PROVING GROUND
TECHNICAL APPENDIX
MARCH 2020

Reduce Absences in Early Grades with Personalized Postcards

Analytic Approach and Model

SAMPLE

We evaluated the impact of personalized postcard interventions designed to reduce absenteeism among students in grades pre-K through grade 2. Following student absences, teachers in the intervention group sent parents a postcard with personalized information about their student's absences. Pilots were carried out by two of our district partners during the Fall and Spring of the 2018–2019 school year. Our analytic sample consisted of 5,602 students in classrooms or schools that participated in the pilot as part of either the control or treatment group. To be included in our analytic sample, students needed to be enrolled at the start and end of the intervention period.

RANDOMIZATION

In District A, we randomly assigned nine schools to the control group and nine schools to the treatment group. We randomized these schools, blocking on levels of the schools' average absence rates in the prior year. We constructed the blocks by evenly dividing schools into three groups based on their average absence rate and then randomly assigned half the schools in each block to treatment and the other half to control. School treatment assignments applied to both the Fall and Spring rounds of piloting. In District B, 54 classrooms were randomly assigned to the treatment and 53 were assigned to the control group. We randomized classrooms, blocking on special education classroom status, school, and grade. Table A1 provides a summary of the randomization.

Table A1. Randomization Summary

	Number of Units Randomized	Randomization Ratios (# of units)		Blocking (# of blocks)
		Controls	Postcard Intervention	
District A	18 schools	50% (9)	50% (9)	Absence groups (3)
District B	107 classrooms	50% (53)	50% (54)	Special Ed, school, and grade (41)
Combined	125 units	50% (62)	50% (63)	

IMPLEMENTATION

Although the theory of action was consistent across districts, each district tailored implementation to match their own context. For example, in District A the teachers filled out the postcards at the end of each week and the district mailed them to students' homes. In District B, the central office provided blank postcards for teachers to complete. At the end of each day, the teachers would fill out postcards for all the students who were absent that day. The next day that the student was present, the teacher would put the postcard(s) in the student's backpack to take home. Teachers tracked how many postcards they completed and which students received them, as well as the time they spent each day filling out the postcards. Table A2 outlines some of the key areas in which implementation differed across the two districts.

Table A2. Implementation Summary

	District A	District B
Delivery	Regular Mail	Student's backpack
Frequency	Weekly	Daily
Printing & Postage Cost	\$0.14/postcard for printing in house + \$0.35 for bulk mail postage	\$0.16/postcard for printing with contractor
Time Cost	Teachers spent about 15 minutes/week populating the postcard.	Teachers spent about 1 hour/week populating the postcards.
Design	Color-coded guide customized for each quarter.	Color-coded guide static for the entire year.

ESTIMATION

We estimate the impact of the intervention by comparing absences among groups that did and did not receive an intervention. We use the number of days a student was absent over a period of roughly 13 weeks as our outcome measure. We estimated the effect of the postcard intervention on student absences using hierarchical Bayesian methods (see Gelman et al., 2014, Chapter 5). Estimation proceeded in two steps:

In the first step, for each district an estimate of the effect of the treatment (postcard intervention) on absences and its standard error were obtained using maximum likelihood estimation of a Poisson model with a log link. The models used student-level data on days absent during the pilot period as the dependent variable, controlled for strata dummies, student-demographics, prior absence information, and included the days each student was enrolled in the district during the pilot period as an exposure variable (see Table A3 for specific controls included). Thus, absences were modeled as:

$$E(\#days\ absent) = (\#days\ enrolled) * e^{(treatment*\beta+X\gamma)}$$

Standard errors on the treatment effect (β) accounted for clustering at the level of randomization (school or classroom). Models were estimated using R.

In the second step, we obtained Bayesian estimates of the treatment effect in each district by modeling the district-level estimates of the treatment effect from the first stage with a hierarchical Bayesian model. Because the first stage estimates for each district ($\hat{\beta}_j$) were based on large samples, they were assumed to be normally distributed, where $\hat{\beta}_j \sim N(\beta_j, SE_j)$ and SE_j is the estimated standard error for $\hat{\beta}_j$ from the first stage. At the district level, we assume the treatment effects are distributed normally around a common mean, where $\beta_j \sim N(\beta_0, \tau)$. To estimate the model, we used a normal prior for β_0 with mean 0 and standard deviation of 0.04, reflecting our prior belief that effects of this type of light-touch intervention were likely to be small on average (e.g., <10%). We used a half-normal prior for τ^2, based on a normal with mean zero and variance chosen so that the square root of the mean of the half-normal was .03, reflecting our prior belief that variation in the treatment effect across districts was likely to be small. Models were estimated using the R package rstan.

Mean estimates and credible intervals are based on the posterior distribution of β_j. We report posterior means and credible intervals for effects transformed into percentage change: $100 * (exp(\beta_j) - 1)$. This is reported for each district, and for the average percentage change across districts. We calculated the certainty that the treatment reduced absences as the probability that the treatment effect was below 0 based on these posterior distributions. Finally, we calculate annualized estimates of days saved if the intervention were scaled up for a full year to all eligible students as: the estimated percentage change in absences times the total number of eligible students times 180 days (a full school year) times the absence rate of the control group during the study period.

Table A3. Covariates Included as Controls in the First Stage Poisson Model

Category	Control	Description
Demographics	Male	Indicator variable that takes the value of "1" if student is male.
	FRPL Eligibility	Indicator variable that takes the value of "1" if student is eligible to receive free or reduced-price lunch.
	English Language Learner (ELL)	Indicator variable that takes the value of "1" if student is an ELL.
	Special Education	Indicator variable that takes the value of "1" if student has a disability.
	Race	Categorical variable indicating whether student is Black, White, Hispanic, or other. Treated as separate indicator variables.
	Grade Level	Categorical variable indicating student's grade level. Treated as separate indicator variables.
Offset Variable	Days Enrolled	The number of the total days the student has been enrolled during the pilot period.

Table A3. Covariates Included as Controls in the First Stage Poisson Model (Continued)

Category	Control	Description
Prior School-Level Absence Variables	School Prior Average Days Absent	The log of the average number of days absent in the prior year for students at the student's school.
	Missing School Prior Average Days Absent	Indicator variable that takes the value of "1" if the student's school is missing prior average days absent data.
Prior Grade-Level Absence Variables	School-Grade Prior Average Days Absent	The log of the average number of days absent in the prior year for students in the student's current grade and school.
	Missing School-Grade Prior Average Days Absent	Indicator variable that takes the value of "1" if the student is missing school grade prior average days absent data.
Prior Student-Level Absence Variables	Prior Year Absence Rate	Student's absence rate for the previous academic year. Represented by 6 spline variables, with thresholds that are derived empirically from the sample's absence rate distribution. The linear spline has knots at the 20, 40, 60, 80, and 95th percentiles.
	Missing Prior Year Absence Rate	Indicator variable that takes the value of "1" if the student is missing prior absence rate data.
	Pre-Treatment Absence Rate	Student's absence rate for the current school year, prior to treatment. Represented by 6 spline variables, with thresholds that are derived empirically from the sample's absence rate distribution. The linear spline has knots at the 20, 40, 60, 80, and 95th percentiles.
	Missing Pre-Treatment Absence Rate	Indicator variable that takes the value of "1" if the student is missing pre-treatment absence rate data.

Results

At the start of the pilot period 5,891 students were enrolled in randomized classrooms and schools (3,181 in treatment, 2,710 in control). Our final sample size was 5,602 (see Table A4) after excluding 5% of students who were no longer enrolled in the district at the end of the analysis period and an additional five students who were missing key covariate information. Attrition from the sample was balanced across treatment and control groups: 144 (4.53%) and 145 (5.35%) respectively. Student characteristics were balanced between treatment and control groups in the final sample (see Table A5).

We estimate that the pilot intervention **reduced absences by 7.9%** compared to absences in the control group. The 50% credible interval (representing the 25th percentile to the 75th percentile of the posterior distribution) around our estimate is a 6.4% to 9.3% reduction in absences. We estimate that it is more than 99% likely that the intervention is better than the status quo (i.e., greater than 99% of the posterior distribution lies below zero, indicating lower absences due to the intervention). We used these estimates to calculate the magnitude of the impact in terms of the likely number of instructional days gained. Using the average number of absences in the control group over the 13-week period and then extrapolating to a full school year and scaling up to all eligible students, we estimate that the intervention could lead to a gain of about 6,883 instructional days. Table A6 further breaks out our estimates for each of our partner districts: 7.7% reduction in District A's pilot and 8.0% reduction in District B's pilot). To check the sensitivity our estimate to the prior selected, we estimated the impact using a more diffuse prior and found a mean reduction of 9.5% in absences across the districts.

Table A4. Analytic Sample

	Number of Schools or Classrooms in Analytic Sample	Number of Students in Analytic Sample	
		Controls	**Postcard Intervention**
District Pilot A	18 schools	1,590	2,006
District Pilot B	107 classrooms	975	1,031
Combined	125 units	2,565	3,037

Table A5. Balance Table

Baseline Characteristics	Control Mean	Treatment Mean	P-value
Male	0.53	0.52	0.40
ELL	0.09	0.16	0.05
Special Ed	0.16	0.16	0.65
FRPL	0.64	0.67	0.79
Prior Days Absent	19.02 (15.39)	18.12 (14.51)	0.20
School-Grade Prior Avg. Days Absent	17.76 (3.66)	16.94 (3.58)	0.26
School Prior Avg. Days Absent	16.81 (3.03)	16.04 (3.49)	0.49
Pre-treatment Days Absent	3.86 (3.92)	4 (4.41)	0.14

Table A5. Balance Table (Continued)

Baseline Characteristics	Control Mean	Treatment Mean	P-value
Black or African American	0.44	0.41	0.46
Hispanic	0.11	0.11	0.86
White	0.27	0.27	0.86
% Missing Prior Days Absent	0.5	0.54	0.13
% Missing School-Grade Prior Avg. Days Absent	0.22	0.24	0.45
% Missing School Prior Avg. Days Absent	0.03	0.02	0.52
% Missing Pre-treatment Days Absent	0.62	0.66	0.94
N	2565	3037	

Table A6. Estimated Impact and Instructional Days Saved

	# of Students in Analysis	Mean Absence Rate for Controls	Impact Estimate	Probability Reduces Absences	Annual Days Saved
District A	3,596	.0984	-7.7%	> .99	4,904
District B	2,006	.0685	-8.0%	> .99	1,979
Combined	5,602	.0870	-7.9%	> .99	6,883

References

Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., & Rubin, D.B.(2013). Hierarchical models (pp. 101-137). In *Bayesian data analysis* (3rd ed., Texts in statistical science). Boca Raton: CRC Press.