

Assignment 5

Sanja Miklin

11/12/2018

1. Experiments on Amazon Mechanical Turk

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

I found a Study from DiCarlo Lab at MIT, titled *Word recognition – report what you see. Up to 50 cent performance based bonus*.

(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.

The task lists a reward of \$0.33, with up to 50 cent performance bonus. If there is more worker-related information available, I cannot find it, or cannot access it as my account hasn't been verified yet.

The Amazon Mechanical Turk Pricing Site, however, gives a bit more information about the requesters side of things. Relevant to this task, there is a Mechanical Turk fee of 20% on the reward and bonus paid to workers.

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

The only qualification requirement is 'HIT approval rate (%) is greater than 90'.

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

According to the description, the HIT is supposed to take 10 minutes or less, and must be completed in under 25 minutes. Assuming 10 minutes is average time, the implied hourly rate is $\$0.33 \times 6 = \1.98 per hour.

(e) When does this job expire?

The job expires on November 17th.

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

If 1 million people participated, it would cost the creator $1.2(\$0.33 \times 1,000,000) = \$396,000$. However, if everyone got the \$0.50 bonus, it might cost the creator up to $1.2(\$0.83 \times 1,000,000) = \$996,000$.

2. Costa and Kahn (2013)

In their paper, Costa and Kahn (2013) explore the effectiveness of energy-conservation ‘nudges’ with respect to individuals’ political leanings and environmentalism. Specifically, their **research question** is: Do household with different political ideologies respond differently to energy-conservation nudges?”

In order to answer their research question, the researchers draw on a few data sources. First, there is the **residential billing data** (for January 2007 to October 2009) for N= 81,722 households, that contains information regarding purchased kilowatt hours per billing cycle of a certain length, whether the household uses electric heat and whether it is enrolled in a renewable energy program (with which there is an associated additional cost). To this, the researchers also added information about the **mean temperature** for each cycle. Second, the researchers purchased **voter registration and marketing data** on individual persons for March 2009 from Aristotle.com. For registered voters, this included party affiliation and whether they donate to environmental organizations. To account for households who could not be matched to this second data-set, the researchers also drew on **census block data** from 2000, noting the share of liberal registered voters and college-educated individuals. Additionally, the researchers are also drawing on a third, much smaller data-set (N= 1,061) based on a **survey of households** receiving the treatment ‘nudge’, which the electric company conducted in 2009.

Within the main data-set, there is N=48,058 households in the **control group**, and N= 33,664 households in the **treatment group**. Households in both groups satisfy the same set of criteria, including 1) electric utility has been active for at least a year, 2) residence is a house (not an apartment building, of area between 250 and 99,998 square feet. Households were grouped in ‘block batches’ of five contiguous census blocks. A ‘batch’ was assigned to treatment randomly, and then a batch contiguous to it was assigned to control until N of households in a group of about 35,000 was reached. The remaining households were assigned to control. The **treatment** was the Home Electricity Report (HER) that select households started receiving between March 14th and May 9th 2008, and continued to receive at the time of the publication of the article. This report contained 1) information on the household’s electric usage over time and 2) relative to their neighbors, as well as 3) energy saving tips.

Through this methodology, Costa and Kahn (2013) build on previous work that sought to evaluate to effectiveness of energy-conservation ‘nudges’. For example, Schultz et al. (2007) looked at the effects of household energy conservation normative messaging with respect to household energy consumption levels prior to the intervention. While prior consumption might correlate with **household environmental ideology**, that is more liberal and environmentalist individuals might already have lower levels consumption, Schultz et al. do not consider the ideological factors in their analysis. Costa and Kahn (2013), in contrast, do gather or approximate the individual-level ideological data and are able to control for this extra level of heterogeneity.

Ultimately, the researchers find that **both the individual household and the census block-group ideological characteristics impact the household’s response to the HER**. On an individual level, being a registered liberal, purchasing electricity from renewable sources and donating to environmental causes are all associated with differential treatment effects, and show a greater decrease in consumption in response to the HER. Furthermore, simply living in a census block group with a higher proportion of liberals and college-educated individuals also affects the treatment response in the same direction, irrespective of individual household characteristics. Furthermore, the researchers also found that the liberals were less likely to opt out of receiving the HER, and were less likely to report their dislike of HER on a survey.

3. Analytical exercise

A new experiment aims to estimate the effect of receiving text message reminders on vaccination uptake. One hundred and fifty clinics, each with 600 eligible patients, are willing to participate. There is a fixed cost of \$100 for each clinic you want to work with, and it costs \$1 for each text message that you want to send. Further, any clinics that you are working with will measure the outcome (whether someone received a vaccination) for free. Assume that you have a budget of \$1,000.

(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?

Clearly, due to the fixed cost-per-clinic, one could enroll more patients in the study if working with fewer clinics. For example, if working with only one clinic, we could use \$900 of the budget towards text messages and could split the patient population so as to estimate not only the effect of receiving a text or not, but also of receiving multiple text-reminders. For example, assuming all 600 patients have cellphones, with 200 patients randomly assigned to a control group, 200 patients assigned to a 1-text condition, and 200 assigned to a 3-text condition, we would spend \$800 on the texts themselves and still be under budget. In contrast, if we want to enroll a 100 individuals in the treatment group, we could spread them out over at most 9 (out of possible 150) clinics.

Our choice of how to spread out our resources depends on a few conditions. First, there are the current vaccination rates and hypothesized impact of the treatment—if the **baseline rates** are either really low or really high, we would need to enroll a larger number of participants in order to evaluate statistical significance, in which case we would want to focus our resources on a small number of clinics. Second, there is the question of the **heterogeneity of treatment effects**. Clinic patient-populations are often bound geographically and might share a variety of characteristics, such as race or SES. If we are interested in (average or differing) effects across populations, we might want to spread the resources out. Third, there is the questions of **dependence**. While randomized trial experiments assume independent observations, patients at a clinic might not be independent of one another (e.g. they might be family members or close neighbors and also remind each other of vaccination after receiving a text). If patients in each clinic are networked together, it would be best to spread our resources as much as possible (and possibly randomly assign treatment and control by clinic matched pairs).

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

The effect size we can reliably detect is limited by the standard error of the average treatment effect. That is, the factors that determine the smallest detectable effect size are the factors that determine the size of the SE. The formula for the SE is as follows:

$$SE(\widehat{ATE}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{N_t} + \frac{\hat{p}(1-\hat{p})}{N_c}}$$

The effect size we can detect, therefore, depends on the 1) size of our control and treatment groups (which are constrained by the budget), and 2) pooled outcome proportion \hat{p} , that is the proportion of patients in both groups that got the vaccine (which is also estimate of the proportion of patients who will get the vaccine if the null hypothesis is true).