

Targeting Nudges to Students in the Pandemic: An Adaptive Experiment in Brazil

Pre-analysis Plan

A growing literature documents that nudges to students and their caregivers – from reminders to encouragement messages – can systematically improve educational outcomes. Having said that, when there is substantial heterogeneity in treatment effects, policy-makers have very little guidance on what is the best nudge for each student. In partnership with the São Paulo State Secretariat of Education (SEDUC-SP), in Brazil, this project implements an adaptive treatment protocol to assess heterogeneity in the responses to motivational nudges via text messages, and thereby to assign optimal messaging to students based on their observed characteristics. In a previous experiment (Bettinger et al., 2021) during school closures in the context of the pandemic, students were randomized to motivational nudges via text messages that targeted, specifically, (i) salience of school activities, or (ii) growth mindset, or (iii) beliefs about high returns to effort, or (iv) beliefs about low costs of effort, or (v) risk-taking, or (vi) future-orientation. We have built a machine learning model to predict student-level responses to each treatment arm – short-term effects on access to remote learning activities, and effects on attendance and grade by the end of the school year –, based on their individual characteristics. In this follow-up experiment, we use the ML-predicted ranking of treatments to assign each individual to their personally optimal messages. We evaluate impacts of being assigned the best treatment on average versus the personally optimal treatment on school attendance, grades and school dropouts.

I. Introduction

A growing literature documents that nudges to students and their caregivers – from reminders to encouragement messages – can systematically improve educational outcomes (Bergman, 2019). Having said that, when there is substantial heterogeneity in treatment effects, policy-makers have very little guidance on what is the best nudge for each student. This project implements an adaptive experiment to study the question of optimal targeting of nudges through the application of machine learning techniques.

For this study, we partnered with São Paulo State's Secretary of Education (SEDUC-SP), in Brazil. SEDUC-SP provided access to 1,415,290 mobile phone contacts either from students themselves (grades 10 to 12) or a students' parent or legal guardian (grades 6 to 9) to evaluate different nudges with the potential to improve educational outcomes -- and, in particular, to prevent student dropouts amidst the COVID-19 pandemic.

In this context, Bettinger et al. (2021) is evaluating the impacts of several versions of the growth mindset intervention (Yeager, 2019; Bettinger et al., 2018) via text messages to students and their caregivers. The experiment tries to decompose the original intervention into its underlying economic parameters in trying to single out what are the key drivers of its impacts on educational outcomes (if any). Concretely, the interventions are as follows:

- (i) **Growth mindset:** text messages that try to convey the content of the original intervention, that communicates students that their brain is 'like a muscle', and as such can 'become stronger' as a result of higher effort, that everyone can improve relative to themselves, and that success (failure) is not merely a matter of talent (lack of thereof).
- (ii) **Salience of school activities:** text messages with simple reminders from the school; placebo intervention whereby text messages try to make school activities more salient, without affecting beliefs, risk or time preferences.
- (iii) **High returns to effort:** text messages that emphasize that higher effort leads to better educational outcomes.
- (iv) **Low costs of effort:** text messages that emphasize that studying is not that hard and might even be fun.
- (v) **Risk-taking:** text messages that emphasize the value of taking risks.
- (vi) **Future-orientation:** text messages that emphasize the value of thinking about one's future.

Over the course of the 3rd quarter of 2020, those interventions were piloted at scale, whereby 800,000 students received a single text message, whose content was randomized to one out of the 6 groups above. The project continues into 2021, evaluating the impacts of the interventions delivered through two SMS per week over the course of three months over the course of the first semester.

In 2021, Movva, our implementing partner, produced 6 different sequences of motivational nudges via text messages to be sent over the course of the 1st and 2nd school quarters to 240,000 students. Each sequence is inspired in one of the interventions in Bettinger et al. (2021), although they differ from that study in that they vary in other dimensions (such as the activities they suggest) which could not be changed in the context of a controlled experiment.

In this follow-up experiment, we estimate heterogeneous treatment effects of the interventions piloted in 2020 to assign the nudges inspired on those interventions based on their predicted returns.

On the one hand, we have the average treatment effects of each of the 6 treatment arms of the 2020 pilot. On the other hand, we form the predicted conditional average treatment effects of these groups for any given individual, based on their observed pre-treatment characteristics.

Because short-term treatment effects on access to online remote learning activities and those at the end of the school year on attendance and grades are very different (both when it comes to average and conditional average treatment effects), we record those estimates separately, and experiment with optimal targeting based on different outcomes as well.

II. Research Questions

Our key research question is: **Does individualized targeting of growth mindset interventions improve on academic and behavioral learning outcomes more than an aggregated policy learning approach?** We hypothesize that the answer is yes, because the CATE's computed from the first-round experiment indicate that the optimal message varies significantly across individuals.

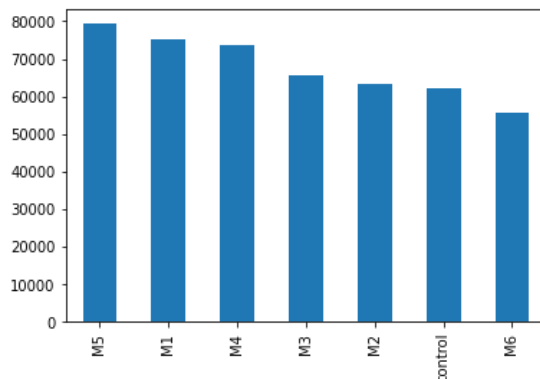
Additional questions that we will consider (some of which might or might not be able to answer depending on access to data in the future and ability to run additional experiments):

1. Does the answer to our key research question vary by grade?
2. Are the targeting gains persistent?
3. Do additional rounds of targeted interventions induce larger impacts?

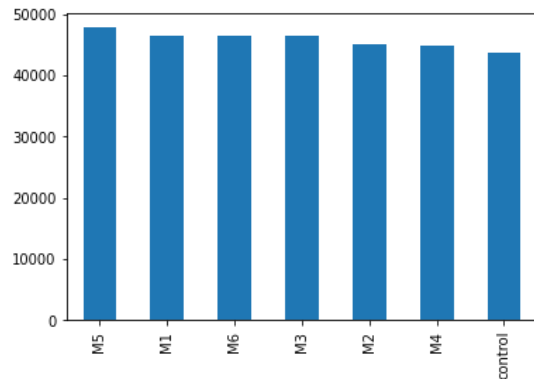
III. Learning Heterogeneous Effects.

We apply the T-Learner method to learn heterogeneous treatment effects (Kunzel et al 2019). From the pilot experiment, we have 6 datasets corresponding to each of the 6 SMS interventions. Within each dataset, we train a regression model to predict outcomes (see below) based on the individual's predetermined characteristics. We use gradient boosting with early stopping to encourage out-of-sample fit. We can then apply each of the 6 trained models to each datapoint, to produce 6 predicted counterfactual outcomes corresponding to each treatment group. We rank these outcomes and choose the message that is predicted to have the best outcome. The figures below showcase the distribution of highest conditional average treatment effects, by outcome on which predictions are based.

Panel A – Online access



Panel B – Summary measure of attendance and grades



IV. Experimental Design

The basic design of this new experiment comprises 5 groups:

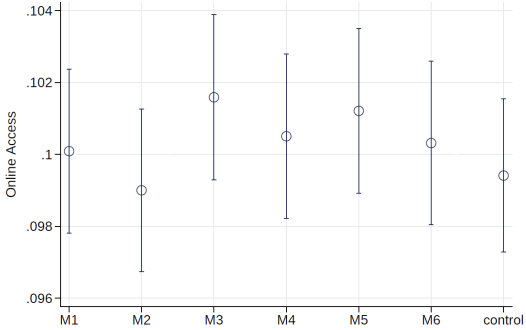
- 1) A **random assignment group**, in which students are randomly assigned to 1 out of the 6 sequences inspired by the interventions in the pilot experiment;
- 2) A **standard targeting based on online access group**, in which students are assigned to the sequence inspired in the intervention with the highest *average treatment effect* in the pilot, based on treatment effects on access to online activities;
- 3) A **standard targeting based on attendance and grades group**, in which students are assigned to the sequence inspired in the intervention with the highest *average treatment effect* in the pilot, based on treatment effects on 4th-quarter attendance and grades;
- 4) An **individual targeting based on online access group**, in which students are assigned to the sequence inspired in the intervention with the the highest *conditional average treatment effect*, learned from the pilot results and predicted using the new cohort's personal characteristics, based on treatment effects on access to online activities; and
- 5) An **individual targeting based on attendance and grades group**, in which students are assigned to the sequence inspired in the intervention with the the highest *conditional average treatment effect*, learned from the pilot results and predicted using the new cohort's personal characteristics, based on treatment effects on 4th-quarter attendance and grades.

The study comprises 240,000 students. Randomization is undertaken at the classroom level. Whenever there are multiple classrooms within a grade at the school, we assign all of them to the same treatment arm to minimize concerns with spillovers. We assign approximately 48,000 students to each group, as follows:

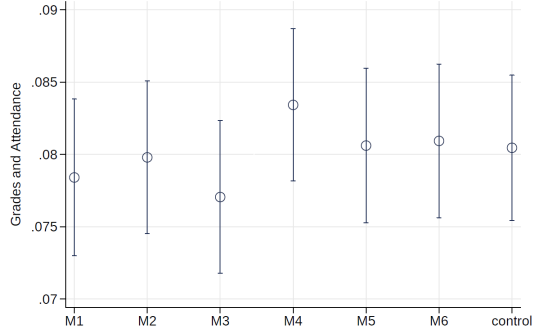
Group	Sample size
Random assignment	1,183 classrooms 47,561 students
Standard targeting based on online access: High returns to effort	1,187 classrooms 48,761 students
Individual targeting based on online access	1,203 classrooms 46,464 students
Standard targeting based on attendance and grades: Low costs of effort	1,182 classrooms 49,505 students
Individual targeting based on attendance and grades	1,190 classrooms 48,567 students

As the table highlights, the pilot intervention with the highest average treatment effect on online access was that highlighting high returns to effort, whereas the one with the highest average treatment effect on a summary measure of Math and Portuguese attendance and grades (Kling, Liebman and Katz, 2017) was that highlighting the value of future-orientation. The figures below showcase average treatment effects for each outcome¹:

Panel A – Online access



Panel B – Summary measure of attendance and grades



¹ In Panel B, one can see that students for whom we have phone numbers are positively selected on attendance and grades with respect to the universe of students (as their summary measure averages above zero even for the control group). The extent of selection is large, equivalent to one additional quarter when it comes to differences in test scores. This is not a concern for our predictions because, in the follow-up experiment, we also restrict attention to students with active phone numbers in the Secretariat’s dataset.

V. Outcomes

We will assess how the different groups of nudges impact the outcomes of students enrolled in grades 6 to 12. To do so, we will use administrative records on:

- Weekly access to the distance learning platform;
- Weekly time online on the distance learning platform;
- Weekly attendance in-person attendance (by school subject);
- Quarterly grades (by school subject);
- Student dropouts.
-

VI. Estimation

We estimate OLS regressions as follows:

$$Y_{icst} = \alpha + \beta_1 Std_{cs}^1 + \gamma_1 Ind_{cs}^1 + \beta_2 Std_{cs}^2 + \gamma_2 Ind_{cs}^2 + \theta_{1i} + \theta_{2i} + \varepsilon_{icst},$$

where Y_{icst} is the outcome of interest for student i at classroom c at school s at time t ; $Std_{cs}^1 = 1$ if classroom c at school s is assigned to the optimal standard targeting based on online access, and 0 otherwise; $Ind_{cs}^1 = 1$ if classroom c at school s is assigned to the optimal individual targeting based on online access, and 0 otherwise; $Std_{cs}^2 = 1$ if classroom c at school s is assigned to the optimal standard targeting based on attendance and grades, and 0 otherwise; $Ind_{cs}^2 = 1$ if classroom c at school s is assigned to the optimal individual targeting based on attendance and grades, and 0 otherwise; θ_{1i} are fixed-effects for the best predicted intervention for student i based on online access; θ_{2i} are fixed-effects for the best predicted intervention for student i based on attendance and grades; and ε_{icst} is an error term.

The excluded category is the random assignment group. We are interested in testing $\beta_1, \beta_2 \geq 0$, $\gamma_1 \geq \beta_1$ and $\gamma_2 \geq \beta_2$, and $\gamma_2 \geq \gamma_1$. We cluster standard errors at the classroom level, which is the unit of randomization.

Whenever we have access to multiple outcome variables mapped into a single outcome category (e.g. grades for several school subjects), we will build summary measures:

following Kling, Liebman and Katz (2007), we will normalize all outcomes to z-scores, and run seemingly unrelated regressions (SUR) to compute effect sizes for each outcome category.

We will also estimate heterogeneous treatment effects by predicted optimal individual intervention in each case. In particular, because we worry that the impacts of nudges (regardless of content) might be higher with in-person classes than during the pilot experiment, we will estimate treatment effects excluding from all treatment arms students whose optimal individual targeting prescribes no intervention.

REFERENCES

- Bergman, P. (2019). How behavioral science can empower parents to improve children's educational outcomes. *Behavioral Science & Policy*, 5(1):52-67.
- Bettinger, E., Ludvigsen, S., Rege, M., Solli, I. F., & Yeager, D. (2018). Increasing perseverance in math: Evidence from a field experiment in Norway. *Journal of Economic Behavior & Organization*, 146, 1-15.
- Bettinger, E., Cunha, N., Lichand, G. & Madeira, R. (2020). Are the Effects of Informational Interventions Driven by Salience? *University of Zurich Department of Economics Working Paper Series*, No. 350, February 2021.
- Kling, J. R., Liebman, J. B., and Katz, L. F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75(1):83–119.
- Künzel, Sören R, Jasjeet S Sekhon, Peter J Bickel, and Bin Yu (2019). Metalearners for estimating heterogeneous treatment effects using machine learning. *Proceedings of the National Academy of Sciences*, 116(10):4156–4165.
- Yeager, D. S., Hanselman, P., Walton, G. M., Murray, J. S., Crosnoe, R., Muller, C., ... & Paunesku, D. (2019). A national experiment reveals where a growth mindset improves achievement. *Nature*, 573(7774), 364-369.