.070)*	(0.067)**	(0.082)
GI * continued	0.029	0.167	0.117
	(0.073)	(0.089)*	(0.087)
II * discontinued	0.224	0.149	0.098
	(0.082)***	(0.087)*	(0.095)
II * continued	0.166	0.443	0.458
	(0.078)**	(0.095)***	(0.111)***
Observations	10707	9794	4879
R-squared	0.196	0.233	0.249
II continued = II discontinued	0.56	0.01	0.01
GI continued = GI discontinued	0.24	0.93	0.89
Estimation Sample			
Cohort	4,5	4,5	5
Year	3	4	5
Grade	3,4	3,5	5

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%

Estimating Gross TE with all cohorts (OLS)

$$T_{ijkm}(Y_n) = \alpha + \gamma_{j(Y_{n-1})} \cdot T_{ijkm}(Y_{n-1}) + \beta \cdot Z_m + \varepsilon_k + \varepsilon_{jk} + \varepsilon_{ijk}$$

- Use control schools to estimate the coefficient on the lagged test score in a normal value-added model (standard in literature)

- Use this to transform dependent variable into 'gross' value-addition

$$T_{ijkm}(Y_n) - \hat{\gamma}_{j(Y_{n-1})} \cdot T_{ijkm}(Y_{n-1}) = \alpha + \delta_{II} \cdot II + \delta_{GI} \cdot GI + \beta \cdot Z_m + \varepsilon_k + \varepsilon_{jk} + \varepsilon_{ijk}$$

- Main advantage – no longer jointly estimating delta and gamma

- Limitations:
  - Have to assume same gamma in treatment in control
  - Have to assume uniform gamma at all parts of the test score distribution
  - Inconsistent estimates of gamma?
  - But each of these issues can be mitigated and this is what the entire literature on EPF's does

Average Non-Parametric Treatment Effect

- Idea is to compare the Y(n) scores for treatment and control students who *start at the same Y(n-1) score*. Implemented as follows:

$$ATE = \frac{1}{100}\sum_{i=1}^{100}\left[\overline{T(Y_n(BG))} - \overline{T(Y_n(C))} \middle| T(Y_{n-1}(BG)), T(Y_{n-1}(C)) \in P_{i,n-1}(C)\right]$$

- X-axis plots Y(n-1) score by percentile of the control distribution
- Y-axis plots the Y(n) score for the control students, and the Y(n) score for treatment students who are *in the same percentile* of the Y(n-1) distribution
- Difference is the non-parametric TE at that percentile – integrate over the pdf of the *control distribution* for the average treatment effect (simple average here).
- Assumptions:
  - Decay only depends on current score (and not on how you got there)
  - Same treatment effects at different percentiles of unobservables
- Advantages :
  - Do not need gamma to be constant at all points of test score distribution
  - Do not even need to estimate gamma for the gross TE (like in Y1 of experiment)

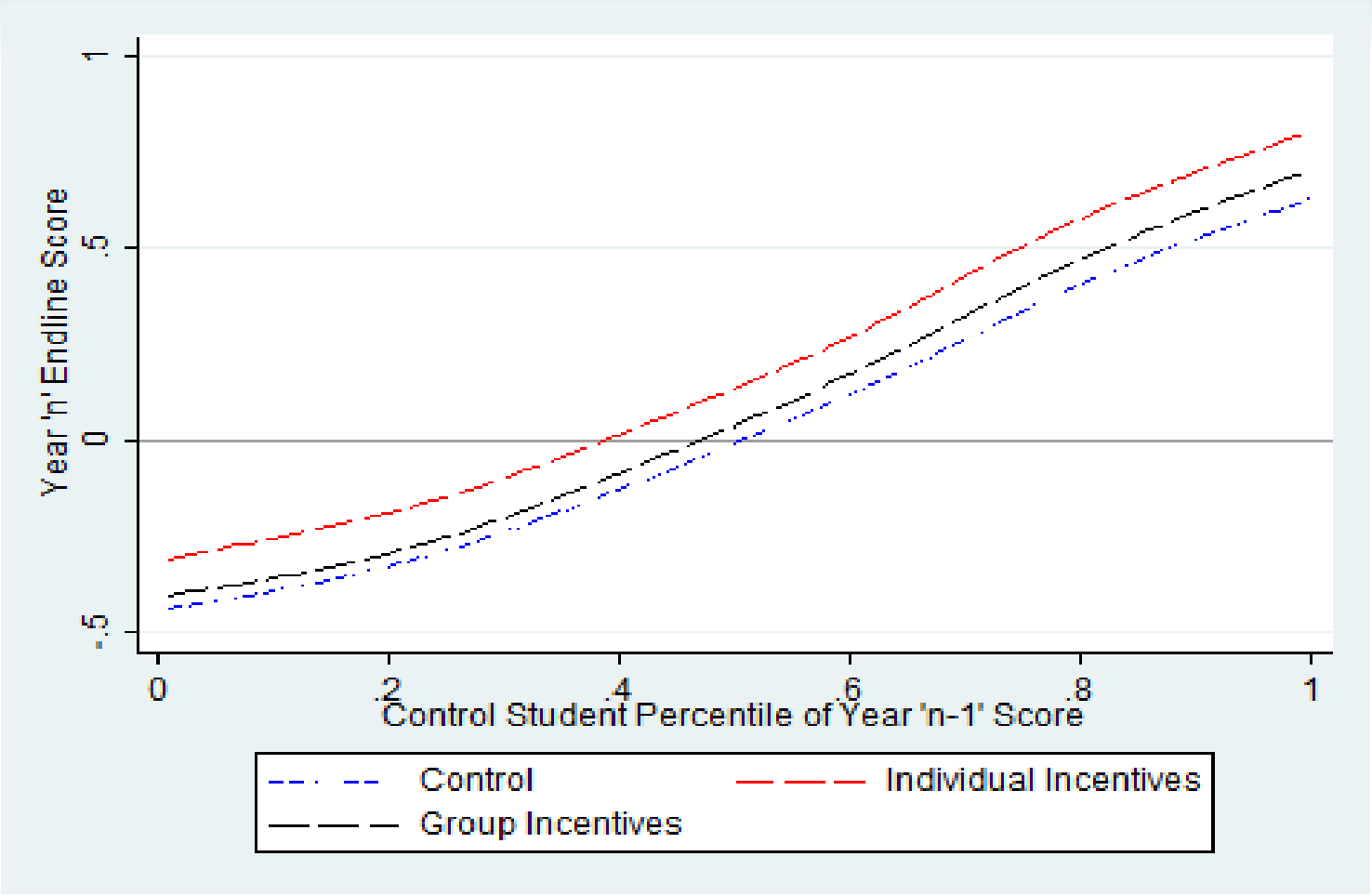
One-year Gross TE

Table 11 : Average "Gross" One-Year Treatment Effect of Teacher Incentive Programs

	Panel A: OLS with Estimated gamma			Panel B: Average non-parametric Treatment Effect (Based on Figure 4)		
	Combined	**Maths**	**Telugu**	**Combined**	**Maths**	**Telugu**
II	0.135	0.164	0.105	0.150	0.181	0.119
	(0.031)***	(0.036)***	(0.027)***			
95% CI	[0.074 , 0.196]	[0.093 , 0.235]	[0.052 , 0.158]	[0.037 , 0.264]	[0.051 , 0.301]	[0.009 , 0.228]
GI	0.064	0.086	0.043	0.048	0.065	0.032
	(0.028)**	(0.031)***	(0.026)			
95% CI	[0.009 , 0.119]	[0.0252 , 0.147]	[-0.008 , 0.094]	[-0.058 , 0.149]	[-0.047 , 0.176]	[-0.083 , 0.145]
Constant	-0.030	-0.029	-0.032			
	(0.018)	(0.021)	(0.017)*			
Observations	165300	82372	82928			
R-squared	0.046	0.054	0.041			
II = GI	0.0288	0.0364	0.0299			

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%

Average Non-Parametric Treatment Effect (with 25 1-year comparisons)



Hysteresis/Discouragement Effect of Discontinuation

Table 12: One-Year Effect of Discontinuation (Y4 on Y3; Cohorts 4-8)

	Panel A: OLS with estimated gamma		
	Combined	Math	Telugu
	(1)	(2)	(3)
II*continue	0.264	0.354	0.175
	(0.056)***	(0.074)***	(0.046)***
II*discontinue	-0.014	-0.017	-0.012
	(0.054)	(0.062)	(0.051)
GI*continue	0.105	0.105	0.105
	(0.060)*	(0.070)	(0.056)*
GI*discontinue	0.049	0.101	-0.003
	(0.054)	(0.065)	(0.050)
N	25706	12832	12874
R-sq	0.140	0.182	0.114
II*continue = II*discontinue (p-value)	0.00	0.00	0.00
GI*continue = GI*discontinue (p-value)	0.43	0.95	0.10

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%

Cost Effectiveness

- The most relevant policy comparison may be with class-size reductions (which is what is being implemented under the “Right to Education” act in India).
- Estimates from OLS and panel data suggest that halving class size at the school level improves test scores by ~0.20 – 0.25 SD
- Halving class size would cost around Rs. 450,000 per school/year
- The average spending in the individual incentive program was Rs. 10,000 per school/year (~Rs. 15,000 including administrative costs)
- Suggests that implementing an individual performance pay program may be around 15-20 times more cost effective than the default “school quality” intervention of reducing class sizes with regular teachers

Concluding Thoughts

- Performance pay for teachers (at least at the individual level) appears to have a large effect on student learning outcomes in primary schools in Andhra Pradesh (0.54 SD in math and 0.35 SD in language for the 5-year cohort)
  - Continued gains on all components of tests as well as non-incentive subjects
  - Long-term data results suggest sustained increases in teacher effort/effectiveness

- The divergence of GI and II is quite interesting – especially since the groups are quite small (median school has 3 teachers)
  - Cannot reject that the GI is (1/N) times as effective as II (N = No. of Teachers)
  - Context of low complementarity of effort? Difficult to monitor effort intensity?

- Importance of accounting for decay in long-term experimental evaluations
  - The N-year treatment effect may significantly understate the true effect

- No evidence of positive hysteresis or de-motivation when incentives withdrawn

- Highlights potential for compensation reforms in improving public sector productivity (MUCH more cost effective than default policies to improve schooling)