Green // Experimental Research // 2014

# Final Exam

1. Identification of Terms. Briefly define and state the significance of the following terms or phrases.
2. Complete random assignment
3. Trimming bounds vs. extreme value bounds
4. Within-subjects design
5. Mediation
6. Short answer
7. Briefly summarize the implications of clustered random assignment for experimental design and analysis.
8. Explain (preferably using a bit of algebra) why rejecting the null hypothesis that implies rejection of the null hypothesis of homogeneous treatment effects (i.e., ).
9. Modeling and data analysis

The table below shows the results of a recent experiment in which 630,640 subjects were randomly sent a “social pressure” mailing immediately prior to an election in June of 2012. The remaining 33,380 subjects were sent nothing. Turnout in that election is indicated by the variable votedS. This variable equals 1 when a subject voted; 0 otherwise. Later that year, a presidential election occurred, and subjects voted or abstained (see the variable votedG).

Suppose you sought to estimate the “downstream” effect of votedS on votedG.

1. Show algebraically how one can identify the average causal effect among those who vote in the June election if and only if they are encouraged by the mailer.
2. Explain and critically evaluate the excludability assumption required to obtain this identification result.
3. Use the results below to estimate this average causal effect. (Don’t worry about estimating standard errors.)

Subjects assigned to the control group

| votedS

votedG | 0 1 | Total

-----------+----------------------+----------

abstained | 7,990 1,275 | 9,265

| 69.96 5.81 | 27.76

-----------+----------------------+----------

voted | 3,431 20,684 | 24,115

| 30.04 94.19 | 72.24

-----------+----------------------+----------

Total | 11,421 21,959 | 33,380

Subjects assigned to the treatment (mail) group

| votedS

votedG | 0 1 | Total

-----------+----------------------+----------

abstained | 147,147 24,721 | 171,868

| 70.46 5.86 | 27.25

-----------+----------------------+----------

voted | 61,691 397,081 | 458,772

| 29.54 94.14 | 72.75

-----------+----------------------+----------

Total | 208,838 421,802 | 630,640

1. Interpreting results

Suppose that researchers in Kenya, a society where more than 90% of schoolchildren test positive for (communicable) intestinal parasites, conduct a randomized experiment to assess the effect of de-worming medicine on educational attainment. The dependent variable is whether a child later graduates from high school. The treatment is randomly assigned at two levels: first, from a list of 1000 high schools, 100 high schools are randomly selected for an intervention. In the schools chosen to receive the intervention (which vary in size from 50 to 500 pupils), 25 children will be selected to receive de-worming medicine. Assume that the design is implemented as planned, and that outcomes are observed for all children in all schools.

* 1. Define the relevant potential outcomes (using appropriate notation) and the causal estimands of interest. Explain the substantive meaning of each estimand.
  2. Explain whether these estimands are identified given the design and whatever assumptions you see fit to invoke. Be sure to make your assumptions explicit.
  3. What special complications does this design pose for estimation of average causal effects and hypothesis testing?