Online Appendix

Incentivizing School Attendance in the Presence of Parent-Child Information Frictions

Damien de Walque and Christine Valente

A. Appendix Tables

Table A-1: Sample Comparison with Representative Samples

		AIS 2015	
	Study Sample	Manica	Mozambique
Female-headed Household	0.20	0.43	0.40
Education of Head:			
No Education	0.15	0.28	0.24
Primary Education	0.56	0.46	0.54
Secondary Education or Above	0.29	0.26	0.22
Ownership of Household Assets:			
Bicycle	0.49	0.36	0.31
Motorbike	0.18	0.08	0.08
Radio	0.66	0.43	0.47
TV	0.34	0.32	0.34
Cell phone	0.80	0.59	0.64
Car	0.05	0.03	0.06
Type of Toilet Facility:			
None	0.09	0.32	0.24
Unimproved Latrine	0.56	0.43	0.37
Improved Latrine or Toilet	0.35	0.25	0.40
Number of Observations	2950	619	7169

Source: baseline household survey (de Walque and Valente, 2022) and MISAU, INE and ICF (2016). The number of observations refers to the size of the largest sample for which the variables are non-missing.

Table A-2: Descriptive Statistics and Balance at Baseline

	(1) Control Mean	(2) Information Mean	(3) Parent Cash Mean	(4) Child Incentive Mean
Household Head:				
Female	0.19	0.19	0.19	0.17
No Education	0.18	0.15	0.13	0.14
Primary Education	0.57	0.57	0.61	0.58
Secondary or Higher Education	0.26	0.28	0.25	0.27
Agriculture	0.53	0.48	0.55	0.50
White Collar	0.14	0.13	0.13	0.11
Other Occupation	0.33	0.39	0.31	0.39
Household wealth[1]:				
Lowest Tercile	0.42	0.36	0.37	0.37
Middle Tercile	0.32	0.34	0.30	0.35
Highest Tercile	0.26	0.30	0.33	0.28
Language:				
Portuguese	0.10	0.07	0.10	0.09
Ndau	0.21	0.21	0.26	0.28
Shona	0.11	0.13	0.13	0.14
Chiute	0.28	0.21	0.24*	0.20**
Chibarue	0.12	0.14	0.12	0.13
Other Language	0.18	0.24**	0.14	0.16
Religion:				
Catholic	0.12	0.07	0.11	0.12
Protestant	0.20	0.22	0.19	0.25*
Christian	0.16	0.21*	0.15	0.18
Zioni	0.20	0.21	0.28*	0.17
Atheist	0.15	0.12	0.10	0.14*
Other Religion	0.18	0.17	0.17	0.13
Girl Characteristics:				
Age	12.70	12.61	12.55	12.73
Consumption of Personal Goods[2]	967.08	887.45	998.58	937.30
High Empowerment[3]	0.40	0.42	0.34	0.42
Enrolled in 2015	0.97	0.98	0.98	0.96
Ever Married	0.02	0.01	0.02	0.02

Monitoring:				
Parent-Reported Absences[4]	1.12	0.93	0.76**	0.66***
High Monitoring Quality[5]	0.86	0.88	0.90*	0.88
Thinks a Weekly Attendance Report Card Would be Useful	0.84	0.82	0.81	0.80
N (Schools)	44	41	44	44

Source: baseline household survey (de Walque and Valente, 2022). *, ** and *** denote p-values significant at 10, 5 and 1% respectively obtained by estimating Equation (4). [1]Based on a principal component analysis score using information on ownership of household items and housing characteristics. [2] Value, in meticais, of non-food items purchased by any household member over the 12 months preceding the baseline survey and personally consumed by girls who, if they were to enroll in 2016, would enroll in Grades 6 or 7. [3]Share of girls with an above-median predicted score based on a principal component analysis of answers to questions about whether the girl would be able to keep some item of clothing given to her in exchange of work done, and whether she is involved in decisions concerning her healthcare, visiting relatives, attending school, and working outside the house. [4] Number of days absent from school during October 2015, if enrolled, as reported by the parent/guardian. [5]Share of girls with an above-median predicted score based on a principal component analysis of parent/guardian answers to three questions: whether they fully/partly agree that, at the end of each day, they know whether their daughter/ward was (i) at school, (ii) in the classroom; and whether it has ever happened that one day, they thought the girl was at school but actually she was not.

Table A-3: Effect on Additional Self-Reported Schooling Outcomes

Panel A: Girls' Own Outcomes				
	(1)	(2)	(3)	(4)
	Years of education	Primary education completed	Years of education	Primary education completed
Information	-0.0625	-0.0550	-0.0338	-0.0342
	(0.0589)	(0.0464)	(0.0615)	(0.0480)
Parent Cash	0.0474	0.00831	0.0425	0.000716
	(0.0580)	(0.0457)	(0.0607)	(0.0474)
Girl Voucher	0.0284	-0.00744	0.0610	0.0155
	(0.0580)	(0.0457)	(0.0611)	(0.0477)
Observations	173	173	173	173
Mean Y	6.32	0.44	6.32	0.44
SD Y	0.301	0.224	0.301	0.224
p info=parents	0.065	0.177	0.224	0.475
p info=girls	0.127	0.310	0.128	0.305
p girls=parents	0.742	0.729	0.759	0.754
Panel B: Girls' Siblings Self-Reported Enrollment				
	(1)	(2)	(3)	(4)
	Sisters	Brothers	Sisters	Brothers
Information	0.112	0.0994	0.0933	0.0767
	(0.0473)	(0.0387)	(0.0501)	(0.0390)
Parent Cash	0.0778	0.0555	0.0882	0.0272
	(0.0466)	(0.0381)	(0.0497)	(0.0387)
Girl Voucher	0.0932	0.0377	0.104	0.0202
	(0.0466)	(0.0381)	(0.0498)	(0.0387)
Observations	171	172	171	172
Mean Y	0.65	0.71	0.65	0.71
SD Y	0.211	0.197	0.211	0.197
p info=parents	0.470	0.258	0.920	0.211
p info=girls	0.693	0.113	0.828	0.150
p girls=parents	0.737	0.637	0.744	0.855
Baseline Characteristics	No	No	Yes	Yes

Source: de Walque and Valente (2022) household survey (endline for outcomes, and baseline for controls). Panel A: Dependent variables are highest completed grade in columns (1) and (3) and an indicator for completing primary education defined as having completed at least 7 grades in columns (2) and (4). Panel B: Dependent variables are enrollment of siblings of the eligible girls (sisters in columns 1 and 3; brothers in columns 2 and 4) as reported by parents. Baseline characteristics are the school sample averages for the following variables: self-reported (by parents) number of missed school days in October 2015 among girls enrolled,

binary indicator for high self-reported monitoring quality, five language indicators and five religion indicators. All regressions include a constant and district fixed effects. Standard errors in parentheses. “p arm_i=arm_j” denotes the p-value of a test of equal treatment effects between treatment arm i and treatment arm j. Standard errors in parentheses.

Table A-4a: Individual and Joint Tests of Treatment Effects Based on Randomization Inference

Table	Baseline Char.?	Outcome	Randomization p-values					
			(1)	(2)	(3)	(4)	(5)	(6)
			Info	Parents	Girls	Joint (equation)	Joint (table)	Joint (all 3*14=42 treatment effects)
Table 2	No	Share present at spot check	0.055	0.01	0.000	0.003		
	No	Share self-reported enrollment	0.659	0.069	0.827	0.165		
	No	Average ASER score	0.051	0.826	0.028	0.05		
	Yes	Share present at spot check	0.05	0.017	0.000	0.005		
	Yes	Share self-reported enrollment	0.762	0.204	0.643	0.356		
	Yes	Average ASER score	0.038	0.977	0.051	0.044	0.034	
Table 3	No	Class teacher presence rate	0.219	0.321	0.748	0.592		
	No	Share ever married	0.083	0.328	0.688	0.358		
	No	Share high self-reported monitoring quality	0.736	0.208	0.788	0.633		
	No	Share high self-reported empowerment	0.623	0.959	0.356	0.734		
	Yes	Class teacher presence rate	0.114	0.413	0.529	0.49		
	Yes	Share ever married	0.238	0.361	0.942	0.556		
	Yes	Share high self-reported monitoring quality	0.969	0.232	0.872	0.501		
	Yes	Share high self-reported empowerment	0.593	0.96	0.385	0.795	0.459	0.072

Source: de Walque and Valente (2022). Authors' calculations using Alwyn Young's randcmd program with 2000 randomization iterations. Randomization-t p-values in columns (1), (2), (3) and (4). Randomization-c p-values in columns (5) and (6). Baseline characteristics are the school sample averages for the following variables: self-reported (by parents) number of missed school days in October 2015 among girls enrolled, binary indicator for high self-reported monitoring quality, five language indicators and five religion indicators.

Table A-4b: Individual and Joint Tests of Treatment Effect Differences Based on Randomization Inference

Table	Baseline Char.?	Outcome	(1) Info =Parents	(2) Info =Girls	(3) Joint (equation)	(4) Joint (table) [no ASER eq.]	(5) Joint (both tables) [no ASER eq.]
Table 2	No	Share present at spot check	0.52	0.089	0.236		
	No	Share self-reported enrollment	0.167	0.499	0.105		
	No	Average ASER score	0.081	0.817	0.088		
	Yes	Share present at spot check	0.689	0.137	0.3		
	Yes	Share self-reported enrollment	0.364	0.452	0.216	0.103	
	Yes	Average ASER score	0.034	0.852	0.058	[0.215]	
Table 3	No	Class teacher presence rate	0.809	0.375	0.653		
	No	Share ever married	0.455	0.201	0.442		
	No	Share high self-reported monitoring quality	0.397	0.92	0.562		
	No	Share high self-reported empowerment	0.553	0.697	0.598		
	Yes	Class teacher presence rate	0.481	0.37	0.64		
	Yes	Share ever married	0.79	0.289	0.533		
	Yes	Share high self-reported monitoring quality	0.221	0.923	0.319		
	Yes	Share high self-reported empowerment	0.632	0.759	0.708	0.236	0.053 [0.191]

Source: de Walque and Valente (2022). Authors' calculations using Alwyn Young's randcmd program with 2000 randomization iterations. Randomization-t p-values in columns (1), (2), and (3). Randomization-c p-values in columns (4) and (5). Baseline characteristics are the school sample averages for the following variables: self-reported (by parents) number of missed school days in October 2015 among girls enrolled, binary indicator for high self-reported monitoring quality, five language indicators and five religion indicators.

Table A-5: Inverse Probability Weighting Attrition Correction

	(1)	(2)	(3)	(4)	(5)
	Average ASER math score	Share self-reported enrollment	Share ever married	Share high self-reported monitoring quality	Share high self-reported empowerment
Panel A: No controls for baseline characteristics					
Information	0.171	0.0000403	-0.00792	-0.0146	-0.0273
	(0.0958)	(0.0161)	(0.0105)	(0.0291)	(0.0418)
Parent Cash	0.0306	0.0319	-0.00221	0.0270	-0.00339
	(0.0943)	(0.0158)	(0.0104)	(0.0287)	(0.0412)
Child Incentive	0.191	-0.00797	0.0115	-0.00970	-0.0430
	(0.0943)	(0.0158)	(0.0104)	(0.0287)	(0.0412)
Panel B: Controlling for baseline characteristics					
Information	0.183	0.000310	-0.00424	-0.0297	-0.0276
	(0.0980)	(0.0171)	(0.0112)	(0.0304)	(0.0441)
Parent Cash	0.0182	0.0256	0.00179	0.0235	-0.00567
	(0.0963)	(0.0169)	(0.0110)	(0.0299)	(0.0432)
Child Incentive	0.168	-0.00792	0.0156	-0.0238	-0.0411
	(0.0955)	(0.0168)	(0.0110)	(0.0297)	(0.0430)
Panel C: Attrition					
Attrition rate in control group	.13	.072	.072	.072	.16
P-value of differences between arms	.488	.153	.153	.153	.512
Observations	173	173	173	173	173

Source: Household survey (de Walque and Valente, 2022). School averages and shares obtained after weighting each observation by the inverse of its predicted probability of being observed at endline as a function of all baseline characteristics listed in Table A-2. Regressions in Panel B also include school sample averages for the following baseline characteristics: self-reported (by parents) number of missed school days in October 2015 among girls enrolled, binary indicator for high self-reported monitoring quality, five language indicators and five religion indicators. All regressions include a constant and district fixed effects. The attrition rate varies across dependent variables due to non-response at the math test and empowerment questions. The p-values reported in the last row correspond to an F-test of joint significance of the treatment variables in a regression of the school-level attrition rate on the three treatment indicators and district fixed effects. Standard errors in parentheses.

Table A-6: ANCOVA Estimates

	(1) Share present at spot check	(2) Share self-reported enrollment	(3) Share ever married	(4) Share high self-reported monitoring quality	(5) Share high self-reported empowerment
Information	0.0431	0.00204	-0.00623	0.0121	-0.0198
	(0.0226)	(0.0144)	(0.00419)	(0.0260)	(0.0384)
Parent Cash	0.0559	0.0231	-0.000547	0.0357	-0.000994
	(0.0225)	(0.0142)	(0.00412)	(0.0258)	(0.0380)
Child Incentive	0.0778	-0.00160	-0.000183	0.00860	-0.0341
	(0.0227)	(0.0141)	(0.00410)	(0.0255)	(0.0379)
Parent-reported missed school days at baseline	-0.0101				
	(0.00994)				
Baseline outcome		0.420			
		(0.105)			
Baseline outcome			1.073		
			(0.0389)		
Baseline outcome				-0.0848	
				(0.0869)	
Baseline outcome					-0.0552
					(0.0769)
Observations	173	173	173	173	173
p info=parents	0.576	0.145	0.177	0.365	0.630
p info=girls	0.131	0.802	0.151	0.893	0.712
p girls=parents	0.321	0.082	0.929	0.287	0.386

Source: de Walque and Valente (2022). Household survey, except for the outcome variable in the first column, which comes from the attendance spot checks data. Any pupil listed on the class roll and not present in the class at the time of the attendance check is coded as absent. Parent-reported missed school days at baseline is the school average number of days parents said their daughter was absent from school during October 2015 (if enrolled in 2015). All regressions include a constant and district fixed effects. Standard errors in parentheses. "p arm_i=arm_j" denotes the p-value of a test of equal treatment effects between treatment arm i and treatment arm j.

Table A-7: Individual-Level Estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	=1 if present at spot check	=1 if self-reported enrollment	ASER score	=1 if Class teacher present	=1 if Ever married	=1 if High self-reported monitoring quality	=1 if High self-reported empowerment
Information	0.0418	0.00611	0.166	0.0210	-0.0169	-0.00664	-0.00861
	(0.0211)	(0.0143)	(0.0806)	(0.0215)	(0.00786)	(0.0196)	(0.0306)
Parent Cash	0.0513	0.0209	0.00358	0.0160	-0.0152	0.0134	-0.000725
	(0.0191)	(0.0122)	(0.0842)	(0.0190)	(0.00787)	(0.0184)	(0.0306)
Child Incentive	0.0597	-0.0110	0.210	0.0133	-0.0128	-0.0178	-0.0352
	(0.0162)	(0.0146)	(0.0730)	(0.0184)	(0.00789)	(0.0207)	(0.0322)
Observations	94746	2793	2600	96501	2793	2793	2520
No. of Clusters	173	173	173	173	173	173	173
Mean Y	0.68	0.95	2.19	0.92	0.03	0.91	0.28
SD Y	0.468	0.216	1.083	0.273	0.169	0.292	0.447
p info=parents	0.666	0.191	0.064	0.817	0.798	0.268	0.776
p info=girls	0.352	0.234	0.556	0.693	0.532	0.587	0.370
p girls=parents	0.619	0.010	0.010	0.873	0.709	0.109	0.247

Source: de Walque and Valente (2022). Outcome variables for Columns (1) and (4): unannounced spot checks attendance data. Any pupil listed on the class roll and not present in the class at the time of the attendance check is coded as absent. All other outcome variables: household survey (endline). The unit of observation in Columns (1) and (4) corresponds to one girl observed during one spot check. The unit of observation in all other columns corresponds to one girl interviewed during the endline household survey. All regressions include a constant and district fixed effects. School-level clustered standard errors in parentheses. "p arm_i=arm_j" denotes the p-value of a test of equal treatment effects between treatment arm i and treatment arm j.

Table A-8: Effect on Attendance, Sample Restricted to Girls Registered at First Spot Check

	(1) Share present at attendance check	(2) Share present at attendance check
Information	0.0419	0.0455
	(0.0228)	(0.0243)
Parent Cash	0.0604	0.0592
	(0.0225)	(0.0240)
Child Incentive	0.0810	0.0823
	(0.0225)	(0.0242)
Baseline Characteristics	No	Yes
Observations	173	173
Mean Y	0.65	0.65
SD Y	0.128	0.128
p info=parents	0.421	0.581
p info=girls	0.090	0.135
p girls=parents	0.359	0.333

Sources: de Walque and Valente (2022). Dependent variable: attendance spot checks, sample restricted to girls with an exact name match in the class roll used in the first spot check of the year (which took place between 02/25/16 and 03/31/16). Any pupil listed on the class roll and not present in the class at the time of the attendance check is coded as absent. Baseline characteristics: household survey. Baseline characteristics are the school sample averages for the following variables: self-reported (by parents) number of missed school days in October 2015 among girls enrolled, binary indicator for high self-reported monitoring quality, five language indicators and five religion indicators. All regressions include a constant and district fixed effects. Standard errors in parentheses. “p arm_i=arm_j” denotes the p-value of a test of equal treatment effects between treatment arm i and treatment arm j.

Table A-9: Sample Trimmed of the 5% Smallest and 5% Largest School Samples

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Share present at spot check	Share self-reported enrollment	Average ASER score	Class teacher presence Rate	Share ever married	Share high self-reported monitoring quality	Share high self-reported empowerment
Information	0.0498	0.00634	0.252	0.0443	-0.0153	0.0332	-0.0437
	(0.0235)	(0.0156)	(0.0925)	(0.0280)	(0.0103)	(0.0273)	(0.0405)
Parent Cash	0.0694	0.0254	0.0377	0.0261	-0.0100	0.0354	-0.0131
	(0.0229)	(0.0152)	(0.0899)	(0.0272)	(0.0100)	(0.0265)	(0.0393)
Child Incentive	0.0860	-0.00812	0.250	0.00798	-0.000850	0.00921	-0.0363
	(0.0227)	(0.0150)	(0.0891)	(0.0270)	(0.00996)	(0.0263)	(0.0390)
Observations	157	157	157	157	157	157	157
Mean Y	0.64	0.95	2.14	0.90	0.02	0.89	0.30
SD Y	0.128	0.084	0.564	0.155	0.045	0.169	0.231
p info=parents	0.420	0.238	0.026	0.527	0.620	0.937	0.463
p info=girls	0.135	0.366	0.984	0.206	0.173	0.391	0.859
p girls=parents	0.476	0.031	0.021	0.512	0.369	0.330	0.561

Source: de Walque and Valente (2022). Outcome variables for Columns (1) and (4): unannounced spot checks attendance data. Any pupil listed on the class roll and not present in the class at the time of the attendance check is coded as absent. All other outcome variables: household survey (endline). School size defined by the number of EP2 girls recorded as enrolled as of the first attendance spot check at the school. All regressions include a constant and district fixed effects. Standard errors in parentheses. “p arm_i=arm_j” denotes the p-value of a test of equal treatment effects between treatment arm i and treatment arm j.

Table A-10: Excluding Data Affected by Conflict

	(1) Share present at spot check	(4) Class teacher presence rate
Information	0.0379	0.0300
	(0.0220)	(0.0256)
Parent Cash	0.0546	0.0332
	(0.0216)	(0.0252)
Child Incentive	0.0718	0.0117
	(0.0216)	(0.0252)
Observations	173	173
Mean Y	0.65	0.91
	0.128	
SD Y		0.161
p info=parents	0.450	0.901
p info=girls	0.127	0.479
p girls=parents	0.425	0.394

Source: Unannounced spot checks attendance data (de Walque and Valente, 2022). Any pupil listed on the class roll and not present in the class at the time of the attendance check is coded as absent. School averages obtained after dropping from the database the three spot check rounds for which attendance data could be collected for less than 70% of the district's schools. All regressions include a constant and district fixed effects. Standard errors in parentheses. "p arm_i=arm_j" denotes the p-value of a test of equal treatment effects between treatment arm i and treatment arm j.

Table A-11: Selecting Covariates Based on Their Predictive Power

	(1) Share present at spot check	(2) Share self-reported enrollment	(3) Average ASER score	(4) Class teacher presence rate	(5) Share ever married	(6) Share high self-reported monitoring quality	(7) Share high self-reported empowerment
Information	0.0395	-0.00799	0.202	0.0344	-0.00568	0.0101	-0.0147
	(0.0227)	(0.0143)	(0.0898)	(0.0272)	(0.00430)	(0.0264)	(0.0405)
Parent Cash	0.0647	0.0140	0.0127	0.0186	-0.00263	0.0463	0.00416
	(0.0229)	(0.0144)	(0.0906)	(0.0275)	(0.00434)	(0.0266)	(0.0408)
Child Incentive	0.0790	-0.00989	0.195	0.00302	-0.000157	0.00417	-0.0299
	(0.0227)	(0.0143)	(0.0899)	(0.0273)	(0.00431)	(0.0264)	(0.0406)
Observations	173	173	173	173	173	173	173
Mean Y	0.65	0.95	2.16	0.90	0.03	0.89	0.30
SD Y	0.128	0.087	0.567	0.154	0.048	0.167	0.228
p info=parents	0.285	0.139	0.043	0.576	0.494	0.186	0.653
p info=girls	0.089	0.896	0.940	0.259	0.209	0.826	0.712
p girls=parents	0.535	0.101	0.046	0.573	0.571	0.117	0.408

Source: de Walque and Valente (2022). Outcome variables for Columns (1) and (4): unannounced spot checks attendance data. Any pupil listed on the class roll and not present in the class at the time of the attendance check is coded as absent. All other outcome variables: household survey (endline). All regressions include a constant, district fixed effects, and any baseline variable which has a t-statistic equal to 1.96 or above when regressing the outcome on district fixed effects and the baseline characteristics reported in Table A-2 in the control group. Standard errors in parentheses. “p arm_i=arm_j” denotes the p-value of a test of equal treatment effects between treatment arm i and treatment arm j.

Table A-12: Ex-Post Power Calculations

Outcome	Mean control group	SD control group	MDE	MDE as % of the Mean
Share present at spot check	0.65	0.128	0.078	12%
Share self-reported enrollment	0.95	0.0870	0.053	6%
Average ASER score	2.16	0.567	0.343	16%
Average teacher presence	0.9	0.153	0.092	10%
Share ever married	0.03	0.0476	0.029	96%
Share reporting high monitoring quality	0.89	0.1677	0.101	11%
Share reporting high empowerment	0.3	0.228	0.138	46%

Power calculations for a probability of type I error of 0.05 and a control and treatment group of 44 schools each (which apply to comparisons between any two of the parent cash, child incentive, and control groups). Calculations applying to comparisons between the information treatment arm (41 schools) and any of the other experimental arms have slightly larger MDEs, but differences only appear at the third decimal and are therefore omitted for conciseness.

Table A-13: Attendance Effects and Timing of Spot Checks

	(1)	(2)	(3)	(4)
	Early in Term	Late in Term	Not Harvest Time	Harvest Time
Information	0.0485	0.0536	0.0481	0.0532
	(0.0271)	(0.0227)	(0.0256)	(0.0250)
Parent Cash	0.0626	0.0607	0.0630	0.0521
	(0.0267)	(0.0222)	(0.0252)	(0.0246)
Girl Voucher	0.0898	0.0734	0.0913	0.0573
	(0.0267)	(0.0224)	(0.0252)	(0.0246)
Observations	172	170	173	168
Mean Y	0.64	0.65	0.61	0.78
SD Y	0.151	0.142	0.146	0.155
p info=parents	0.603	0.754	0.562	0.965
p info=girls	0.129	0.386	0.094	0.872
p girls=parents	0.305	0.566	0.260	0.832

Source: de Walque and Valente (2022). Dependent variable: attendance spot checks: any pupil listed on the class roll and not present in the class at the time of the attendance check is coded as absent. All regressions include a constant and district fixed effects. Standard errors in parentheses. “p arm_i=arm_j” denotes the p-value of a test of equal treatment effects between treatment arm i and treatment arm j.

Table A-14: Effect of Treatments on Eligible Girls' Consumption of Personal Items

	Dependent Variable: Consumption of Personal Items Not Purchased With Vouchers (meticais)	
	(1) All observations	(2) Top 1% removed
Information	19.55	47.52
	(72.94)	(64.96)
Parent Cash	-50.08	-41.66
	(71.83)	(63.97)
Child Incentive	-68.40	-89.18
	(71.83)	(63.97)
Constant and District FE	Yes	Yes
Observations	173	173
Mean Y	831.69	783.72
SD Y	517.92	462.06
p info=parents	0.344	0.174
p info=girls	0.232	0.038
p girls=parents	0.798	0.456

Source: Endline household survey (de Walque and Valente, 2022). The dependent variable is the total value of purchases, over the 12 months preceding the survey, of the following items: trousers/skirts, shirt/t-shirt/jumper, school uniform, other ready-made garments, made-to-measure clothing, clothing repairs, shoes, sandals, trainers, other types of shoes, shoe repairs, matches, soap (detergent), soap (personal hygiene), toothpaste, teeth cleaning twig, perfume, deodorant, backpack, travel bag/handbag, batteries, magazines/newspapers, any other good for personal use (e.g., hair extensions, etc…). Standard errors clustered at the school level in parentheses. "p arm_i=arm_j" denotes the p-value of a test of equal treatment effects between treatment arm i and treatment arm j.

Table A-15: Effect on Attendance by Selected School Population Characteristics

	(1) Share present at spot check		(2) Share present at spot check		(3) Share present at spot check
Information	0.0461	Information	0.0418	Information	0.0554
	(0.0239)		(0.0241)		(0.0224)
Information × Poorest	-0.0121	Information × Oldest	0.0247	Information × Furthest	0.0369
	(0.0434)		(0.0446)		(0.0538)
Parent Cash	0.0496	Parent Cash	0.0637	Parent Cash	0.0480
	(0.0301)		(0.0264)		(0.0231)
Parent Cash × Poorest	0.0435	Parent Cash × Oldest	0.00589	Parent Cash × Furthest	0.0547
	(0.0456)		(0.0527)		(0.0615)
Child Incentive	0.0966	Child Incentive	0.0766	Child Incentive	0.0961
	(0.0289)		(0.0311)		(0.0263)
Child Incentive × Poorest	-0.0408	Child Incentive × Oldest	0.0541	Child Incentive × Furthest	0.0140
	(0.0540)		(0.0564)		(0.0640)
Poorest	0.0320	Oldest	-0.0793	Furthest	-0.0184
	(0.0631)		(0.0670)		(0.102)
District FE	Yes		Yes		Yes
Interactions District FE and Poorest or Oldest or Furthest	Yes		Yes		Yes
Observations	173		173		173
P-value 3 interactions=0	0.355		0.758		0.780

Source: de Walque and Valente (2022). Unannounced spot checks attendance data (for outcome variable) and household survey (variables interacted with the treatment indicators). “Poorest”, “Oldest” and “Farthest” are indicator variables equal to one if the school’s share of girls surveyed at baseline that are classified as “poor”, “old”, and “far from school”, respectively, is in the top tercile of the school distribution. “Poor” refers to girls in the lowest household wealth tercile, “old” refers to girls in the highest individual tercile for age (14 and above at baseline) and “far from school” refers to girls in the highest individual tercile for time taken to travel to school (33 minutes and above). Standard errors in parentheses.

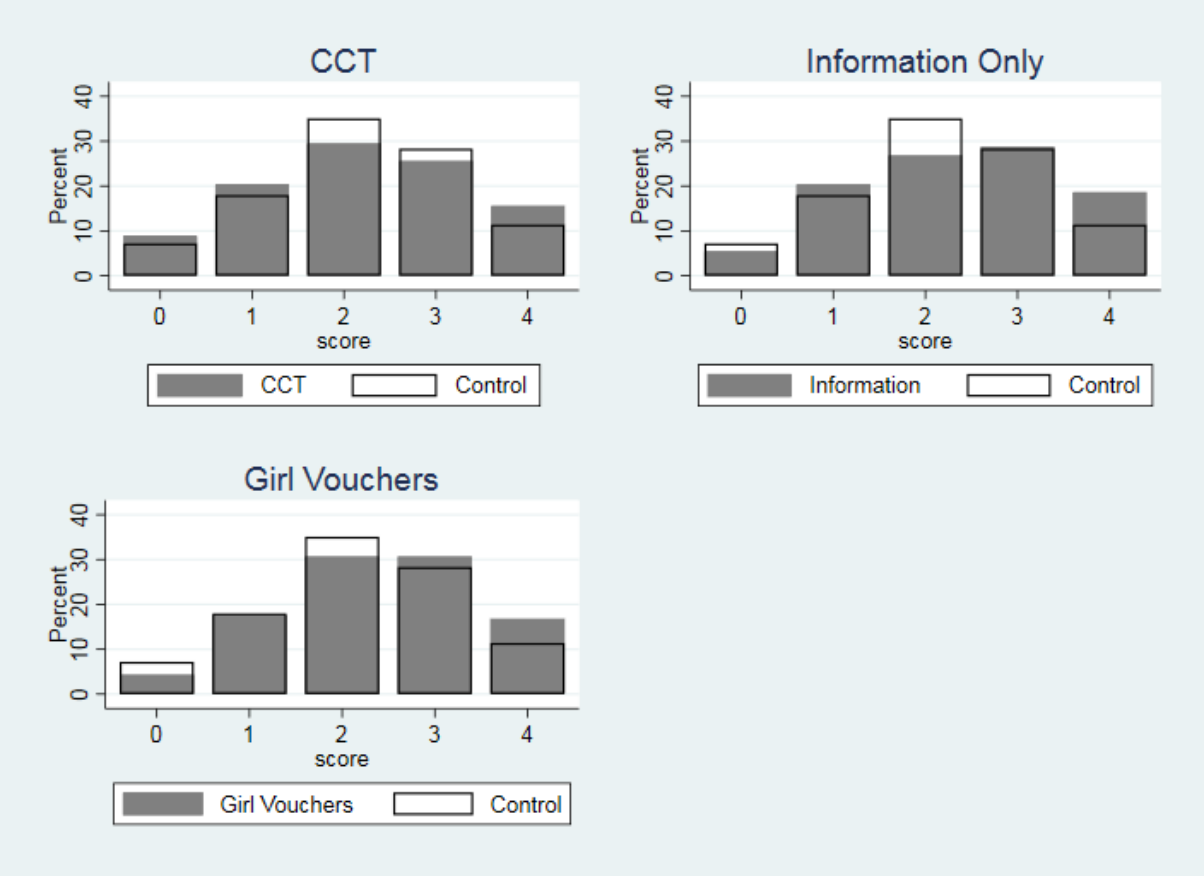
Table A-16: Learning Effects Outside the Mean

Outcome:	(1) score	(2) I[score=4]	(3) I[score>2]	(4) I[score<2]	(5) score	(6) I[score=4]	(7) I[score>2]	(8) I[score<2]
Panel A: School averages								
Information	0.183	0.0699	0.0804	-0.00342	0.195	0.0782	0.0829	-0.00342
	(0.0911)	(0.0281)	(0.0384)	(0.0326)	(0.0909)	(0.0284)	(0.0377)	(0.0326)
Parent Cash	0.0202	0.0471	0.0228	0.0341	-0.00233	0.0427	0.0208	0.0341
	(0.0898)	(0.0277)	(0.0378)	(0.0321)	(0.0897)	(0.0280)	(0.0372)	(0.0321)
Girl Voucher	0.203	0.0338	0.0799	-0.0468	0.178	0.0256	0.0693	-0.0468
	(0.0898)	(0.0277)	(0.0378)	(0.0321)	(0.0904)	(0.0282)	(0.0374)	(0.0321)
Characteristics	No	No	No	No	Yes	Yes	Yes	No
Observations	173	173	173	173	173	173	173	173
Panel B: Individual outcomes:								
Information	0.166	0.0740	0.0776	0.00242	0.171	0.0835	0.0829	0.0105
	(0.0806)	(0.0259)	(0.0371)	(0.0304)	(0.0823)	(0.0274)	(0.0379)	(0.0307)
Parent Cash	0.00358	0.0435	0.0190	0.0413	-0.00673	0.0449	0.0173	0.0494
	(0.0842)	(0.0219)	(0.0359)	(0.0323)	(0.0830)	(0.0217)	(0.0360)	(0.0318)
Girl Voucher	0.210	0.0571	0.0859	-0.0359	0.226	0.0647	0.0966	-0.0356
	(0.0730)	(0.0208)	(0.0325)	(0.0286)	(0.0693)	(0.0206)	(0.0314)	(0.0280)
Baseline Characteristics	No	No	No	No	Yes	Yes	Yes	Yes
Observations	2600	2600	2600	2600	2464	2464	2464	2464
No. of clusters	173	173	173	173	173	173	173	173
Mean individual outcome	2.19	0.11	0.40	0.25	2.20	0.11	0.40	0.25

Source: Endline household survey (de Walque and Valente, 2022). I [] denotes the indicator function. “Score” is the ASER math test score, which takes the following possible values: recognition of single digit numbers (scored 1), recognition of double-digit numbers (scored 2), correct subtraction (scored 3), correct division with remainders (scored 4), or cannot even recognize single digit numbers (scored 0). All regressions include a constant and district fixed effects. Panel A: standard errors in parentheses. Panel B: school-level clustered standard errors in parentheses.

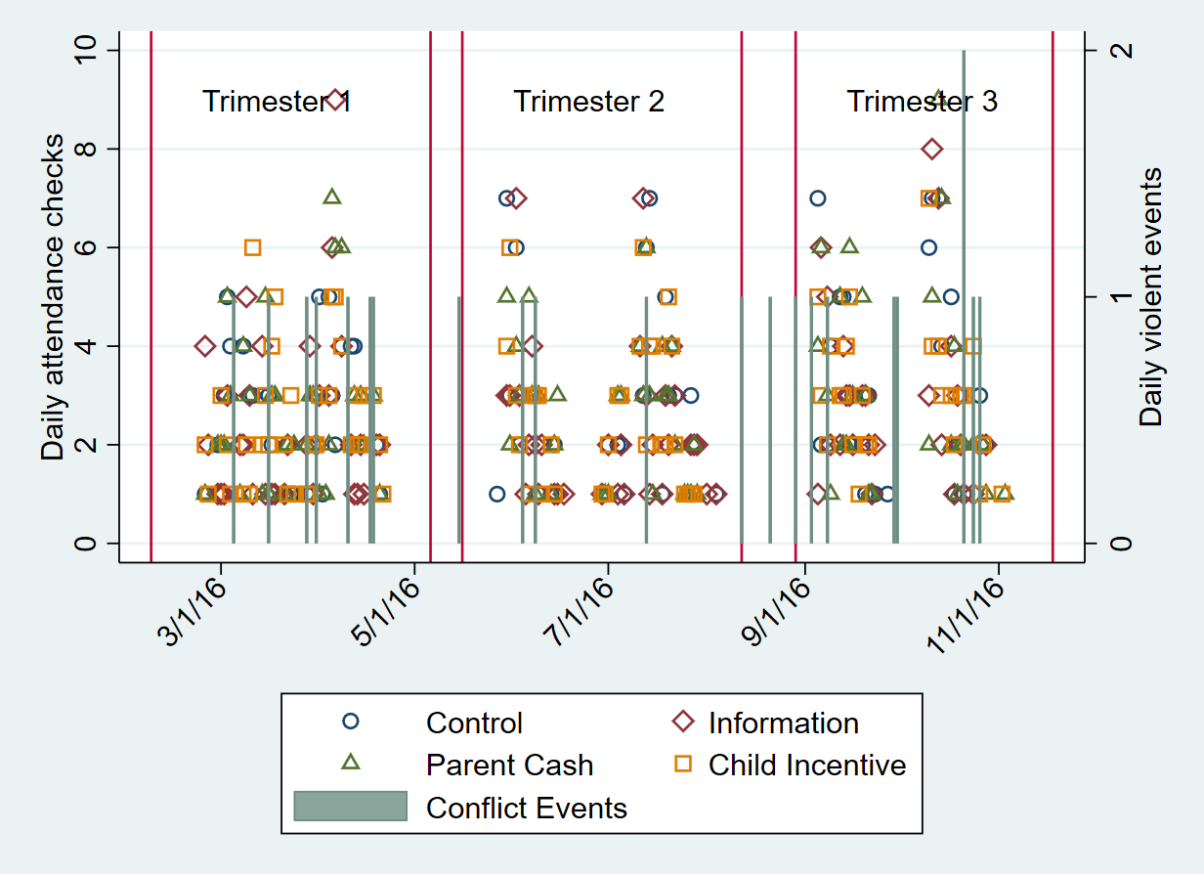
B. Appendix Figures

Figure A-1: Effect of the Treatments on the Distribution of Math Scores



Source: Endline household survey (de Walque and Valente, 2022). "Score" is the ASER math test score, which takes the following possible values: recognition of single digit numbers (scored 1), recognition of double-digit numbers (scored 2), correct subtraction (scored 3), correct division with remainders (scored 4), or cannot even recognize single digit numbers (scored 0).

Figure A-2: Timing of Attendance Spot Checks and Conflict Events



Source: Attendance spot checks (de Walque and Valente, 2022) and ACLED (2021)

C. Further Related Literature

C1. Role of Information

For conciseness, in the main text we focus on the literature interested specifically in the information asymmetry between parents and their children in the area of education. Gallego, Malamud and Pop-Eleches (2018) have documented evidence of information asymmetry between parents and children regarding internet usage, and a rich body of work has shown evidence of misinformation relevant to educational choices that goes beyond parent-children asymmetric information (e.g., Nguyen, 2008; Jensen, 2010; Bettinger et al., 2012; Hoxby and Turner, 2013; Dinkelman and Martínez, 2014; Wiswall and Zafar, 2015; Andrabi, Das and Khwaja, 2017; Dizon-Ross, 2019).

C2. Children Incentives

For experiments incentivizing students, but not parents, for the student to achieve a certain standard at scholastic tests or a range of inputs in this test, see Angrist and Lavy (2009), Kremer, Miguel and Thornton (2009), Jackson (2010), Fryer (2011), Bettinger (2012), Levitt et al. (2016), Burgess, Metcalfe and Sadoff (2016), Hirshleifer (2017). For experimental evaluations of the effect of distributing free school uniforms without conditionality, see Hidalgo et al. (2013), Duflo, Dupas and Kremer (2015) and Evans and Ngatia (2020).

D. Qualitative Evidence

Focus groups interviews took place in May 2014 to obtain a qualitative understanding of the relevance, acceptability, and feasibility of the proposed study in the Manica context. The focus groups consisted of five groups of girls age 11-15 (mostly 6^{th}-7^{th} graders for those currently enrolled in school) and their parents or guardians. The girls and their parents/guardians were interviewed separately to avoid girls' answers being influenced by the presence of their parents or guardians, yielding 10 focus groups in total. One of the authors was present at all the interviews, which were carried out in the local language by a member of our partner NGO, Magariro, with experience in carrying this type of interviews. The interviews were recorded, with the consent of the participants, and then translated to permit analysis by the PI.[1]

There were between 3 and 10 participants per group. The recruitment aimed to purposely include girls who attend school regularly, girls who have dropped out, and girls who miss school a lot. Two sites were chosen to cover a remote rural setting and a setting that is closer to a main town. Both schools practice double shifts (one in the morning and one in the afternoon), with shifts of about 3 hours and a half each.

[1] Ethical clearance was obtained from the University of Bristol.

The same main reasons for missing school days conditional on enrolment were cited by parents and girls, namely: illness, lack of soap to wash themselves or their clothes, lack of decent clothes (which makes some girls ashamed or sad when they go to school and see others with nicer things), some children not liking school or simply preferring doing something else like playing or hanging out with their boyfriends (from 13/14 years old according to respondents). Two parents also said that sometimes the road was not passable due to rain. Only one parent in all the groups said that sometimes they asked their daughter to stay home to help them with some work on the family plot or to look after their siblings when their mother has to travel.

It was also clear that both parents *and* children had influence on the decision to go to school. It was interesting, for instance, to note that several parents contrasted different daughters of theirs, with one sibling going to school regularly without any problem, and another dragging their feet, arriving late at school, making excuses not to go, or simply skipping schools (e.g., seemingly going to school but then turning around and going back home while their parents are out for work). Here are a few illustrative quotes, from both settings:

From a mother of two girls, one age 14 (in 6th grade) and one age 11 (in 4th grade): on her older daughter: *"Ana, she's 14 and this one likes school, she knows how to read and write, she even reads the bible for me and explains it to me. Sometimes she doesn't go to school but it's not her fault but rather because of me asking her to stay home and help with some chores at home or on our plot when there is a lot to do and sometimes if I want to travel I ask her to look after her siblings at home and she gets cross because she likes to study. Her sister Laura doesn't know how to write, she only learnt how to write her name this year, she is lazy and doesn't like to go to school. She is 11 and is in fourth grade."*

From another mother of two girls, one age 14 and one age 12: on her younger daughter: *"Yes, sometimes this one she misses school. She says she's ill, she has no soap to wash her clothes, etc… They are very different. When food is late, Luisa [the older sister], she goes to school all the same, whereas Maria, she waits and sometimes misses school because of that. Luisa only misses school if she's actually ill."*

From yet another mother of two girls, on her older daughter: *"My older daughter, Gabriela, she doesn't like going to school, she misses school a lot. I want her to go but she doesn't like it. You tell her: it's time to go, and she says: I'll go tomorrow. She finds excuses. Lucia, no, she likes going to school. Even if lunch is late, she goes to school anyway [i.e., skipping lunch]."*

And from a father of two girls, on his younger daughter: *"Veronica doesn't like school, but her older sister likes it. Sometimes she says: my clothes are dirty, or my clothes are still wet. Sometimes I'm not even sure she doesn't turn around and goes back home instead of going to school. Sometimes she says she's feeling unwell, but I'm dubious."*

All in all, the overall message seems to be that when parents want children to go to school and children do not want to, many parents' testimonies seem to imply that they had little influence on their children and that in the end children did what they wanted. Of course, some

parents seemed more "in charge". For instance, when asked whether his daughter attended school regularly, the father of one 13-year-old responded that *"Yes, she goes to school regularly, I don't accept her not going to school"*. Several of the better educated parents said they checked their children's notebook ("caderno diario") to check what they had been up to (E.g., *"I check Maria's [notebook] more because I know she doesn't like school. She also tends to be late, so I check on her a lot."*). And there may be other ways for illiterate parents to exert some monitoring: *"I have no way to control them on their way [to school]. I don't know how to read, I can't check their notebook or anything. But I hear her when she's talking to her brothers, what she's talking about"*. However, the monitoring technologies available to parents are certainly imperfect - when asked directly about whether they know what their children do on the way to and from school, two parents replied: *"On the way to and from school, we can't control them" "We don't know."*.

When girls were asked whether they attended school regularly, the vast majority said that they did. However, when asked about the attendance of other girls in general, several girl participants report that there are girls who skip school a lot *("It varies from person to person, there are girls who go every day but also others who miss school two or three days a week."*), or who go to school but do not actually enter the classroom. In what highlights the limitations of the "monitoring technologies" at the disposition of parents, one girl told us that *"many stay in the street chatting to boys and then go and write in their notebook, faking corrections and go back home and show it to their parents when they never even reached the school."*.

E. Pre-Specified Outcomes

The following main and secondary outcomes were registered on the AEA registry in February 2016. Main outcomes: school attendance conditional on enrollment, unconditional attendance, and school enrollment. Secondary outcomes: teacher absenteeism, score at ASER math test and RAVEN test, marital status, self-reported quality of monitoring of daughter's school attendance, and intra-household bargaining power. There was no further pre-analysis plan other than pre-specifying these outcomes. Here we report estimates for all the outcomes which we were able to measure satisfactorily. The two exceptions are: (i) RAVEN test, which ended up not being fielded in the endline questionnaire because pre-tests of the endline questionnaire suggested it was too long and (ii) unconditional attendance. We intended to construct this measure of unconditional attendance by setting attendance to 1 if a girl from the household survey was observed in any of our spot check class rolls and present at a check, and zero if she was matched but absent or if she could not be matched to any spot check record. If, despite being announced after the official school enrollment period, the treatments had had an impact on enrollment, this outcome variable would have allowed us to estimate the effect of the treatments on attendance independently of any selection into school enrollment, albeit on the much smaller household survey sample rather than on the universe of EP2 girls.

While, conditional on being reported by her parent as being enrolled in the endline household survey, the probability of finding a match in one of our 173 school records of 2016 enrollees is high (80%), this probability varies significantly across treatment arms. When estimating Equation (4) on the sample of girls who are reported as being enrolled in 2016 in the household survey, and defining Y_c as the share of girls with a match in our 2016 class rolls, the coefficients associated with the information only arm is -0.05, that associated with the parents cash arm is 0.02, and that associated with the child incentive arm is 0.008. In contrast, the largest absolute effect of our treatments on the share of girls self-reported as enrolled in Table 2 is 0.027, and this effect is shown not to be robust. Since evidence supports the conclusion that our treatments had no robust effect on enrollment or on school switches, while we are unequally successful across experimental arms in matching names of self-reported enrollees from the household survey with those found in school records, analyzing the effect of the treatments on unconditional attendance would be a bad cure for a non-existent ailment.

F. Detailed Description of Further Robustness Checks

Correcting for attrition for outcomes measured through the household survey. As reported in the main text, attrition of girls taking part in the household survey was slightly larger in the control group than in the treated groups, although not jointly statistically significantly so. Our main outcome of interest (independently verified attendance rate at school) and teacher attendance are not affected by any differences in attrition in the household survey. In Table A-5, we present results for the other outcomes, correcting for differences in survey attrition. More precisely, we ran regressions in which the school averages are obtained after weighting each individual observation in the endline survey sample by the inverse of the probability that it is included in the sample, as predicted by all the individual and household baseline characteristics summarized in Table A-2. Reassuringly, reweighting observations by the inverse of the probability that they attrit does not change our conclusions.

Controlling for pre-treatment outcomes. As an additional robustness check, we also present ANCOVA estimates obtained from estimating Equation (4) with an additional regressor capturing- or proxying for baseline outcomes. When the outcome was measured at baseline, we control for the baseline value of the outcome variable. Alternatively, we control for an available proxy of the outcome at baseline when a reasonable proxy exists. When the baseline outcome is available, a commonly used approach is to use a Difference-in-Differences specification. Using an ANCOVA approach is preferable to Difference-in-Differences even when the baseline outcome is available, as there is no loss of power when the correlation between pre- and post-treatment outcomes is low (McKenzie, 2012). Results in Table A-6 show that all our conclusions are robust to the inclusion of these pre-treatment outcomes.

Sample-weighted estimates. The main analysis reported in this paper is carried out at the school level (i.e., averaging variables at the school level) without applying any sampling

weights, so that each school is weighted equally. We repeated the analysis at the individual level (clustering the standard errors at the school level), and thus weighting each school by the size of the school sample. For outcomes measured at spot checks, this essentially implies weighting each school by the size of its female EP2 intake. For outcomes based on the household survey, the sampling target was to interview the same number of observations per school (20), which would have led to the same weighting as in the cluster-level analysis. In practice, there was some variation in the household-survey sample size across schools—but *not* across experimental arms—due to difficulties locating the girls listed in the school records, as discussed in Section II-A. It is therefore less clear how the weighting in these individual-level estimates should be interpreted. Results—shown in Table A-7—are however largely unchanged.

No selection of girls through school switches. The treatments were announced after the official enrollment period closed, and, in most cases, after the start of the school year, so that a negligible effect on enrollment was to be expected, as confirmed in our data analysis. Another potential source of selection of girls into the school registers for which the survey firm recorded spot check attendance data is through school switches. Out of the 2,687 endline survey girls who were reported by their parents as being enrolled for the 2016 school year, only 157 (5.84%) were reported as being enrolled in a school other than the one they were sampled from at baseline. Estimating Equation (4) using, as dependent variable, a binary indicator equal to one if the girl is reported enrolled in a different school to that from which she was sampled and zero if she was reported enrolled in her original school, no treatment indicator is individually significant (nor are they jointly significant).[2] As a further robustness check, we re-estimated the effect of our treatments on attendance, but restricting the sample used to construct the share of girls present to names registered on the class roll at the first spot check. The first spot checks were carried out within the two first months of school (between February 25 and March 31), and so well before any end-of-trimester transfers were paid. The class rolls called by the independent surveyor were slightly updated between spot checks for various reasons. A few girls changed classes or schools during the year, some names were updated to match the girl's used name when it did not match that with which she was recorded in the school register, or to match the name used at home in the case of girls included in the household survey sample. Estimates obtained by restricting the spot checks data to girls with exact name matches from the first attendance check roll are presented in Table A-8. These results are near-identical to those obtained in the main analysis, thus confirming that selection through school switches is unlikely to be biasing our results.

Trimming the school sample. The school-level analysis carried out in the paper is much less sensitive to outliers in terms of school size than individual-level analysis (since each school is given the same weight). Still, in Table A-9, we report results obtained when dropping the 5% largest schools and 5% smallest schools to test whether results are very different in the tails of

[2]Individual coefficients (p-values) are: 0.018 (0.355), -0.007 (0.713), 0.017 (0.355) for the information, parent cash and child incentive arms, respectively, and the joint F-test p-value is 0.457.

the school size distribution. Trimming the school sample in this way tends to increase slightly the magnitude of all the treatment effects without altering any of the conclusions based on the baseline results.

Excluding spot check data where conflict caused substantial disruptions to data collection. Low-level conflict between government and RENAMO forces slowed down but did not prevent household data collection at baseline and endline. School closures due to the conflict are balanced across experimental arms: the p-value of a joint test of no treatment effect in a regression of an indicator for the school having closed at any point in the school year due to the conflict on the three treatment dummies and district (strata) fixed effects is equal to 0.89. Figure A-2 plots the number of daily spot checks by experimental arm against daily violent events recorded in Manica province in the ACLED dataset (Raleigh et al., 2010). It shows that both violent events and attendance spot checks were spread out across the whole school year. Within a spot-check "round", which each lasted about one month, there is also no difference across treatment arms in the probability of the enumerator successfully collecting attendance data (p-value: 0.27). At the peak time of tension in the most affected district (Mossurize), however, many schools were closed so that attendance data collection could not proceed. The schools for which we were able to collect attendance data at those times may therefore be selected (although, as mentioned before, there was no overall difference between treatment arms in the number of times attendance data was collected). Table A-10 reports estimates for the two outcomes based on attendance checks obtained when ignoring data from spot checks for which less than 70% of the district's schools could be surveyed.[3] Point estimates decrease slightly in magnitude—suggesting the treatments may have had larger effects at times of high absenteeism due to the conflict, but the overall picture is unchanged.

Selecting control variables based on their predictive power instead of baseline balance. Throughout the paper, we check the robustness of our findings to controlling for characteristics that were not balanced for at least one treatment arm at baseline. In Table A-11, we instead control for covariates chosen based on their ability to predict the outcomes studied, irrespective of baseline balance. More specifically, in this robustness check we control for any baseline variable which has a t-statistic equal to 1.96 or above when regressing the outcome on district fixed effects and the baseline characteristics reported in Table A-2 in the control group. The effect of the CCT on high self-reported monitoring quality becomes statistically significant at 10% but our overall conclusions are unchanged.

Ex-post power calculations. In Table A-12, we report ex-post power calculations using the means and standard deviations of the outcomes studied in this paper in the control group, for 80% power in detecting differences between any experimental group pair and a Type 1 error of 0.05. In keeping with the main analysis, we present power calculations based on the

[3] Note that where the robustness checks apply to spot check data only, we only report results for outcomes based on spot check data – hence not test scores, which are collected in the endline household survey. For instance, in Table A-10, we exclude spot check data collected at times of substantial conflict disruption, but this does not affect outcomes based on the household survey since the conflict slowed down but did not prevent data collection at baseline or endline.

distribution of school-level averages.[4] The last column reports the Minimum Detectable Effect (MDE) as a share of the control group's mean, showing that the experiment is well-powered for our three schooling outcomes, teacher absenteeism and self-reported monitoring quality, but not for early marriage and self-reported empowerment. This bolsters our confidence in the results for which we find consistent significant effects, while confirming the inconclusiveness of our findings for early marriage and self-reported empowerment.

[4] The standard deviation in the school-average distribution of ASER scores (0.567) is much smaller than the standard deviation in the individual-level distribution (1.083). When computing power for an analysis carried out at the individual level, and taking the mean, standard deviation, and intraclass correlation in the control group as reference parameters, the MDE for 80% power for a 0.05 Type 1 error corresponds to 0.265 of a standard deviation.

Appendix-Only References

ACLED. 2021. "The Armed Conflict Location & Event Data Project". Conflict data for Mozambique [March 1, 2016 to November 3, 2016]. https://developer.acleddata.com/. Accessed July 8, 2021.

Andrabi, Tahir, Jishnu Das, and Asim I. Khwaja. 2017. "Report Cards: The impact of providing school and child test scores on educational markets." *American Economic Review* 107(6): 1535-63.

Angrist, Joshua, and Victor Lavy. 2009. "The effects of high stakes high school achievement awards: Evidence from a randomized trial." *American Economic Review* 99(4), 1384-1414.

Bettinger, Eric P. 2012. "Paying to learn: The effect of financial incentives on elementary school test scores." *Review of Economics and Statistics*, 94(3), 686-698.

Bettinger, Eric, Bridget Long, Philip Oreopoulos, and Lisa Sanbonmatsu. 2012. "The Role of Application Assistance and Information in College Decisions: Results from the H&R Block FAFSA experiment". *The Quarterly Journal of Economics* 127(3), 1205-1242.

Burgess, Simon, Robert Metcalfe, and Sally Sadoff. 2016. "Understanding the response to financial and non-financial incentives in education: Field experimental evidence using high-stakes assessments." (IZA Discussion Papers, No. 10284.

Dinkelman, Taryn and Claudia Martínez A. 2014. "Investing in Schooling in Chile: The role of information about financial aid for higher education." *The Review of Economics and Statistics* 96(2), 244-257.

Dizon-Ross, Rebecca. 2019. "Parents' Beliefs About Their Children's Academic Ability: Implications for Educational Investments." *American Economic Review*. 109(8): 2728-65.

Duflo, Esther, Pascaline Dupas, and Michael Kremer. 2015. "Education, HIV, and Early Fertility: Experimental evidence from Kenya." *The American Economic Review* 105(9), pp.2757-2797.

Evans, David. and Mũthoni Ngatia. 2020. "School Uniforms, Short-Run Participation, and Long-Run Outcomes: Evidence from Kenya." *World Bank Economic Review* 35(3): 705-719.

Fryer, Roland G. 2011. "Financial incentives and student achievement: Evidence from randomized trials." *The Quarterly Journal of Economics* 126(4), 1755-1798.

Hidalgo, Diana, Mercedes Onofa, Hessel Oosterbeek, and Juan Ponce. 2013. "Can Provision of Free School Uniforms Harm Attendance? Evidence from Ecuador." *Journal of Development Economics* 103:43-51.

Gallego, Francisco, Ofer Malamud. and Cristian Pop-Eleches. 2018. "Parental Monitoring and Children's Internet Use: The role of information, control, and cues." NBER Working Paper No. 23982.

Hoxby, Caroline and Sarah Turner. 2013. "Expanding College Opportunities for High-Achieving, Low-Income Students." Stanford Institute for Economic Policy Research Discussion Paper.

Jackson, C. Kirabo. 2010. "A Little Now for a Lot Later: A look at a Texas Advanced Placement Incentive Program." *Journal of Human Resources 45*(3), 591-639.

Jensen, Robert. 2010. “The (Perceived) Returns to Education and the Demand for Schooling.” *The Quarterly Journal of Economics* 125(2), 515-548.

Kremer, Michael, Edward Miguel, and Rebecca Thornton. 2009. “Incentives to Learn.” *Review of Economics and Statistics* 91 (1): 437- 456.

Levitt, Steven D., John A. List, Susanne Neckermann, and Sally Sadoff. 2016. “The behavioralist goes to school: Leveraging behavioral economics to improve educational performance.” *American Economic Journal: Economic Policy* 8(4), 183-219.

Ministério da Saúde (MISAU), Instituto Nacional de Estatística (INE), e ICF. 2015. *Inquérito de Indicadores de Imunização, Malária e HIV/SIDA em Moçambique 2015*. Household Recode (MZHR71FL). Demographic and Health Surveys Program. https://dhsprogram.com/data/dataset/Mozambique_Standard-AIS_2015.cfm?flag=0. Accessed April 6, 2020.

Nguyen, Trang. 2008. “Information, Role Models and Perceived Returns to Education: Experimental Evidence from Madagascar.” Unpublished manuscript.

Raleigh, Clionadh, Andrew Linke, Håvard Hegre and Joakim Karlsen. 2010. “Introducing ACLED-Armed Conflict Location and Event Data.” *Journal of Peace Research* 47(5) 651-660.

Wiswall, Matthew and Basit Zafar. 2015. “Determinants of College Major Choice: Identification using an information experiment.” *The Review of Economic Studies* 82(2), 791-824.