Green // Experimental Research // 2013 **Key –** *Prepared by Al Fang (5/10)*

Final Exam

1. Identification of Terms. Briefly define and state the significance of the following terms or phrases.
   a. Complier Average Causal Effect

      CACE = ATE among Compliers
      CACE under one-sided noncompliance: See Theorem 5.1 (pp. 144, FEDAI)
      CACE under two-sided noncompliance: See Theorem 6.1 (pp. 180, FEDAI)

   b. Trimming bounds vs. extreme value bounds

      Trimming bounds: Given an additional monotonicity assumption, estimate lower and upper bound for the ATE for Always Reporters
      (see pp. 227-8, FEDAI).

      Extreme value bounds: Estimate lower and upper extreme value bounds under assumptions that the missing values in the treatment and control group are imputed using the max (min) and min (max) values of the outcome variable
      (see pp. 226-7, FEDAI).

   c. Non-interference

      The non-interference assumption states that a subject's potential outcomes are unaffected by the treatment assignment of any other subject, or $Y_i(\mathbf{d}) = Y_i(d)$.
      (pp. 253, FEDAI)

   d. Restricted random assignment

      See Box 4.5; pp. 121, FEDAI

2. Short answer

   a. Briefly summarize the implications of clustered random assignment for experimental design and analysis.

      Implications for analysis:
      If cluster sizes are equal, then difference-in-means is an unbiased estimator of the ATE. If cluster sizes are unequal, then the difference-in-means estimator is biased because the denominator is now a random variable, and the ratio of an expectation is not equal to the expectation of a ratio. Use difference-in-totals instead (Eq, 3.24, pp. 83).

      We cannot use standard methods to estimate uncertainty: since the effective N is smaller (due to randomization at the cluster level), the sampling variability generally increases in the variability of the cluster means (pp. 82).

      The true SE of the estimated ATE, assuming fixed cluster means, is described by Equation 3.22 (pp. 82). This quantity cannot be identified because we cannot observe the covariance term, so we estimate standard errors using Equation 3.23 (pp. 83).

      We can form confidence intervals by creating a schedule of potential outcomes under the assumption of constant treatment effects ($\tau_i = ATE_{hat}$); we also apply a degrees of freedom adjustment that expands the width of the interval by the square root of [(k-1)/(k-2)] where k is the number of clusters (pp. 83, including footnote 20).

      Implications for design (based on Eq 3.22 and extensions to the discussion in pp. 57-59):
      We can decrease the SE (i.e. shrink the sampling distribution) by:
      (1) increasing the number of clusters;
      (2) placing greater number of clusters into the treatment group where cluster-level means of potential outcomes associated with that group have a higher variance, but if unknown use a balanced design;
      (3) examine treatments that minimize (or have negative) covariance between average treated potential outcomes and average untreated potential outcomes at the cluster level;
      (4) minimize the cluster level mean treated potential outcome and the cluster level mean untreated potential outcome.

b. Explain (preferably using a bit of algebra) why rejecting the null hypothesis that $\text{Var}(Y_i(1)) = \text{Var}(Y_i(0))$ implies rejection of the null hypothesis of homogeneous treatment effects (i.e., $\text{Var}(\tau_i) = 0$).

See pp. 293, FEDAI.

$$\text{Var}(Y_i(1))$$

$$= \text{Var}(Y_i(0) + \tau_i) \qquad \text{by def of } Y_i(1)$$

$$= \text{Var}(Y_i(0)) + \text{Var}(\tau_i) + 2\,\text{Cov}(Y_i(0), \tau_i) \qquad \text{by properties of variance}$$

and the equality $\text{Var}(Y_i(1)) = \text{Var}(Y_i(0))$ holds when

$$\text{Var}(\tau_i) = -2\,\text{Cov}(Y_i(0), \tau_i) \tag{9.3}$$

Under the null hypothesis that τ_i is constant across subjects, both sides of Equation (9.3) are zero, since the covariance between a variance and a constant is zero. Thus rejecting the hypothesis $\text{Var}(Y_i(1)) = \text{Var}(Y_i(0))$ means rejecting the null hypothesis that $\text{Var}(\tau_i) = 0$.

QED

3. 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 late spring. The remaining 33,380 subjects were sent nothing. Turnout in that election is indicated by the variable `votedS`. 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`. Show algebraically how one can identify the average causal effect among those who vote in the spring election if and only if they are encouraged by the mailer. Explain and critically evaluate the excludability assumption required to obtain this identification result. Use the results below to estimate this average causal effect. (Don’t worry about estimating standard errors.) Use the back of the page, if necessary.

Subjects assigned to the control group

votedG	votedS 0	votedS 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

votedG	votedS 0	votedS 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

The downstream effect of **votedS** on **votedG** may be identified by instrumenting for **votedS** using random assignment to the mailer, or:

Z = random assignment to mailer

D = votedS = voted in spring election

Y = votedG = voted in general election

Assume random assignment of $Z \rightarrow Y \perp Z$, non-interference $Y_i(\mathbf{d}) = Y_i(d)\ \forall\ i \in N$, and that the exclusion restriction holds $Y_i(z,d) = Y_i(d)\ \forall\ z \in Z$ and $\forall\ d \in D$, $\forall\ i \in N$.

Then the estimated downstream effect can be recovered using the Wald estimator (alternatively and equivalently in this case, the IV estimator):

$$\hat{\tau} \quad = \frac{E[Y(Z{=}1)]-E[Y(Z{=}0)]}{E[D(Z{=}1)]-E[D(Z{=}0)]}$$

$$= \frac{\left(\frac{458772}{630640}\right)-\left(\frac{24115}{33380}\right)}{\left(\frac{421802}{630640}\right)-\left(\frac{21959}{33380}\right)}$$

$$= \frac{.7275-.7274}{.6688-.6578} = \frac{.005032}{.010999} = .4575$$

Critically evaluate exclusion restriction assumption: The assumption would be violated if the mailer affected voting in the general election “via” some causal pathway other than voting in the spring. This is likely, for instance, if the mailer was lost in one’s house after receiving it but “found again” just prior to the general election and affected turnout.

4. Interpreting results

Guan and Green (2006) report the results of a canvassing experiment conducted in Beijing on the eve of a local election. Students on the campus of Peking University were randomly assigned to treatment or control groups. Canvassers attempted to contact students in their dorm rooms and encourage them to vote. No contact with the control group was attempted. Of the 2,688 students assigned to the treatment group, 2,380 were contacted. A total of 2,152 students in the treatment group voted; of the 1,334 students assigned to the control group, 892 voted. One aspect of this experiment threatens to violate the exclusion restriction. At every dorm room they visited, even those where no one answered, canvassers left a leaflet encouraging students to vote.

(a) Estimate the ITT.

$$\widehat{ITT} = E[Y_i(Z=1)] - E[Y_i(Z=0)]$$
$$= \left(\frac{2152}{2688}\right) - \left(\frac{892}{1334}\right)$$
$$= .8806 - .6687$$
$$= .1319$$

(b) Assume excludability (i.e., that the leaflet had no effect on turnout). Estimate the CACE.

Let compliance be defined in terms of contact: d=1 if contacted, =0 otherwise. Then:

$$\widehat{CACE} = \frac{\widehat{ITT}}{\widehat{ITT}_D} = \frac{.1319}{E[d_i(z=1)] - E[d_i(z=0)]} = \frac{.1319}{\left(\frac{2380}{2688}\right)} = \frac{.1319}{.8854} = .149$$

(c) Assume that the leaflet raised the probability of voting by one percentage point among both Compliers and Never-Takers. In other words, suppose that the treatment group's turnout rate would have been one percentage point lower had the leaflets not been distributed. Write down a model of the expected turnout rates in the treatment and control groups, incorporating the average effect of the leaflet.

Let τ equal the complier average causal effect of *canvassing* on turnout.

Let α_j denote the proportion of subjects who are in principal stratum *j* where j = {C, NT} and C =compliers, NT = never takers. By definition, $\alpha_C + \alpha_{NT} = 1$, so

$$\alpha_{NT} = 1 - \alpha_C$$

In the original case (under the assumptions that the leaflets do NOT raise the probability of voting by one percentage points among both Compliers and Never Takers), the expected potential outcomes in the treatment and control groups may be written as the following weighted averages:

$$E[Y(Z=0)] = \alpha_C\, E[Y(0)|C] + (1-\alpha_C\, E[Y(0)|NT]$$
$$E[Y(Z=1)] = \alpha_C\, E[Y(0)+\tau\, |C] + (1-\alpha_C\,)E[Y(0)|NT]$$

When we violate the exclusion restriction such that leaflets DO raise the probability of voting by one percentage points among both Compliers and Never Takers, we rewrite this set of equations as:

$$E[Y(Z=0)] = \alpha_C\, E[Y(0)|C] + (1-\alpha_C)\, E[Y(0)|NT]$$
$$E[Y(Z=1)] = \alpha_C\, E[Y(0)+\tau+0.01\, |C] + (1-\alpha_C\,)\, E[Y(0)+ 0.01|NT]$$

(d) Given this assumption about the ATE of leaflets, estimate the CACE of canvassing.

$$ITT = E[Y(Z=1)] - E[Y(Z=0)]$$

$$= \alpha_C\, E[Y(0)+\tau+0.01\, |C] + (1-\alpha_C)E[Y(0)+0.01\, |NT] - \alpha_C E[Y(0)|C]$$
$$-(1-\alpha_C)\, E[Y(0)|NT]$$

$$= \alpha_C\, \tau + 0.01\, \alpha_C + 0.01 - 0.01\alpha_C$$
$$= \alpha_C\, \tau + 0.01$$

Then, solve for τ :

$$\tau = \frac{E[Y(Z=1)] - E[Y(Z=0)] - 0.01}{\alpha_C}$$

$$= \frac{.1319 - .01}{\frac{2380}{2688}} = \frac{.1219}{.8854} = .1377 = CACE$$