Pre-analysis Plan

Daiki Kishishita, Atsushi Yamagishi, Tomoko Matsumoto

February 18, 2022

We describe our pre-analysis plan for the project “Public Goods, Welfare Policies and Government Size."

1 Experimental design

We will post the survey openly on Amazon Mechanical Turk (MTurk). Our targeted number of observations is 2000.

In the survey, following a question detecting a satisficer, respondents will be asked to answer questions on demographics and political attitudes. Then, half of the respondents will be randomly assigned to the treatment. After the treatment, they will be asked to answer the questions on preferences for taxation and public spending.

In particular, we have four groups:

1. Treatment group who receives the information emphasizing the governmental role of providing public goods
2. Control group for the comparison with 1
3. Treatment group who receives the information emphasizing the governmental role of helping the poor
4. Control group for the comparison with 3

2 Analysis plan

Balance test By using demographic questions (e.g., gender, age, income), we conduct a balance test to ensure that randomization is successfully done.

2.1 Public goods provision

We first focus on the samples belonging to either group 1 or 2.

Manipulation check We conduct a manipulation check, which examines whether our treatment changes each respondent's perception on the governmental role in providing public goods. Across groups, we ask "How much of your taxes do you think are used for public goods and services that benefit all of you?" We consider the following linear regression model:

$$y_i = \beta_1 treated_i + \gamma X_i + \epsilon_i, \tag{1}$$

where y_i is the answer to this question, $treated_i$ is the treatment status of person i, X_i are the observed characteristics, and ϵ_i is the error term. By estimating (4) by OLS, we investigate the causal effect of the treatment β_1. The randomization ensures that β_1 indeed captures the causal effect. In the survey, we also collect information on age, gender, childbearing status, marital status, race, state of residence, urbanity of residence, education, employment status, and income. We collect these variables as observed characteristics that could be included in the control variables X_i.

Our hypothesis is $\beta_1 > 0$. By testing this hypothesis, we conduct manipulation check.

Main outcomes The aim of this study is to identify the effect of our treatment on policy preferences for taxation and public spending. In the survey, we ask preferences for the government size (e.g., whether one agrees to increase tax rates), preferences for the desirable tax system (e.g., progressive vs. uniform vs. lump-sum), and preferences for public spending (e.g., whether it should be used for helping the poor). These are the main outcomes of our study.

We analyze the causal effect of the treatment on these main outcomes as in the case of manipulation check. That is, we consider the following linear regression model:

$$y_i = \beta_1 treated_i + \gamma X_i + \epsilon_i \tag{2}$$

where y_i is the outcomes of interest. As a specific example, consider the support for a hypothetical proposal for tax increase. We ask "Imagine that the government says it needs to increase everyone's tax bill by 1%. Do you agree with this tax increase?" Agreement is coded as 1 and disagreement is coded as 0. Then, (5) is interpreted as the linear probability model for agreeing with the tax increase. The details of our main outcomes can be seen in the questionnaire. We conduct a hypothesis test regarding whether $\beta_1 = 0$ for each of our main outcomes.

Heterogeneous treatment effect We would also investigate the potential heterogeneity of the treatment effect. We estimate the equation of the following form:

$$y_i = \beta_1 treated_i + \beta_2 treated_i \times c_i + \gamma X_i + \epsilon_i, \tag{3}$$

where c_i is the characteristic in which we suspect the heterogeneous impact. For example, c_i can be a dummy variable coding whether person i trusts in the government. β_2 captures the differential treatment effect between those who trust the government and those who do not. We conduct a hypothesis test on the null hypotheses $\beta_1 + \beta_2 c = 0$. We also investigate the heterogeneity with respect to socioeconomic status (e.g., income) and the political position (i.e., liberal vs conservative).

2.2 Welfare policy

We next focus on the samples belonging to either group 3 or 4.

Manipulation check We conduct a manipulation check, which examines whether our treatment changes each respondent's perception on the governmental role in helping the poor. Across groups, we ask "How much of your taxes do you think are used for helping the poor?" We consider the following linear regression model:

$$y_i = \beta_1 treated_i + \gamma X_i + \epsilon_i, \tag{4}$$

where y_i is the answer to this question, $treated_i$ is the treatment status of person i, X_i are the observed characteristics, and ϵ_i is the error term. By estimating (4) by OLS, we investigate the causal effect of the treatment β_1. The randomization ensures that β_1 indeed captures the causal effect. In the survey, we also collect information on age, gender, childbearing status, marital status, race, state of residence, urbanity of residence, education, employment status, and income. We collect these variables as observed characteristics that could be included in the control variables X_i.

Our hypothesis is $\beta_1 > 0$. By testing this hypothesis, we conduct manipulation check.

Main outcomes The aim of this study is to identify the effect of our treatment on policy preferences for taxation and public spending. In the survey, we ask preferences for the government size (e.g., whether one agrees to increase tax rates), preferences for the desirable tax system (e.g., progressive vs. uniform vs. lump-sum), and preferences for public spending (e.g., whether it should be used for helping the poor). These are the main outcomes of our study.

We analyze the causal effect of the treatment on these main outcomes as in the case of manipulation check. That is, we consider the following linear regression model:

$$y_i = \beta_1 treated_i + \gamma X_i + \epsilon_i \tag{5}$$

where y_i is the outcomes of interest. As a specific example, consider the support for a hypothetical proposal for tax increase. We ask "Imagine that the government says it needs to increase everyone's tax bill by 1%. Do you agree with this tax increase?" Agreement is coded as 1 and disagreement is coded as 0. Then, (5) is interpreted as the linear probability model for agreeing with the tax increase. The details of our main outcomes can be seen in the questionnaire. We conduct a hypothesis test regarding whether $\beta_1 = 0$ for each of our main outcomes.

Heterogeneous treatment effect We would also investigate the potential heterogeneity of the treatment effect. We estimate the equation of the following form:

$$y_i = \beta_1 treated_i + \beta_2 treated_i \times c_i + \gamma X_i + \epsilon_i, \tag{6}$$

where c_i is the characteristic in which we suspect the heterogeneous impact. For example, c_i can be a dummy variable coding whether person i trusts in the government. β_2 captures the

differential treatment effect between those who trust the government and those who do not. We conduct a hypothesis test on the null hypotheses $\beta_1 + \beta_2 c = 0$. We also investigate the heterogeneity with respect to socioeconomic status (e.g., income) and the political position (i.e., liberal vs conservative).