Assignment 5

Jiaxu Han

1. Experiments on Amazon Mechanical Turk.

(a) The experiment I chose is called “Visual memory experiment (Bonus up to $1.50!) [cs-19fb]”.

(b) The payment structure was mentioned in two places. In the job description, it says: “Remember the exact position of a small dot inside a large image (bonus of up to $1.50)”. Additionally, they mentioned in the consent form that “the task will pay $0.20 + bonus. You can win a bonus of $1.50 extra depending on your accuracy”.

(c) There are two required qualifications for accepting the work: (1) Location is U.S. (2) HIT approval rate (%) is not less than 95. There are two additional requirements in the consent form of the experiment: the participant must be (1) at least 18 years old and (2) a fluent English speaker.

(d) The task will take about 15 minutes. The implied hourly rate ranges from $\frac{\$0.20}{15} \times 60 = \$0.80$ to $\frac{\$1.70}{15} \times 60 = \$6.80$

(e) It expires on December 8th, 2018

(f) If 1 million people participated in the task and all of them get the highest bonus possible, the project would cost the HIT experiment creator $1{,}000{,}000 \times \$\ 1.70 = \$\ 1{,}700{,}000$.

2. Costa and Kahn (2013).

Costa and Kahn (2013) conducted a randomized field experiment to answer the research question: how do political and social views influence household response to energy conservation “nudges”? (Costa & Kahn, 2013 p. 682)

The authors acquired and merged data from three major sources and one ancillary data source. The primary data set is “residential billing data from January 2007 to October 2009” (Costa & Kahn, 2013 p. 685). These data provide “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 electric utility’s program to purchase energy from renewable sources” (Costa & Kahn, 2013 p.682). Another data set is treatment and control data which “contain information on when the household began to receive the HERs, square footage of the house, whether the home heats with electricity or natural gas, and the age of the house” (Costa & Kahn, 2013 p.682). The third data set is “voter registration information and marketing data for March 2009”, from which the authors would “know party affiliation, whether the individual or a household has donated to environmental organizations, and whether the household has signed up for the company’s renewable power program prior to the treatment” (Costa & Kahn, 2013 p.682). Additionally, the authors also accessed to an ancillary data set where they obtained information on “age and the fraction of liberals in the block group and to households who were not in minor parties we could not classify as liberal or conservatives” (Costa & Kahn, 2013 p.685 - p.686).

To solve the problems of confounders, researchers often conduct a randomized controlled experiment where a researcher intervenes for some people and not for others and the decision of who receives treatment and who doesn't is totally random (Salganik & Matthew, 2018 Chapter 4.1). Therefore, the group of people who are intervened by the treatment is called the treatment group and the group of people who don't get treatment is called the control group. In this paper, around "35,000 households compose of the treatment group" (Costa & Kahn, 2013 p683). "The electric utility sent the first Home Electricity Reports (HER) to the treatment group between March 14 and May 9, 2008, and these households are still receiving the report on a quarterly or monthly basis" (Costa & Kahn, 2013 p683) when the authors were working on this paper. In contrast, "a control group of roughly 49,000 households has never received HER" (Costa & Kahn, 2013 p683). In addition, "both treatment and control households had to have a current account with the electric utility that had been active for at least one year, could not be living in apartment buildings, and had to be living in a house with square footage between 250 and 99,998 square feet" (Costa & Kahn, 2013 p683). Therefore, the treatment in HER experiment is receiving Home Electricity Reports from utility company and the report compares the electricity usage of the current household with all neighbors as well as the electricity usage in the current month with the same month in prior year and some tips for saving energy (Costa & Kahn, 2013 p 682).

Schultz et al. (2007) studied the heterogeneity effect of door-hangers on heavy users and light users. However, their study, like most analog experiments, focused on average treatment effects and failed to consider other participant heterogeneity possibly because they don't know much additional information about each participant except for whether they are heavy/light users (Salganik & Matthew, 2018 Chapter 4.4.2). In the current paper, Costa and Kahn are mainly interested in the effectiveness of HER on groups with different political ideologies "which are measured either by political party affiliation, donations to environmental organizations, or the purchase of green energy" (Costa & Kahn, 2013 p. 690). Therefore, besides controlling for heavy/light users, they also control for a series of participant heterogeneity. In "Econometric Framework" section (Costa & Kahn, 2013 p. 689), we can see that the regression analysis for the treatment effect also controlled for "household and month/year fixed effects", mean daily temperature within the billing cycle, if the house is an electric house, individual characteristics, block characteristics, and house characteristics (Costa & Kahn, 2013 p. 689). To estimate who opts out of the treatment and report reaction, the regression also controlled for the age of the head of the household (Costa & Kahn, 2013 p.689).

As a result, Costa and Kahn found that "political ideology… is associated with differential treatment effects" (Costa & Kahn, 2013 p. 690). More specifically, they found that (1) environmental "nudge" (HER treatment) is "two to four times more effective with political liberals than with conservatives" and (2) "political conservatives are more likely than liberals to opt out of receiving the home electricity report and to report disliking the report" (Costa & Kahn, 2013 p. 680).

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, June 2013, 11 (3), 680–702.

Salganik, Matthew J., Bit by Bit: Social Research in the Digital Age, Princeton University Press, 2018.

Schultz, P. Wesley, Jessica M. Nolan, Robert B. Cialdini, Noah J. Goldsteinand, and Vladas Griskevicius, "The Constructive, Destructive, and Re- constructive Power of Social Norms," Psychological Science, 2007, 18 (5), 429–434.

3. Analytical exercise.

(a)

To answer the research question of what is the effect of receiving text message reminders on vaccination uptake, the researchers may design a randomized controlled experiment. In the experiment, there is a treatment group in which patients receive text message reminders and a control group in which patients do not receive text message reminders. The treatment, therefore, is receiving text message reminders about vaccination uptake. Now we have a budget of $1,000 to conduct this experiment, but working with each clinic costs $100 in addition to the costs of text messages. Therefore, given the budget constraint, we need to decide how to allocate our resources. To make a reasonable decision, it is important to check whether the first part of stable unit treatment value assumption (SUTVA) about "no interference" or "no spillovers" (Salganik, 2018 Chapter 4 in mathematical notes) is valid in this experiment. In another word, we would like to know if participant i is not impacted by the treatment given to other people.

We might test this assumption in a pilot study and see if the following equation is valid:
$Y_i(W_i, \mathbf{W_{-i}}) = Y_i(W_i) \quad \forall \mathbf{W_{-i}}$ (Salganik, 2018, equation 4.5)

Where $\mathbf{W_{-i}}$ is a vector of treatment statuses for everyone except for person i, $Y_i(1)$ indicates treatment condition and $Y_i(0)$ indicates control condition, and W_i is either 0 or 1 (Salganik, 2018 Chapter 4 in mathematical notes).

If there is no spillover effect, we could include as many patients as possible in a smaller number of clinics as long as we have enough budget. However, when, for example, one person received a text message and she told her friends who live in the same area and go to the same clinic but in the control group, then "the treatment from one person spills over onto another person" (Salganik, 2018 Chapter 4 mathematical notes), either positively (the friend in the control group will go as well) or negatively (the friend decides to never get a vaccine because he/she felt left out) and the SUTVA assumption is violated. Under such a situation, it would be better if we collect samples from a wider range of clinics and control the sample size collected from each to limit the spillover effect.

(b)

Reliably detecting the smallest effect size relies on sufficient statistical power which may depend on several factors. Normally, setting a more conservative significance criterion and increase the magnitude of the effect of an experiment can increase the power (https://en.wikipedia.org/wiki/Power_(statistics)#Factors_influencing_power), but those are not the main concern here. Since the effect size is already small, there are another two options to increase the power to detect the effect.

First, increasing sample size can boost the statistical power of a test (https://en.wikipedia.org/wiki/Power_(statistics)#Factors_influencing_power) and may allow researchers to detect a smaller effect size. Especially if we use the difference-in-means estimator, we can decrease the standard error of the average treatment effect by increasing the number of patients. Though we have a budget for the experiment, we can try to recruit as many participants as possible as the budget allows.

Second, increasing the precision of the effect that is measured also influences power. One way to increase precision is to find the optimal allocation of participants to conditions. Salganik (2018 Chapter 4 mathematical note) discussed the mathematical expression of the standard error of the average treatment effect and how it informs researchers to allocate the same number of participants in control and treatment group if and the costs of treatment and control are the same. However, we don't know if the variance for each condition is the same or not in the current experiment and the costs for control and treatment is clearly different here. Therefore, we can estimate the variance of each condition from a pilot study and allocate participants based on the equation:

$\frac{n_1^*}{n_0^*} = \sqrt{\frac{c_0}{c_1}} \frac{\sigma_1}{\sigma_0}$ (List, Sadoff, and Wagner, 2011)

Where c0 and c1 is the cost of applying control and treatment respectively and is the variance of the effect for treatment and control group.

In addition, we can also change the experiment design to increase precision. Salganik suggests that we can "get more precise estimates from a difference-of-differences approach than from a difference-of-means one" (2018 Chapter 4 mathematical note). If we can acquire information of each participant about vaccination uptake in the prior year, we then can estimate the average treatment effect by:

$ATE' = \frac{1}{N} \sum_{i=1}^{N} ((Y_i(1) - X_i) - (Y_i(0) - X_i))$ (Salganik, 2018 Chapter 4)

Since there is a lot of natural variation in how often people take vaccination and that makes comparing treatment and control conditions quite difficult, now if we "difference-out this naturally occurring variability, then there is much less variability, and that makes it easier to detect a small effect." (Salganik, 2018 Chapter 4).

References

List, John A. and Sadoff, Sally and Wagner, Mathis Christoph, So You Want to Run an Experiment, Now What? Some Simple Rules of Thumb for Optimal Experimental Design (January 2010). NBER Working Paper No. w15701.

Salganik, Matthew J., Bit by Bit: Social Research in the Digital Age, Princeton University Press, 2018.

https://en.wikipedia.org/wiki/Power_(statistics)#Factors_influencing_power