**Work Sample Instructions**

This work sample task represents the type of work that Research Coordinators at the Behavior Change for Good Initiative (BCFG) are asked to do. Your goal is to analyze a dataset and present your results as if writing for an academic paper.

**Background**

We are providing you with a dataset from a simulated megastudy (i.e., we made it up!). In this imaginary megastudy, we partnered with a pharmacy to test seven different interventions to encourage pharmacy patients to get vaccinated against RSV, along with a control condition that consisted of the pharmacy’s business as usual text messages. The interventions consisted of a single text message sent to pharmacy patients. Vaccinations were measured within one month of patients receiving the intervention. Details of the intervention messages are included in the codebook, linked below.

**Pre-registered analysis plan**

For the main analysis, we will run an ordinary least squares regression (OLS) to predict whether a given participant received an RSV vaccination (a binary indicator variable). The primary predictor variables in our regression will be 7 indicators for assignment to each of the study’s 7 experimental conditions, with an indicator for the placebo control condition omitted. Our regression will include the following control variables:

* Participant age
* Indicators for participant race (Black, Asian, white, other/unknown; white omitted)
* Indicator for participant ethnicity (Hispanic; non-Hispanic omitted)
* Indicators for participant gender (male, female, other/unknown; female omitted)
* Indicators for pharmacy region (Northeast, Midwest, Southeast, West, Southwest; Northeast omitted)
* Median household income of the county in which the pharmacy is located

**What you’ll need**

1. A participant-level dataset: here
2. A pharmacy-level dataset, which you’ll need to merge with the participant level dataset
3. A codebook explaining the variables in each of the datasets

**The task**

For this task, you’ll need to do the following:

1. **Analyze the datasets.**

Based on the pre-registration proposal, analyze the data using the software program with which you are most comfortable. Ensure that your code is written in a clear and comprehensible manner, so that other team members can easily follow your logic and decisions.

1. **Present your results.**

The first table should contain the results from the pre-registered main analysis, following the format of Table 1 in [Milkman et al.'s 2022 paper](https://www.pnas.org/doi/10.1073/pnas.2115126119). The second table should display patient characteristics. Tables should be formatted to effectively communicate with other team members. Include any necessary notes beneath the tables to clarify the analyses you conducted.

1. **Write up your method and results.**

Explain the analyses that you ran and what you found. Your writeup should include the following:

1. An explanation of the analyses you ran. This should be no more than 100 words, and should be written as if for an academic publication.
2. A summary of your findings, explaining what you found. This should be written as if for an academic publication. This should use no more than 200 words.

**What you should send**

Please send us the following:

1. The code you used to clean and analyze the data
2. The output from your analyses
3. A word document or PDF with your tables and writeup of your analyses and results