Cosette L. Hampton
Assignment #5
MACS 30000, Dr. Evans
Due: Monday, Nov. 12 at 11:30am

1. Experiments on Amazon Mechanical Turk (3 points)

a. Selected the one titled "Learning without instructions" by StanfordNeuroAILab

b. The base pay of $0.50 is the amount in the "Reward" column, however according to the instructions on the preview, you can earn a maximum of a little over $5 based on your learning speed via bonuses. The amount of money you receive after base pay then, is based on the bonus meter. When you respond incorrectly the meter decreases and when you respond correctly it increases. Essentially, this experiment is performance-based and they are hoping that the respondent tries their best and responds quickly because you can get more than $5 if you're ~20 minutes or less, but you won't earn more than $5 if you take ~45 minutes.

c. Qualifications: Respondent must be located in the U.S. and HIT approval rate cannot be less than 90%, while they may have the game elsewhere, in order to complete it here you obviously have to be registered on Amazon Mechanical Turk. Eligibility Requirements: You must be at least 18 years old. Restrictions: You can only complete the experiment with Chrome, you cannot look at the experiment until you accepted the HIT, and you can only complete it one time (refreshing disqualifies you).

d. You have up to 120 minutes to take it, but after 5 minutes the internal task times out. They say it can take as little as 20 minutes if you learn fast or 45-120 if you're a slow learner. We do not know the average amount of time it takes a person to complete the tasks and we don't know how much more than $5 one can make with bonuses, so for the sake of the question I'll use ranges between 20-120 minutes. Let's take the maximum of $5 at 20 minutes, and the minimum at $0.50 at 120 minutes. So, the maximum hourly wage is $15/hour if you learn fast, and $0.25/hour if you learn slow.

e. The job expires in 3 hours! From 11/11/18.

f. Holding constant the maximum reward of $5, if 1 million fast learners participated in this task, the HIT experiment creator would have to pay $5 million. This is unlikely for obvious reasons.

2. Costa and Kahn (2013) (4 points). See "composition" on the next page:

The research question of the Costa and Kahn (2013) paper is the following: **(a)** Does an individual's political ideology influence the effectiveness of "nudges" geared toward energy conservation, and if so, for whom? (Costa & Kahn, p. 4). In the context of this experiment, "nudges" are Home Electricity Reports (HERs) that are sent to households in the treatment group with information about the household's energy consumption in comparison to all neighbors with similar size homes and heat type, households in the bottom 20th percentile of electricity usage, and to one's own household using the previous month and year's consumption (Costa & Kahn, p. 5). Along with the HER, three energy-saving tips and a dollar amount in energy savings is provided. Households receive a message about their consumptions stating "great," "good," or "room for improvement," and green stars for months where they consumed less energy than they did in the same month in the previous year.

In order to observe differences in consumption according to the presence of a nudge or lack thereof and the political ideology of the head of the household, households accessed were randomized into either a treatment and control group. The data about these households came from various different sources. The setting in which the "randomized field experiment" (Costa & Kahn, p. 3) took place is in a utility district in California. **(b)** Residential billing data provided by the unnamed California "electric" utility district, and voter registration and marketing data was purchased by Costa and Kahn (p. 10) from www.aristotle.com. They also used an "ancillary data set" (p. 11) generated by the electric utility company in 2009 to get more information about ideology.

(c) The treatment group are households that received a HER (24,028 monthly or 9,636 quarterly) report since March 14, 2008, and the control group are those who did not receive one (~49,000 homes). At the time the paper was written (April, 2010) HERs were still being sent to homes in the treatment group. Both experimental groups were selected from "85 census tracts with a high density of single-family homes… have a current account with the electric utility that had been active for at least one year, [do not live] in apartment buildings," and have a home with a square foot size between 250-99,998 (Costa & Kahn, p. 6). Continuous census blocks or "block batches" were used for both so that perhaps there would be greater discussion between neighbors who were receiving HERs.

According to Costa and Kahn, they, **(d)** "…control for household and month/year fixed effects, a cubic in mean daily temperature within the billing cycle, and an interaction of the cubic mean daily temperature with a dummy indicator if the house is an electric house…" (Costa & Kahn, p. 14). They also controlled for the "fraction of liberals in the block group," and other housing characteristics (Costa & Kahn, p. 16) in order to help answer their research question beyond Costa and Kahn. Alcott et al. (2014, p. 10) discussed that they controlled for baseline usage, weather, the interaction between daily treatment effects and control group energy usage, and "five respondent characteristics: gender, age, whether homeowner or renter, education and annual income" (Alcott et al., p. 26).

Costa and Kahn found that **(e)** political ideology does have a significant impact on energy consumption and the success of nudges (HERs) in reducing energy consumption compared to those who did not receive HERs, and those who did are more likely to obtain a rebate from their utility bill for purchasing an "energy efficient durable," (pg. 19). Additionally, those who are labelled "environmentalists" or politically liberal are four times as likely to reduce energy usage with nudges than political conservatives (pg. 18).

References:

Alcott, Hunt and Todd Rogers, "The Short-run and Long-run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation," American Economic Review, 2014, 104 (10), 3003–3037.

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.

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

3. Analytical exercise (3 points).

Our "estimand" (Salganik, p. 205) in this case is the effect of receiving text message reminders on vaccination uptake, and using many research subjects can help determine the "average treatment effect," (ATE) specifically because, as Salganik states, we cannot observe both when a single patient does or does not get a vaccination.

At a budget of $1,000, we can engage at minimum two clinics and at maximum 9 clinics, depending on how many patients we want to send text messages to. Some things to consider are the following validity issues, or if experiment results would be different if done in different situations (Salganik, p. 211):

- The ability to randomize within each clinic
- Compliance (internal validity)—did people actually receive the text? Did they lie about getting a vaccine or not?
- The similarity of the population within each clinic and between clinics (socioeconomic status, demographic and household characteristics, etc.)
- The possible outcomes (Salganik, p. 206), like getting a vaccine or not, and when... anything else?
- Any spillover effects that may result from patients attending the same clinic and effects of text message receipt among those in the control group, who don't receive texts but do interact with patients who do.

The control is N – m and the treatment group is "m" according to Salganik, and if the variance between the two are approximate according to the factors above, we want m = N/2. So if N is 90,000, then we want m=45,000. This isn't possible at our budget, so consider N being the size of a clinic at 600, then we want m=300. At that rate we could have 300 treatment group at two

clinics and 100 participants in the treatment group at one clinic and spend all of our budget for 700 participants at three clinics. We consider this the largest size then.

(a) Considering the above discussion, a large (~700) number of participants help with randomization and help create "representativeness" as a sample population, allow us to compare observations better, and help variance decrease so that Var(treatment) ≈ Var(control) and we can keep with the calculations above. (Evans slides, Oct. 29). If the groups at clinics are very different, variance may be high and it would be better to use more clinics, and controlling for interactions in a small sample population may dilute the ATE. There may also be ethics issues with only making texts available to some people at a clinic. So if we assume that everyone at each clinic will need to receive a text, then we only have an unusable $300 left. In this case it may be good to use similar clinics and have one treatment, one control, instead of randomizing within the clinic.

(b) The smallest effect size that I would be able to reliably detect with your budget would be based on some of the same things stated above related to validity and differences in the sample. This is compliance, whether everyone in treatment group receives the text message, whether they receive the *same* text message, whether outcomes are measured among all those treated, say families/friends attend the same clinic so some people in the control group receive the treatment too, or another kind of spillover, selection bias, and of course the level of variance being minimized among control and treatment groups (Salganik, p. 212). All of these things help improve validity, and with the aforementioned things and reducing bias, and controlling heterogenaiety we can get to a small effect size.

Reference:

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