

Problem Set #5

MACS 30000, Dr. Evans

Nan Ge

Question 1. Experiments on Amazon Mechanical Turk

(a) Search for an experiment on MTurk that interests you.

The experiment's name is: Smeeble day 3, continue to learn a small part of a new language.¹

(b) Describe the full payment structure of this experiment. That is, the reward column says an amount, but there is a lot more information available as to what that amount means.

This experiment is the third part of a series experiment study. This section alone takes 30-45 minutes to complete, and the participant would be reward \$4-\$6 depending on his/her progress in the experiment. The further a participant gets, the more he/she will be paid. Those who progress through the whole test will unlock another HIT in the next stage of the experiment as an extra bonus.

(c) Describe any qualifications, eligibility requirements, or restrictions (or lack thereof).

1. The participant must have successfully finished the first and the second part to unlock this HIT.
2. The participant must have read the information letter to begin the experiment.
3. The participant must voluntarily agree to participate.
4. The participant must be at least 18 of age.
5. The participant must be a native speaker of English.
6. The participant must have device and environment to support audio listening.

(d) How long does this job take? What is the implied hourly rate (dollars per hour)?

This job takes 30-45 minutes to finish. The implied hourly rate varies between \$5.3 to \$8 dollars per hour.

(e) When does this job expire?

This job expires at 6 A.M., Nov 8th 2018.

(f) What is the most this project would cost the HIT experiment creator if 1 million people participated in the task?

¹https://worker.mturk.com/projects/3ZCNYL2D77MCS6GJ88QJON5PYI76MX/tasks?ref=w_pl_prvw (No longer available)

The maximum reward for this experiment is \$8. Therefore, if 1 million people participated, and all participants completed the experiment successfully, it would cost 8 million dollars.

Question 2. The Nudging Experiment

Here is the million-dollar question: how do we get people to consume less electricity? There have been numerous campaigns trying to reach the same goal, but heterogeneity is always a headache for policy makers and electricity companies. Different cohorts have surprisingly divergent responses to the same message. Sometimes campaign impacts would cancel each other on the aggregate level. Thus, it's vital to understand how does heterogeneity interplay with these campaign messages. In this paper, the authors raised this question: how does ideology affect people's response to electricity conservation nudges?

They explored the proposition with a compound and detailed dataset. First part of the data comes from the electric utility. It consists of residential billing data from January 2007 to October 2009. (Costa and Kahn, 2013, p.685) First Home Energy Report (HER) report was sent out between March 14 and May 9 2008, so the billing dataset covers information with approximately equal time length before and after the treatment. (Costa and Kahn, 2013, p. 683) It provides information on kilowatt hours purchased per billing cycle, the length of the billing cycle (measured in days), whether the house uses electric heat, and whether the household is enrolled in the renewable energy program. (Costa and Kahn, 2013, p. 685) Main body of the experiment data contains information on when the household first received HER report, and attributes about the house, such as square footage of the house, whether the home heats with electricity or natural gas, and the age of the house. (Costa and Kahn, 2013, p. 685) Key information about the household's ideology comes from a third sub-dataset consisting of voter registration and marketing data for March 2009. (p.685) It reports party affiliation, and whether the individual donates to environmental organizations. (Costa and Kahn, 2013, p. 685) Ideological characteristics of the block were then merged from census data, including the share of registered liberal voters and share of the college-educated in the block group. (Costa and Kahn, 2013, p. 686)

Candidates of the experiment are current customers with the electric utility. All of them live in a house with square footage between 250 and 99,998 square feet. (Costa and Kahn, 2013, p.683) They are selected into the experiment based on 85 census tracts to ensure a high density of single-family homes. (Costa and Kahn, 2013, p.683) Then the households were randomly assigned to treatment and control group. The researchers assigned five contiguous census blocks, i.e. a census block batch into the treatment group, then the next batch into the control group, and repeated this process until both groups reached roughly 35000 households. (Costa and Kahn, 2013, p.683) The remaining households went to the control group. Those in the treatment group would receive monthly or quarterly HERs about their absolute electricity consumption and relative consumption level in the neighborhood. The researchers then tested whether the ideology has a role to play in the treatment effect.

This paper builds on a growing literature about the nudging effect on electricity

consumption. [Schultz et al. \(2007\)](#) confirmed of the boomerang effect of normative messages. They also reported injunctive message could eliminate the boomerang effect. [Costa and Kahn \(2013\)](#) contributes to the literature by controlling for heterogeneity in ideology of the household. In particular, they used four measurements to outline ideology: their registered party, whether they live in a liberal or conservative community, whether they previously paid for electricity generated by renewable resources, and whether they donated to environmental organizations. ([Costa and Kahn, 2013](#), p. 681)

The experiment conveys very rich information. Overall, the authors have convinced us that ideology does affect the nudging effect. In a baseline analysis, registered liberals decrease electricity consumption by 0.7% more than registered conservatives. ([Costa and Kahn, 2013](#), p.691 Table3) After controlling for behavioral measurement of ideology, the effect of party affiliation becomes insignificant, but people who donate to environmental organizations, and who purchase renewable electricity reduced their consumption significantly more than others. Also, receiving the HER reports would lead to a 3.8% more reduction if the household is in a census block where the share of liberals is in the top 75th percentile. Combining these causes, registered liberal households who paid for renewable electricity, donated to environmental group and in a liberal block reduced 3.6%-4.8% more electricity consumption. ([Costa and Kahn, 2013](#), p.691 Table4) As a second step, the authors employed probit model to predict the possibility of turning down air-conditioning in summer and the possibility of opting out of the HER program. The results show that liberals are 15% less likely to opt out of the treatment, and they are more likely to turn down air-conditioning in summer. ([Costa and Kahn, 2013](#), p. 682) The behavioral measurements yield similar predictions. Lastly, the authors found that whether the HER reports consist of descriptive message or normative message doesn't have a significant difference on their nudging effect.

Question 3. Analytical Exercise

(a) Under what conditions might it be better to focus your resources on a small number of clinics and under what conditions might it be better to spread them more widely?

In this experiment, the unit is each patient in a clinic. The treatment is receiving text message reminders on vaccine uptake. However, we can't observe the natural state of the treatment group. That is, we don't know if the patient in the treatment group would take a vaccine upshot had he/she not received the reminder. Therefore, one key assumption is that the patients in the control and treatment group are identical and randomly assigned. If the assumption holds, we can estimate the average treatment effect as:

$$\widehat{ATE} = \frac{1}{N_t} \sum_c \sum_{i:W_i=1} Y_{ci}(1) - \frac{1}{N_c} \sum_c \sum_{i:W_i=0} Y_{ci}(0)^2 \quad (1)$$

²Salganik 2018, p. 208

where c denotes the clinic, and i is the identity of the patient. The condition of the randomization assumption is that the patients in the society are randomly distributed among clinics, i.e. each clinic's patient is a random draw from the whole cohort. However, if patients are allowed to choose their clinic, there would exist a selection bias among different clinics' patients, which could contaminate our randomization assumption. For example, clinic in a neighborhood with good socioeconomic profile may have patients that pay more attention to immunization, and would take upshots regardless of the text message reminder. Therefore, we might observe the treatment effect is smaller in that clinic. If the selection bias exists, we would need expand our clinic sample to cover a variety of clinics. If the selection bias is not an issue, we could focus on less clinics and spend the budget on spending more text messages.

(b) What factors would determine the smallest effect size that you will be able to reliably detect with your budget?

The effect size is decided by mean difference between control and treatment group divided by standard error of average treatment effect.³ The effect size is calculated with:

$$Effect\ Size = \frac{\mu_1 - \mu_2}{SE(\widehat{ATE})} \quad (2)$$

$$SE(\widehat{ATE}) = \sqrt{\frac{1}{N-1} \left(\frac{m\text{Var}(Y_i(0))}{N-m} + \frac{(N-m)\text{Var}(Y_i(1))}{m} + 2\text{Cov}(Y_i(0), Y_i(1)) \right)} \quad (3)$$

As indicated by equation (2) and (3), several factors could influence the smallest effect size we are able to detect. First, the smallest effect size rely on the difference between the average vaccine upshot rate in the treatment group and the control group. Second, the standard deviation is determined by the sample size of the treatment group (m), the sample size of the control group ($N-m$), their ratio ($m/N-m$), and the degree of freedom ($N-1$). Third, the smallest effect size depends on the chosen confidence level. Fourth, the smallest effect size is also determined by the statement of the null hypothesis, i.e. whether we need a one tail or two tail hypothesis test.

References

Costa, Dora L and Matthew E Kahn, "Energy conservation nudges and environmentalist ideology: Evidence from a randomized residential electricity field experiment," *Journal of the European Economic Association*, 2013, 11 (3), 680–702.

³There are many measurements to the effect size. Here I stand with Cohen's d to begin my discussion. https://en.wikipedia.org/wiki/Effect_size

⁴Salganik 2018, p. 209

Schultz, P Wesley, Jessica M Nolan, Robert B Cialdini, Noah J Goldstein, and Vidas Griskevicius, "The constructive, destructive, and reconstructive power of social norms," *Psychological science*, 2007, 18 (5), 429–434.