Appendix A   Further tables and figures

Table A.1 Treatment Status Regressions on Attrition Status

	(1) Attrite	(2) Age	(3) Female	(4) Reading score	(5) Teacher attrition
Attrite		0.169**	-0.0647**	-0.0175	
		(0.0678)	(0.0309)	(0.0717)	
Training	0.00605	0.0859	-0.0309	-0.204*	-0.0239
	(0.0222)	(0.0535)	(0.0241)	(0.121)	(0.0363)
Coaching	-0.0136	-0.0251	-0.0139	0.0822	-0.0234
	(0.0183)	(0.0514)	(0.0232)	(0.152)	(0.0378)
Attrition x Training		-0.0518	0.0900*	-0.0262	
		(0.102)	(0.0504)	(0.113)	
Attrition x Coaching		0.0176	0.00665	-0.103	
		(0.0961)	(0.0531)	(0.113)	
Strata fixed effects?	Yes	Yes	Yes	Yes	Yes
Observations	3,539	3,523	3,518	3,539	2,951
R-squared	0.010	0.018	0.003	0.059	0.013
Mean attrition	0.168				0.208

Notes: Each column represents a separate regression. Column headings indicate the dependent variable. "Attrite" is a dummy variable equal to one if the pupil was not surveyed at endline. "Teacher attrition" is a dummy variable equal to one if the pupil's teacher was not surveyed at endline. Standard errors are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1.

Table A.2 Comparing lesson observation schools with full sample

	(1)	(2)	(3)	(4)	(5)
	Pupil reading proficiency				*Location*
	Value-added	Endline	Midline	Baseline	Rural
In sample	0.0594	-0.00586	0.0200	-0.0284	-0.250***
	(0.0724)	(0.0814)	(0.0748)	(0.119)	(0.0692)
Observations	3,148	3,148	3,337	3,539	180
R-squared	0.001	0.000	0.000	0.000	0.087
Sample mean	0.0368	0.00873	0.0304	-0.0180	0.633

Notes: Each column represents a separate regression on a dummy variable indicating whether the pupil/school is in the sample where we conducted the lesson observation. In columns (1) to (4) the data is at the individual level; in column (5) the data is at the school level. In column (1) the dependent variable is midline reading proficiency, and the regression includes the full set of controls used in Table 2. Standard errors are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1

Table A.3 Descriptive and balance statistics - Lesson observations sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Control	*Training*		*Coaching*			
	Mean	Coef.	Std error	Coef.	Std error	Obs	R-squared
Pupil Characteristics							
Age	6.481	0.117	(0.0781)	0.0263	(0.0767)	1,194	0.021
Female	0.479	-0.0634	(0.0423)	-0.0653*	(0.0356)	1,191	0.008
Reading proficiency	0.0404	-0.244	(0.253)	0.171	(0.224)	1,198	0.157
Teacher Characteristics							
Diploma or degree	0.947	0.0451	(0.0444)	0.0559	(0.0547)	88	0.117
Age	48.92	0.108	(2.882)	0.368	(2.875)	89	0.103
Female	1	-0.0320	(0.0307)	0.00641	(0.0153)	87	0.213
Class size	42.17	-3.470	(2.692)	-7.309**	(3.057)	87	0.253
Multi-grade	0.0619	-0.0570	(0.0379)	-0.0183	(0.0382)	88	0.407
Comprehension test	0.663	-0.0663	(0.0741)	0.0198	(0.0808)	88	0.128
School characteristics							
Majority parents - highschool	0.443	-0.247	(0.179)	-0.00595	(0.170)	59	0.277
Rural	0.850	-0.0312	(0.151)	-0.144	(0.172)	60	0.239
Bottom quintile (SES)	0.463	0.0412	(0.108)	-0.0935	(0.0818)	60	0.791
Pass rate (ANA)	55.35	-0.215	(1.446)	0.773	(1.771)	60	0.542

Notes: Each row indicates a separate regression on treatment dummies controlling for strata indicators. Column one shows the control mean, columns (2) and (4) the coefficient on the two treatment dummies. Standard errors (columns (3) and (5)) are clustered at the school level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A.4 Results with trimmed sample for Training

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		Endline				Baseline		
Training	0.116	0.117	0.113	0.124	-0.209*	-0.166	-0.116	-0.0730
	(0.0791)	(0.0792)	(0.0786)	(0.0791)	(0.118)	(0.118)	(0.119)	(0.120)
Coaching	0.242***	0.242***	0.243***	0.243***	0.0666	0.0665	0.0665	0.0665
	(0.0778)	(0.0778)	(0.0778)	(0.0776)	(0.146)	(0.146)	(0.146)	(0.145)
Percentile trimmed	0	5th	10th	15th	0	5th	10th	15th
Observations	2,951	2,918	2,878	2,843	3,539	3,499	3,445	3,399
R-squared	0.169	0.167	0.167	0.165	0.058	0.053	0.049	0.048

Notes: Each column represents a separate regression. In columns (1) to (4) the outcome variable is endline aggregate reading proficiency, and the regressions are estimated using the same set of controls as Table 3. In columns (5) to (8) the outcome variable is baseline reading proficiency, and regressions are estimated using the same set of controls as Table 1. Moving from left to right, a larger share of the sample of students in the Training arm are excluded: the bottom 5 percent, 10 percent, and 15 percent respectively, in terms of baseline aggregate reading proficiency.

Table A.5. Dynamic impacts in terms of standard deviations

VARIABLES	(1) Aggregate score	(2) Phon. awareness	(3) Letters	(4) Words	(5) Non-words	(6) Paragraph	(7) Writing
Training	0.129	0.0919	0.0714	0.0682	0.108**	0.0925*	0.235**
	(0.0798)	(0.0716)	(0.0755)	(0.0520)	(0.0518)	(0.0518)	(0.0918)
Training x endline	-0.0124	0.0478	-0.0316	0.0201	0.00301	0.0213	-0.139*
	(0.0620)	(0.0799)	(0.0715)	(0.0696)	(0.0678)	(0.0653)	(0.0811)
Coaching	0.141*	0.222***	0.124	0.0492	0.0843*	0.0642	0.181*
	(0.0804)	(0.0725)	(0.0805)	(0.0509)	(0.0507)	(0.0500)	(0.104)
Coaching x endline	0.100	-0.0934	0.0596	0.194***	0.199***	0.199***	-0.0407
	(0.0661)	(0.0851)	(0.0923)	(0.0672)	(0.0705)	(0.0625)	(0.0915)
Endline	-0.00645	0.856***	0.633***	0.780***	0.765***	0.762***	0.0461
	(0.0429)	(0.0549)	(0.0611)	(0.0481)	(0.0445)	(0.0430)	(0.0499)
Observations	6,190	6,190	6,190	6,190	6,190	6,191	6,190
R-squared	0.171	0.251	0.232	0.296	0.275	0.284	0.124
Training x Endline=Coaching x Endline:P-value	0.0956	0.108	0.246	0.0123	0.00974	0.00883	0.322
Midline	-0.174	-1.452	0.0204	-1.002	-1.177	-1.173	-0.416
Endline	-0.0479	0.394	0.311	0.383	0.352	0.358	-0.0287

Notes. See Table 4.

Table A.6 Breakdown of variable costs by treatment arm

	Training		Coaching	
	Sub-total	%	Sub-total	%
Materials Provision	$ 20,985.90	18%	$ 19,799.69	12%
Transport	$ 3,687.70	3%	$ 32,268.53	20%
Accommodation and venue	$ 42,197.15	37%	$ 540.19	0%
Catering	$ 82.63	0%	$ 82.63	0%
Salary				
---Program management	$ 30,180.00	26%	$ 50,300.00	31%
---Coaches	$ -	0%	$ 55,384.62	35%
---Trainers	$ 12,038.46	11%	$ -	0%
---Motivational visits / calls	$ 3,192.31	3%	$ -	0%
---Training of trainers/coaches	$ 1,846.15	2%	$ 1,846.15	1%
Total	**$ 114,210.31**	100%	**$ 160,221.80**	100%

Table A.7. Mediation analysis using sequential g estimator

VARIABLES	(1)	(2)	(3)
Training	0.224***	0.208**	0.187**
	(0.0850)	(0.0949)	(0.0838)
Coaching	0.284***	0.122	0.0909
	(0.0861)	(0.103)	(0.0818)
Group-guided reading		0.122	
		(0.0884)	
Print-richness in classroom		-0.0883*	
		(0.0512)	
Routine		0.0382	
		(0.0724)	
Curriculum coverage		-0.0441	
		(0.0413)	
Group-guided reading x Training		0.156	
		(0.138)	
Group-guided reading x Coaching		0.318**	
		(0.142)	
Observations	2,295	2,295	2,295
R-squared	0.166	0.186	0.155

Notes. Each row represents a separate regression, including the same set of controls as Table 1, restricting the sample to students for whom we have teacher survey data. The dependent variable in columns (1) and (2) is the aggregate reading score. The dependent variable in column (3) is the demediated outcome, calculated using equations (6) and (7) in the Appendix. Standard errors are clustered at the school level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Figure A.1: Timeline of study

Baseline data collection	Grade 1 teacher training	Grade 1 teacher training	Midline data collection	Grade 2 teacher training	Grade 2 teacher training	Classroom observations	Endline data collection
February	March	July	November	February	July	September	November

Grade 1 teacher coaching (March–Nov)

Grade 2 teacher coaching (Feb-Nov)

2015

2016
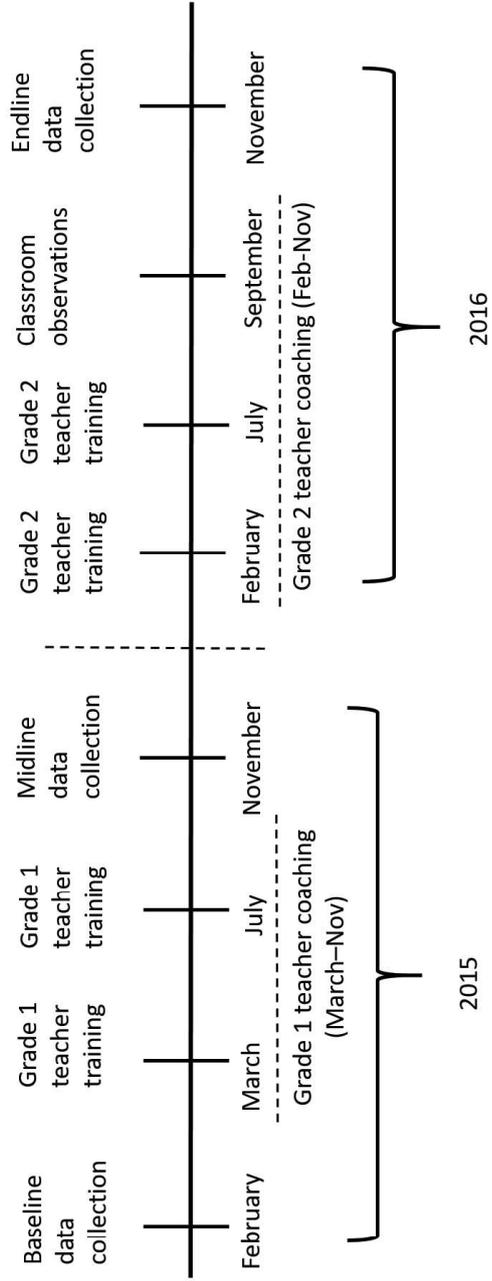
Figure A.2: Attrition and repetition rates across treatment arms
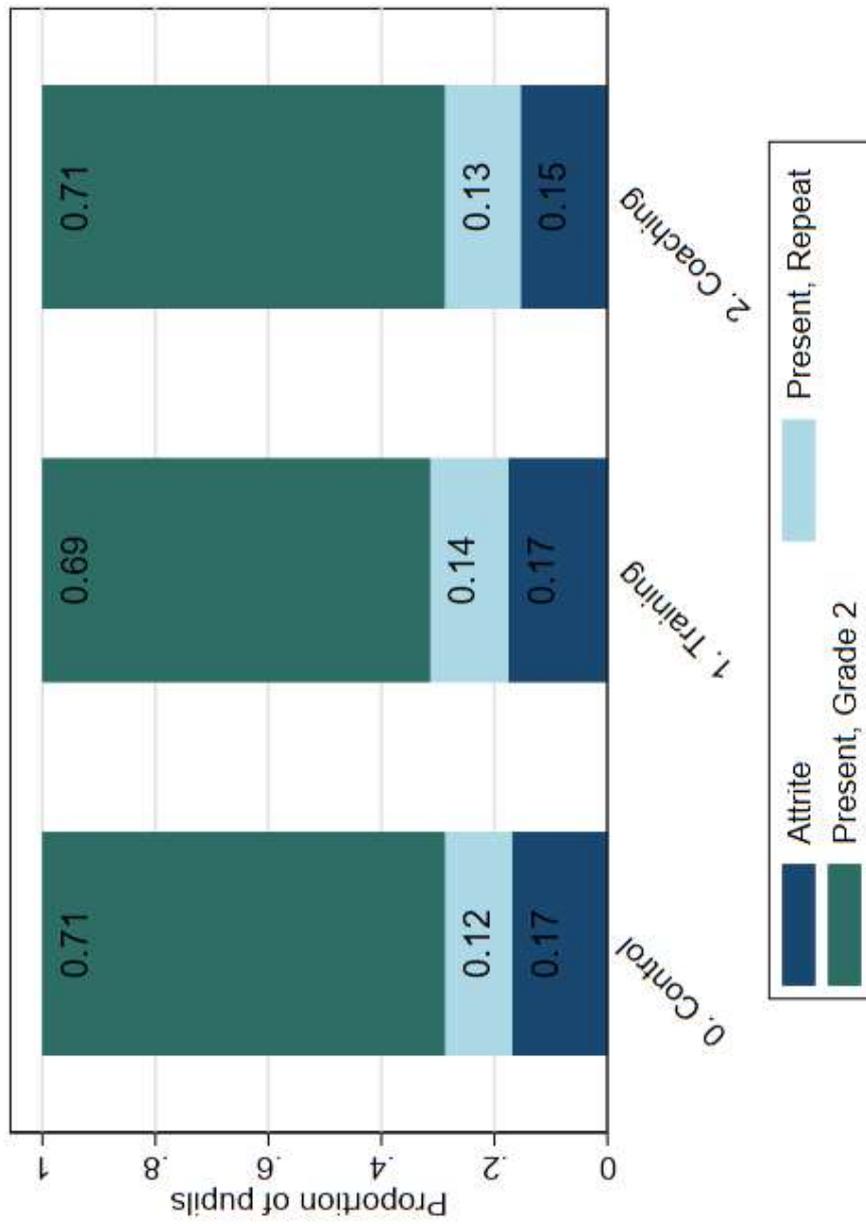
Note: The figure shows the proportion of surveyed pupils by treatment group who: (i) were not present at endline for the reading assessment; (ii) were present, but are repeating grade one; (iii) were present and are in grade two

Figure A.3: Baseline distribution of reading proficiency scores
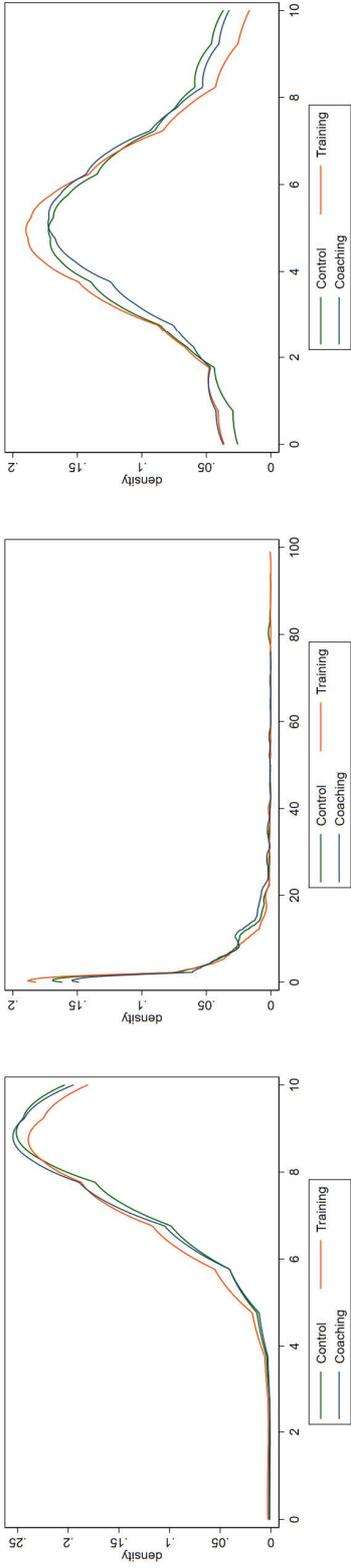
(a) Picture comprehension test

(b) Letter recognition

(c) Digit span test
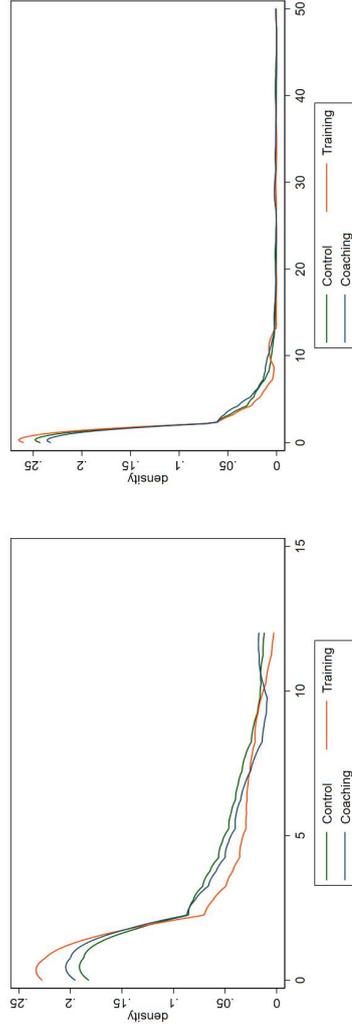
(d) Phonological awareness

(e) Words correct
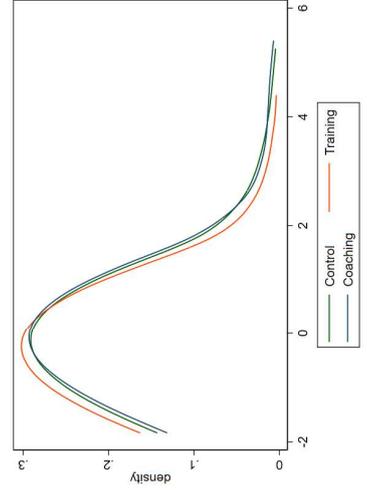
(f) Mean index

Figure A.4: Post-intervention distribution of aggregate reading proficiency
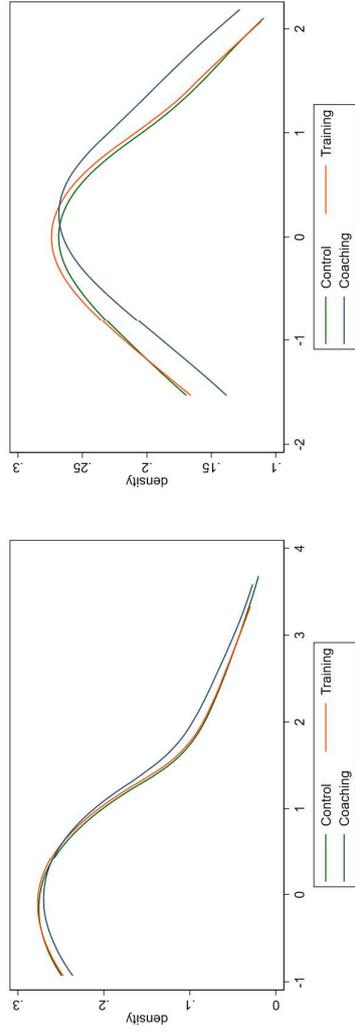
(a) Midline

(b) Endline

Appendix B   Mediation analysis

Appendix B.1   Linear Structural Equations Model

The Linear Structural Equations Model (LSEM) compares the regression result from equations 1 and 3 with a regression that includes both the treatment dummies and the mediator, M_{cs}:

$$y_{icsb1} = \beta_0 + \beta_1(\text{Training})_s + \beta_2(\text{Coaching})_s + \beta_3 M_{cs} + X'_{isb0}\Gamma + \rho_b + \varepsilon_{icsb1}, \tag{5}$$

Under some strong assumptions, the reduction in the estimated treatment impacts, $\hat{\beta}_1$ and $\hat{\beta}_2$, between equations 1 and 5 can be interpreted as the mediation effect.

Appendix B.2   Sequential g estimation

The sequential g estimation strategy, as proposed by Acharya et al. (2016), is considered an improvement to the above since it allows one to control for all potential post-treatment confounders. (Intuitively, the mediating variable of interest, M_{cs}, could be correlated with another post-treatment variable, Z_{cs}, that is correlated with both y_{icsb1} and treatment. Not including this variable would lead to a biased estimate of the contribution of M_{cs}.)

This estimation strategy consists of three steps. The first step regresses the outcome indicator on the mediation variables of interest, treatment dummies, pre-treatment confounders, and post-treatment confounders:

$$\begin{aligned} y_{icsb1} = {} & \beta_0 + \beta_1(\text{Training})_s + \beta_2(\text{Coaching})_s + \beta_3 M_{cs} + \beta_4(\text{Training} \times M)_{cs} \\ & + \beta_5(\text{Coaching} \times M)_{cs} + \mathbf{Z_{cs}}'\Delta + X'_{isb0}\Gamma + \rho_b + \varepsilon_{icsb1}, \end{aligned} \tag{6}$$

where $\mathbf{Z_{cs}}$ is a vector of potential post-treatment confounders. We follow the recommendation of Acharya et al. (2016) and also interact the mediator with treatment. Column (2) in Table A.7 shows the result for this regression, where the mediator is our index for group-guided reading, we also include the mean indices for curriculum coverage, routine, and print-richness in the classroom as other potential post-treatment confounders.

Figure A.5: Comparing the distribution of pupil performance between lesson observation sample and the remaining sample
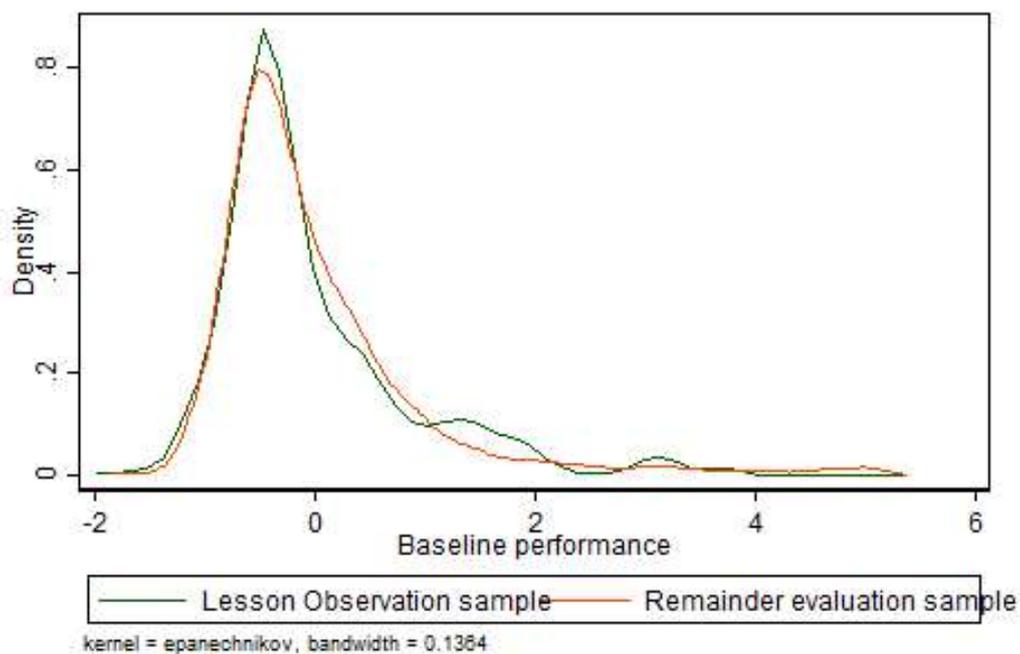

(a) Baseline pupil performance
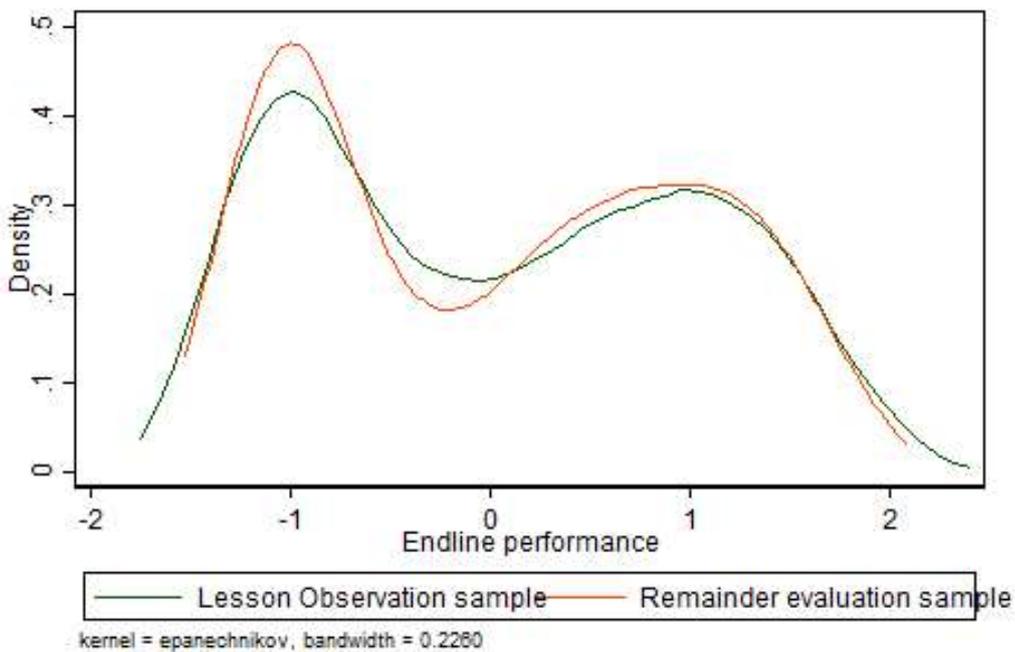
(b) Endline pupil performance

Figure A.6: Heterogeneous treatment impacts
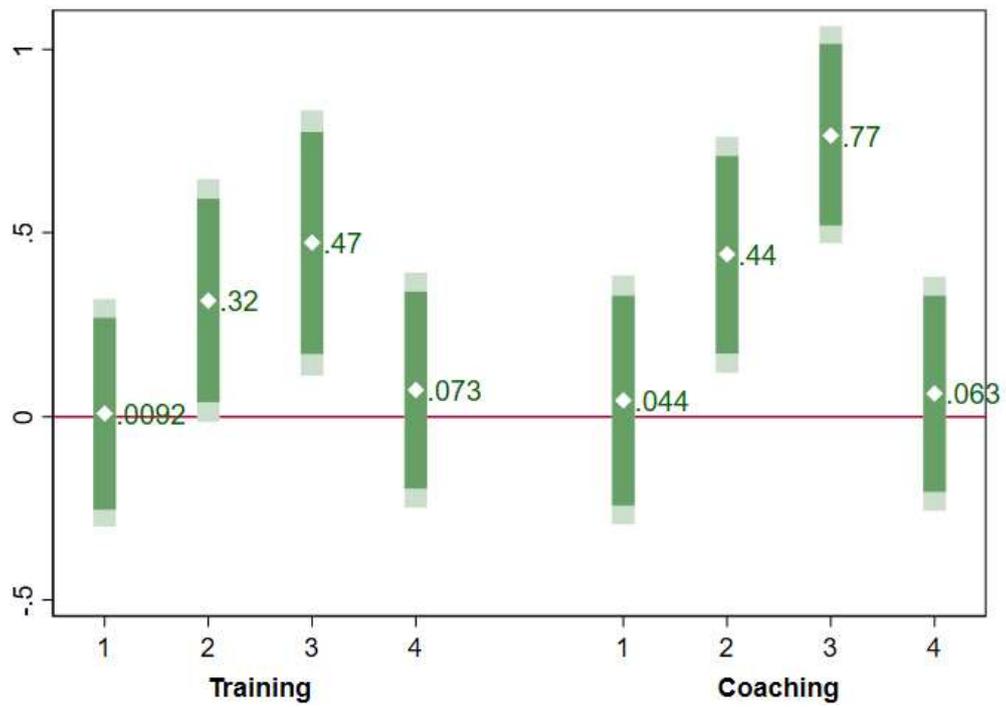
(a) By quartiles of class size
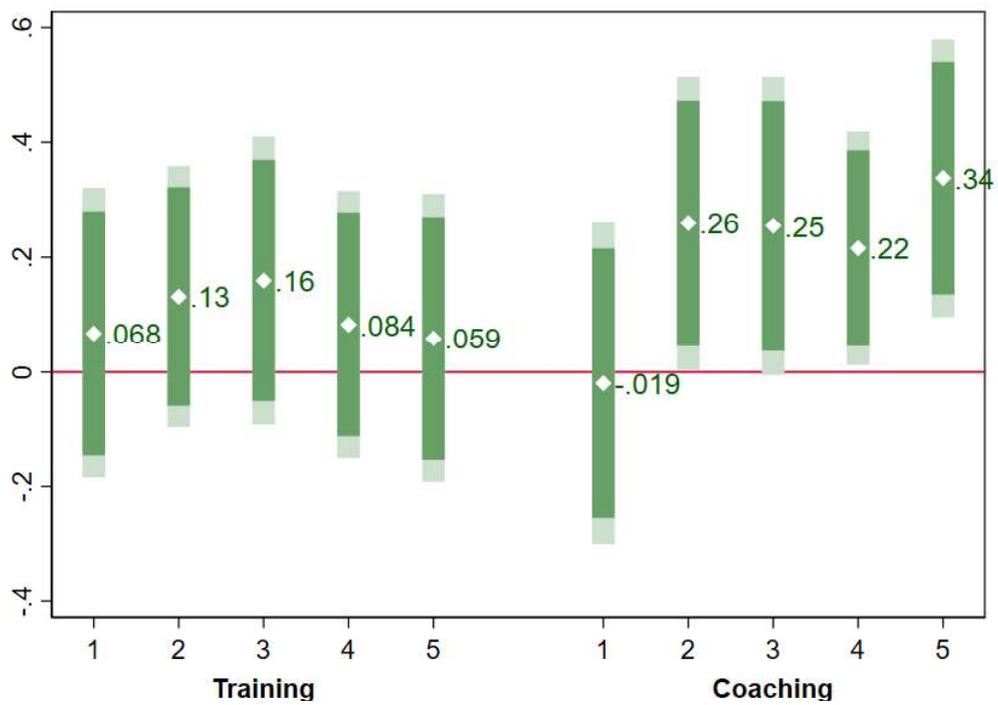
(b) By quintiles of baseline pupil performance

In the second step we demediate the outcome:

$$\tilde{y}_{icsb1} = y_{icsb1} - \hat{\beta}_3 M_{cs} - \hat{\beta}_4 (\text{Training} \times M)_{cs} - \hat{\beta}_5 (\text{Coaching} \times M)_{cs}. \tag{7}$$

Finally, we re-run our main regression on the demediated outcome, including the pre-treatment confounders (i.e. the initial set of controls from equation 1):

$$\tilde{y}_{icsb1} = \beta_0 + \beta_1 (\text{Training})_s + \beta_2 (\text{Coaching})_s + X'_{isb0} \Gamma + \rho_b + \varepsilon_{icsb1}. \tag{8}$$

The treatment impacts from equation 7 can be interpreted as the Average Controlled Direct Effect (ACDE): It is what the treatment impact would have been, if the value of the mediating variable was set to zero (in our case, this is the same as setting the mediating variable equal to the mean in the control). The difference between $\hat{\beta}_1$ and $\hat{\beta}_1$ in equations 1 and 8 can therefore be interpreted as the indirect impact of the treatment through the mediator— i.e. the contribution of the mediator to the overall treatment impact.

Column (3) in Table A.7 shows the regression results from equation 8. As a comparison, column (1) shows the regression results from 1, when restricting the sample to the same set of observations as in equations 7 and 8. The reduction in treatment impact is much larger when using the sequential g estimator, compared to the LSEM. The reason for this is the large positive interaction between Coaching and group-guided reading.