Pre-Analysis Plan for "Predicting Biased Polls"

Andy Brownback* Nathaniel Burke† Tristan Gagnon-Bartsch‡

December 2018

1 Choice Data

In the first stage of the study, we will collect data on choices over each of eight different binary actions, labeled by A. (For an example of such an action, we will ask participants if they would like to donate $\$x$ to a given organization.) We will collect this data from two groups: in one group (labeled "IC"), we will elicit choices in an Incentive-Compatible fashion; in the other (labeled "H"), we will instead elicit hypothetical choices. For each action A and participant i, we denote i's choice regarding action A by $IC_{i,A}$ if i is in the IC group and by $H_{i,A}$ if i is in the H group.

The only difference between the elicitations for the IC and H groups is the incentive-compatibility of the IC group's elicitation. (For instance, if a participant in this group chooses to donate $\$x$ to a given organization, then that participant will really lose $\$x$ dollars and the organization will actually receive a donation of $\$x$ dollars.) The H group has no such incentive to make choices that reveal their true preferences since their stated choices are purely hypothetical. Thus, for each action A, we will define socially desirable responding (SDR) as the gap between the average choice in the two groups, denoted by

$$SDR_A = H_A - IC_A.$$

Action A is socially desirable if the subjects in the H group state a higher demand for taking that action—since they do not have to pay the cost of taking the action—than subjects in the IC group.

2 Prediction Data

We are primarily interested in the ability of Predictors to guess the behavior of subjects in the IC group. That is, how well can people guess the incentivized choices of other people? Moreover, we want to examine how these guesses respond to information about others' choices when it comes from the incentivized group (IC) versus the unincentivized, hypothetical group (H). That is, to what extent do people appreciate the different informational content of others' incentivized choices relative to their stated hypothetical choices.

*University of Arkansas: ABrownback@walton.uark.edu.

†University of Arkansas: NBurke@walton.uark.edu

‡Harvard University: gagnonbartsch@fas.harvard.edu.

All Predictors will begin the study by stating guesses about the choice rates of the IC group for all eight actions. Let $\textsc{Guess}_{i,1,A}$ denote Predictor i's initial guess about this choice rate for action A.

Predictors will then be randomly-assigned to receive information from either the IC group or the H group. They will then receive a "signal" in the form of the choices of ten randomly-chosen individuals from their assigned group. In this way, we provide two sources of randomness: the source of the information (IC or H group), and the idiosyncratic variation in the ten observed individuals.

After receiving their signal, Predictors will make a second set of guesses about the behavior of the IC group. Let $\textsc{Guess}_{i,2,A}$ denote Predictor i's revised guess about the choice rate for action A.

3 Mechanical Turk Extension (Planned)

In addition to considering how Predictors drawn from the same population (University of Arkansas undergraduates) predict the behavior of subjects in the IC group, we also hope to extend the recruitment of Predictors to subjects from a different population. This will reveal the extent to which "social distance" interferes with predictions and inferences about the behavior of a target group.

This secondary sample of Predictors will participate in the same tasks as the original Predictors. From them, we will elicit $\textsc{Guess}_{i,1,A}$ and $\textsc{Guess}_{i,2,A}$.

4 Sentiment

Alongside the Choice and Prediction aspects of the study, we will also collect information on the sentiment associated with each of the eight actions. More specifically, subjects will answer the following questions on a scale of 1-10:

1. How do you feel about taking this action yourself? (1 – very negative, 10 – very positive)
2. How do you feel about other people who take this action? (1 – very negative, 10 – very positive)
3. How do you think most other people feel about people who take this action? (1 – very negative, 10 – very positive)

For each action A, let $Q_{i,j,A}$ denote participant i's answer to question $j \in \{1, 2, 3\}$ above. From these, we will construct $V_{i,A} = \frac{Q_{i,1,A}+Q_{i,2,A}+Q_{i,3,A}}{3}$ and use it as an index of the perceived "social-desirability" of action A.

5 Data

- Choices

  - Subjects: 100 for the IC group; 100 for the H group
  - Population: University of Arkansas undergraduates
  - Data: For each action A and subject i, we will record their choices—either $H_{i,A}$ or $IC_{i,A}$.
  - Note: the set of all $H_{i,A}$ and $IC_{i,A}$ will be used to construct the signals of behavior that the Predictors will receive. Specifically, we will construct each signal by drawing 10 observations at random with replacement. Each Predictor will have a signal that was drawn independently.

- Predictions
  - Subjects: 300 – 100 each from the IC and H groups, and 100 new subjects)
  - Population: University of Arkansas undergraduates
  - Data: For each action A and Predictor i, we will elicit two guesses—$\textsc{Guess}_{i,1,A}$ will be elicited prior to receiving a signal about choice behavior and $\textsc{Guess}_{i,2,A}$ will be elicited after.

- Sentiment
  - Subjects: 40
  - Population: University of Arkansas undergraduates who participated in neither the Choices nor Predictions parts
  - Data: For each action A and subject i, we will measure sentiment, $V_{i,A}$.

- Mechanical Turk Extension (Planned)
  - Subjects: 1,000
  - Population: Mechanical Turk workers
  - Data: For each action A and Predictor i, we will capture two guesses. $\textsc{Guess}_{i,1,A}$ will be before receiving a signal about choice behavior and $\textsc{Guess}_{i,2,A}$ will be after receiving the signal.

6 Hypotheses

6.1 Hypothesis 1 (SDR): The gap between the H and IC groups will grow as the social desirability of an action grows.

Recall that we defined "socially desirable responding" (SDR) as the gap between the H and IC groups. To buttress our claim that this gap is indeed due to a social desirability bias, we will examine how it relates to the social-desirability index $V_{i,A}$ collected from other subjects in the same undergraduate population as the H and IC groups.

We will run the following regression using each of the eight actions as an observation:

$$SDR_A = \beta_0 + \beta_1 \bar{V}_A + \epsilon_A, \tag{1}$$

where $\bar{V}_A$ is the average sentiment for an action A across all subjects.

We will test if $\beta_1 > 0$. This will reveal if sentiment correctly predicts our measure of SDR. This test will be under-powered, as it only has 8 observations, but should provide cross-validation for our chosen metric of SDR.

6.2 Hypothesis 2 (Accuracy): Predictors will be more accurate with information from the IC group.

This hypothesis simply states that revised guesses *about* the IC group will become more accurate for Predictors who receive their signal *from* the IC group rather than the H group.

If predictors knew the value of SDR_A for each action, then they could infer the choice rates of the IC group from observing the choice rates of the H group. However, any inaccuracies in the predictions about SDR_A will result in less accurate guesses after receiving a signal from the H group rather than the IC group.

To test this prediction, we will look at how both the absolute errors ($ABS_{i,t,A} = |IC_{t,A} - \textsc{Guess}_{i,t,A}|$) and squared errors ($SQ_{i,t,A} = (IC_{t,A} - \textsc{Guess}_{i,t,A})^2$) in the Predictors' guesses change after receiving their signal. We will use a random-effects linear regression to test for the impact of the signal source on the accuracy of the guesses:

$$ABS_{i,2,A} = \beta_0 + \beta_1 ABS_{i,1,A} + \beta_2 \mathrm{IC}_i + \delta_A + \nu_i + \epsilon_{i,A}, \tag{2}$$

$$SQ_{i,2,A} = \beta_0 + \beta_1 SQ_{i,1,A} + \beta_2 \mathrm{IC}_i + \delta_A + \nu_i + \epsilon_{i,A}, \tag{3}$$

where IC_i are indicator variables equal to 1 if Predictor i received a signal from the IC group, and ν_i are Predictor random-effects (meaning they will not be individually identified). We will cluster standard errors at the individual level.

Our test of increased accuracy amounts to a test of $\beta_2 < 0$.

6.3 Hypothesis 3 (Updating): Predictors' updates will be more sensitive to incentive-compatible information.

This hypothesis posits that Predictors will recognize that signals from the IC group are more predictive of the IC group's behavior than signals from the H group. In practice that would mean that Predictors' guesses update more strongly toward signals from the IC group than from the H group.

Members of the IC and H groups face different incentives with their responses, thus we anticipate that the distribution of choices from the two groups will differ. This means that signals of behavior from each of the two groups will be drawn from different distributions. As such, a Predictor's response to their signal is not identified separately from the average distance of the distribution of signals from the Predictor's original guess.

Recall that each Predictor only observes the responses of a subset of the assigned group – ten randomly-selected subjects from either the IC or H group. This randomly-selected subset helps us identify the distinct effect of each signal source. Each subset will be randomly drawn, meaning that the signals will have randomness from the sampling variation. After controlling for the mean choice rate in the assigned group (either IC or H), the residual variation in signals is mechanically random. The response of a Predictor to this residual randomness will causally identify how the different types of signals affect Predictors' revised guesses. We will test this using the following random-effects linear regression:

$$\textsc{Guess}_{i,2,A} = \beta_0 + \beta_1 \textsc{Guess}_{i,1,A} + \beta_2 S_{i,A} + \beta_3 S_{i,A} \times \text{IC}_i + \beta_4 \text{IC}_i + \beta_5 \bar{S}_{T,A} + \delta_A + \nu_i + \epsilon_{i,A}, \quad (4)$$

where $S_{i,A}$ is the signal received by Predictor i for item A (i.e., the fraction of subjects from i's random sample of 10 who took action A), and $\bar{S}_{T,A}$ is the mean of the distribution of signals from group T (either IC or H) for item A. By controlling for $\bar{S}_{T,A}$, we are able to use $S_{i,A}$ to identify the effect of a change in the signal that is derived only from sampling variation in the selection of the ten observed subjects—that is, a mechanically-random change in the signal. Again, ν_i are Predictor random-effects, and we will cluster standard errors at the individual level.

To test for differential sensitivity to signals from the IC group, we will test if $\beta_3 > 0$.

6.4 Hypothesis 4 (SDR Updating): Predictors' updates will be increasingly sensitive to incentive-compatible information as the social-desirability of an action becomes more extreme.

If the Predictors can correctly infer the actions for which SDR will be most extreme, then we expect them to discount signals from the H group about these actions by a greater extent. That is, the hypothetical statements that are prone to greater bias should be discounted by more.

To test this hypothesis, we will use the same random-effects linear regression specification as in Equation 4, but will include terms interacted with the absolute value of our measure of SDR:

$$\begin{aligned}\textsc{Guess}_{i,2,A} = \beta_0 + \beta_1 \textsc{Guess}_{i,1,A} + \beta_2 S_{i,A} + \beta_3 S_{i,A} \times \text{IC}_i + \beta_4 \text{IC}_i + \beta_5 S_{i,A} \times |\text{SDR}_A| \quad (5)\\ + \beta_6 S_{i,A} \times \text{IC}_i \times |\text{SDR}_A| + \beta_7 \text{IC}_i \times |\text{SDR}_A| + \beta_8 |\text{SDR}_A| + \beta_9 \bar{S}_{T,A} + \delta_A + \nu_i + \epsilon_{i,A}\end{aligned}$$

Here, we have interacted all of the relevant terms from Equation 4 with the absolute value of our measure of SDR for the action A, $|SDR_A|$. We will test if signals from the IC group are weighted more heavily (relative to signals from the H group) as SDR becomes more extreme. This amounts to testing if $\beta_6 > 0$.

6.5 Hypothesis 5 (Experience): Predictors who previously participated in the IC or H groups will discount signals from the H group more than those who did not.

The subjects who participated in the H and IC groups will return to become Predictors. They will join other Predictors who have no such experience. Thus, some of our Predictors will have first-hand experience with the choices and other Predictors will not. This will allow us to test if those with experience are more capable of de-biasing the signals than Predictors who had no such experience.

To test this hypothesis, we will use the same random-effects linear regression specification as in Equation 4, but will include terms interacted with the role previously played by the Predictor:

$$\begin{aligned}\textsc{Guess}_{i,2,A} = \beta_0 + \beta_1 \textsc{Guess}_{i,1,A} + \beta_2 S_{i,A} + \beta_3 S_{i,A} \times \mathrm{IC}_i + \beta_4 \mathrm{IC}_i + \beta_5 S_{i,A} \times \textsc{Exp}_i \\ + \beta_6 S_{i,A} \times \mathrm{IC}_i \times \textsc{Exp}_i + \beta_7 \mathrm{IC}_i \times \textsc{Exp}_i + \beta_8 \textsc{Exp}_i + \beta_9 \bar{S}_{T,A} + \delta_A + \nu_i + \epsilon_{i,A}\end{aligned} \tag{6}$$

In this equation, $\textsc{Exp}_i$ is an indicator variable equal to one if the Predictor has previous experience participating in the IC or H group. Again, we will test for a significant interaction effect by testing if $\beta_6 > 0$.

We will repeat this analysis looking at members of the H and IC groups separately. This will reveal heterogeneity in the learned experience of the two groups.