Analysis Plan: "How COVID-19 Information affects Behavior and Policy Preferences"

Jonathan Rothwell

Gallup

Michael E. Darden

Johns Hopkins University and NBER

May 6, 2022

The purpose of the research is twofold. First, our goal is to better understand how information affects attitudes and behavior. Our study will contribute to this literature in the context of COVID-19, with an emphasis on behavior that has clear economic implications (e.g., eating at restaurants, traveling, etc.). Second, our work will shed light on the role of beliefs in explaining hesitancy to get vaccinated against COVID-19. Lack of vaccine uptake remains a serious problem around the world, and our experiment will help test which messages might be effective or counter-productive in encouraging people to take the vaccine.

We will randomize survey respondents into 12 different information arms (plus one control arm in which no information is provided). Each information arm includes some factual information on COVID-19 vaccines and COVID-19 transmission. We will then follow up with several behavioral and attitudinal questions, including plans for in-person dining, in-person gathering, and, if not already, plans for receiving the COVID-19 vaccine. We will test how the effect of information on behavioral and attitudinal beliefs varies by a.) prior beliefs and b.) political affiliation. Prior beliefs will be measured both with numerical assessments of the COVID-19 hospitalization rate as well as categorical assessments on the relative severity of the Omicron variant. For example, we will ask individuals to assess the severity, in terms of hospitalization and mortality rates, of the Omicron variant of COVID-19 relative to previous variants.

We will start by assessing balance in demographic, socioeconomic, and prior belief characteristics across the treatment arms. Next, we will estimate two main regression models. First, to understand the treatment effects of the 12 types of information, we will estimate:

$$y_i = \beta_0 + \sum_{j=1}^{12} \beta_j 1[InformationArm = j] + \epsilon_i, \tag{1}$$

where β_0 captures the mean level of attitude or belief or behavior y_i for the control group who received no information, and β_j captures how the mean value of y_i differs for information treatment j. We will estimate this regression with and without controls for baseline demographic and socioeconomic characteristics. The second regression will explore heterogeneity in the treatment effects. We will estimate:

$$y_i = \beta_0 + \sum_{j=1}^{12} \beta_j 1[InformationArm = j] + \gamma_0 Z_i + \sum_{j=1}^{12} \gamma_j 1[InformationArm = j] * Z_i + \epsilon_i, \qquad (2)$$

where the γ parameters capture how the treatment effects differ for those individuals with characteristic Z_i. These models will also be estimated with and without individual level controls.

The next step in our analysis will be to utilize followup longitudinal information. After two-months, we will re-survey our original sample to test if the effects of information last only within the experiment itself or for as long as two-months after the experiment, when participants are re-interviewed. The econometric model will assess the consistency of attitudes and beliefs by treatment arm.

Finally, we will ask respondents to voluntarily reveal whether they use Twitter and, if so, whether they will share their Twitter handle (ID). For the selected sub-sample who agree to provide their Twitter identity, we will merge Twitter handles with internal Twitter information that is made freely available to scholars. These data have been used by other researchers to study, for example, the spread of misinformation. To proceed, we will first characterize individuals' past behavior on Twitter with such metrics as an index score of retweeted or liked tweets, an index score of average account activity, a measure of whether past tweets mention COVID or related words, and a measure of percent of tweets sent outside home. We will use these past Twitter behaviors as Z_i characteristics to which we can better understand heterogeneity in our treatment effects. Next, we will model subsequent measurements of these Twitter variables (ex post) as behaviors that may have been influenced by the information treatment.