**Experimental Research: Design, Analysis, and Interpretation**

W4368

Spring 2013

Take-Home Midterm Exam

This exam reviews concepts from FEDAI Chapters 1-7. The questions below are drawn from the chapter exercises, although I have changed the empirical examples.

Work on this assignment independently. If you run into snags, confer with Al or me. In order to make your datasets unique, insert and run the code from subset data.R after reading in the relevant dataset for each question. Use your 4-digit UNI as the random number seed.

Turn in a hard copy of your exam at the start of lecture on Monday, April 22, 2013. Attach a hard copy of the R code you used as an appendix; please do not paste any code in your written solutions. By the start of lecture, send Al, via email, the R program file you used as an attachment. Please include your name in the file name. We should be able to alter the file path, run the entire program straight through, and reproduce your results exactly.

Problem 1. The FEDAI book presents results from a simplified version of the Clingingsmith et al. dataset. The actual dataset has multiple observations in each household, and random assignment of visas occurred at the level of the household (i.e., everyone in the household was entered in the lottery together, and everyone won or lost as a cluster). The data for this problem may be found in q1hajj\_sub.dta. The assignment (success), treatment (hajj2006) and outcome variables (views) are otherwise the same as the dataset used in the book. As noted in Chapter 6, the study encountered two-sided noncompliance. Thus, when analyzing these data, bear in mind the complications of clustered assignment and two-sided noncompliance.

1. Conduct a randomization check using the pre-treatment covariates in the dataset (age, female, literate). Interpret the results.
2. Construct a table or graph to illustrate the ITT.
3. Estimate the ITT with and without covariate adjustment. Provide a 95% confidence interval for the ITT. Test the sharp null hypothesis that the ITT is zero for all subjects.
4. What are the pros and cons of estimating the ITT controlling for the pre-treatment covariates? Does controlling for covariates have a material effect on the results in part [b]? Present a res-res plot of the covariate-adjusted ITT (see Figure 4.3).
5. Because cluster size varies, assess whether the results you found above change when you change the estimator of the ITT from difference-in-means to difference-in-totals.
6. Explain what the term “CACE” means in the context of this study.
7. Explain what the monotonicity assumption means in the context of this study. Assuming monotonicity, estimate the share of Compliers, Never-Takers, and Always-Takers.
8. Estimate the CACE with and without covariate adjustment, and interpret the results.
9. Another version of this dataset may be found in q1hajj.dta; this version includes observations for which outcomes are missing. Use this dataset to calculate extreme value bounds for the ITT.
10. Does missingness appear to be related to treatment assignment? Suppose the missingness rates were identical for the assigned treatment and control groups; under monotonicity, would the upper trimming bound and lower trimming bound be identical?

Problem 2. In 2003, ACORN conducted an experiment in which 5,761 registered voters in Maricopa County were randomly assigned to be canvassed in advance of a municipal election. The data for this problem may be found in q2acorn.dta. Subjects resided in one-voter and two-voter households. Although the assignment was intended to take place at the individual level (see the variable treatment), ACORN decided that it was easier for canvassers to treat everyone in a household if any of its members were assigned to the treatment group. The effective assignment is the variable treat2; ACORN’s procedure amounted to a blocked and clustered assignment, where blocks are defined by household size and clusters are all subjects living at the same address. The variable hhid gives the cluster identifier. The variable persons indicates the number of voters in each household. The outcome variable in this study is voter turnout, vote03. Previous votes, precinct, and age are included as covariates. ACORN’s canvassers only reached some of the subjects they sought to canvass; the variable contact indicates that the treatment was actually administered. When analyzing these data, bear in mind the complications of blocked and clustered assignment as well as one-sided noncompliance.

1. Estimate the probability of treatment assignment for each block.
2. Construct a table showing the relationship between assigned treatment and voter turnout for each size household. Use this table to calculate the ITT for each block, and interpret the results.
3. Pooling both blocks, estimate the overall ITT, and estimate a 95% confidence interval. Test the sharp null hypothesis of no ITT effect. Interpret the results.
4. Because cluster size varies, assess whether the results in part [c] change when you change the estimator of the ITT from difference-in-means to difference-in-totals.
5. Construct a table showing the relationship between assigned treatment and actual treatment for each size household. Use this table to calculate the ITTD for each block, and interpret the results.
6. Define the CACE in this context. Estimate the CACE and its 95% confidence interval. To what extent are these results changed when you control for covariates?
7. Conduct a randomization check in which treat2 is predicted by age, precinct, and voting in previous elections. Interpret the results.