Field Experiments: Design, Analysis and Interpretation
Solutions for Chapter 13 Exercises

Alan S. Gerber and Donald P. Green*

Question 1

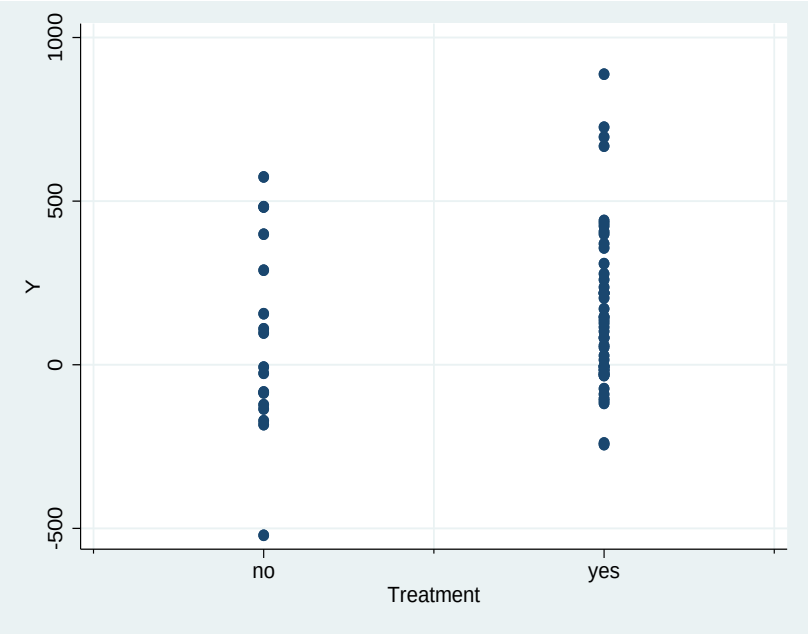
Middleton and Rogers report the results of an experiment in which ballot guides were mailed to randomly assigned precincts in Oregon prior to the 2008 November election. The guides were designed to encourage voters to support certain ballot measures and oppose others. Load the example dataset from `http://isps.research.yale.edu/FEDAI`. The dataset contains election results for 65 precincts, each of which contains approximately 550 voters. The outcome measure is the number of net votes won by the sponsors of the guide across the four ballot measures that they endorsed or opposed. The treatment is scored 0 or 1, depending on whether the precinct was assigned to receive ballot guides. A prognostic covariate is the average share of the vote cast for Democratic candidates in 2006.

a) Estimate the average treatment effect, and illustrate the relationship between treatment and outcomes graphically using an individual values plot.
Answer:

```
In [1]: qui import delim ./data/chapter13/Middleton_Rogers_AI_2010, clear

        rename relevant_measures_net Y
        gen int Z=.
        replace Z = 1 if treatment ==1
        replace Z = 0 if treatment ==0

        qui mean Y if Z==1
        scalar avg_treat = _b[Y]
        qui mean Y if Z==0
        scalar avg_control = _b[Y]
        global tau = avg_treat - avg_control

        disp "average treatment effect = " $tau

average treatment effect = 90.20098


In [2]: twoway (scatter Y Z), ///
        xlabel(0"no" 1 "yes") xmtick(-0.5(1)1.5,grid) xtitle("Treatment")
```

*Solutions prepared by Peter M. Aronow and revised by Alexander Coppock

```
graph export ../results/chapter13/exercise_13_1_a_graph.pdf
```

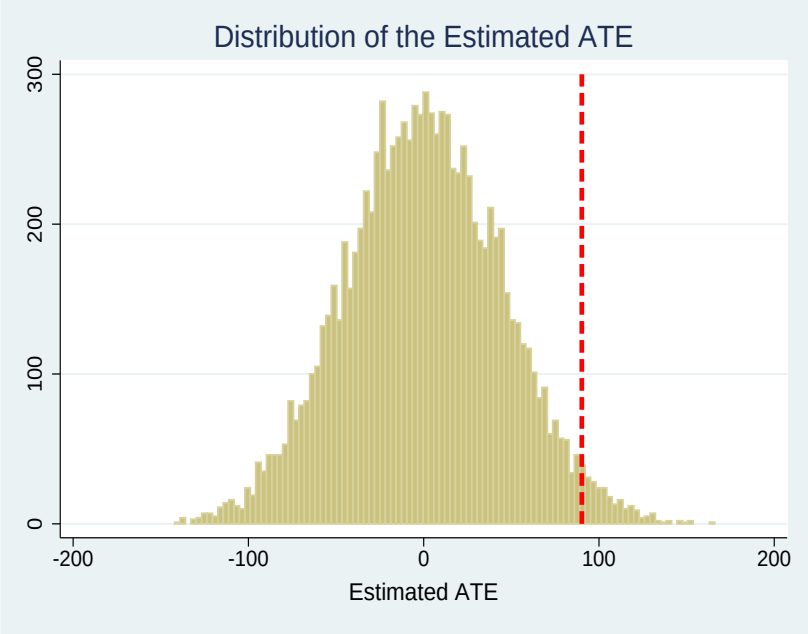


b) Interpret the graph in part (a).
Answer:
The mean of the treatment observations (164) is higher than the mean of the control observations (74), suggesting that the the treatment led to 90 more Democratic votes per precinct. The amount of dispersion around the mean is similar in both groups.

c) Use randomization inference to test whether the apparent difference-in-means could have occurred by chance under the sharp null hypothesis of no treatment effect for any precinct. Interpret the results. Answer:

```
In [3]: ritest Z ate_sim = _b[Z], ///
                reps(10000) sav(13_1_distout.dta, replace) right nodots: ///
                regress Y Z


      Source |       SS           df       MS      Number of obs   =        65
-------------+----------------------------------   F(1, 63)        =      1.50
       Model |  102140.815         1  102140.815   Prob > F        =    0.2248
    Residual |  4281899.43        63  67966.6576   R-squared       =    0.0233
-------------+----------------------------------   Adj R-squared   =    0.0078
       Total |  4384040.25        64  68500.6288   Root MSE        =     260.7

------------------------------------------------------------------------------
           Y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           Z |   90.20098   73.57996     1.23   0.225    -56.83684    237.2388
```

```
       _cons |   73.88235   63.23005     1.17   0.247    -52.47281    200.2375
------------------------------------------------------------------------------

      command:  regress Y Z
      ate_sim:  _b[Z]
  res. var(s):  Z
   Resampling:  Permuting Z
Clust. var(s):  __000003
     Clusters:  65
Strata var(s):  none
       Strata:  1

------------------------------------------------------------------------------
T             |     T(obs)       c       n   p=c/n   SE(p) [95% Conf. Interval]
--------------+---------------------------------------------------------------
     ate_sim |   90.20098    1119   10000  0.1119  0.0032   .105785    .118242
------------------------------------------------------------------------------
Note: Confidence interval is with respect to p=c/n.
Note: c = #{T >= T(obs)}


In [4]: // one-tail p-value
        di %8.4f el(r(p), 1, 1)

  0.1119


In [5]: set more off
        preserve
        use "13_1_distout", clear
        //historgam

        graph twoway (histogram ate_sim,frequency bin(100)) ///
        (scatteri 0 $tau 300 $tau, c(l) lc(red) lw(thick) lp(dash) m(i)), legend(off) ///
        b1title("Estimated ATE") title("Distribution of the Estimated ATE") ///
        xtitle("")

        graph export ../results/chapter13/exercise_13_1_c_graph.pdf

        restore
```



A one-tailed test is appropriate here given that the campaign sought to increase its votes. Randomization inference applied to 10,000 simulated randomizations shows that one-tailed p-value of the estimated ATE is 0.119. This figure is short of the conventional 0.05 threshold.

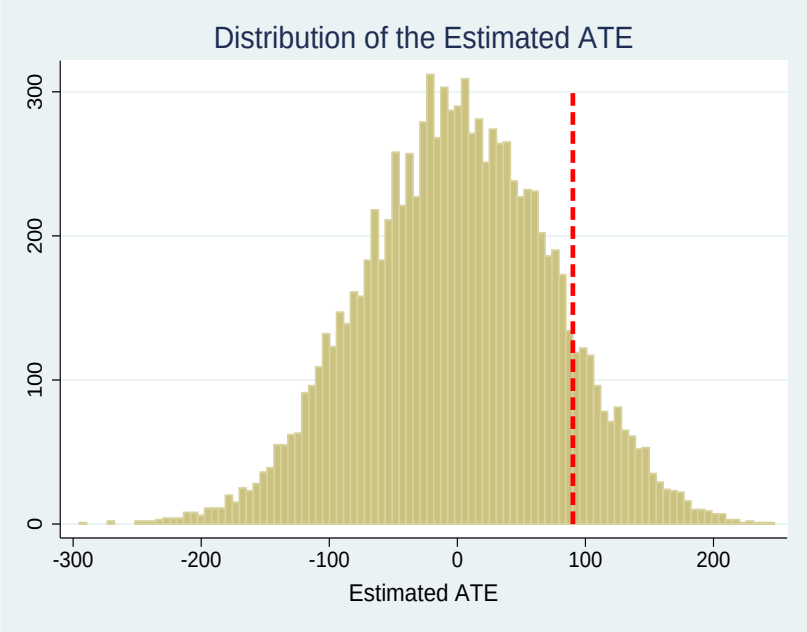
d) Suppose it were the case that when randomly assigning precincts, the authors used the following screening procedure: no random allocation was acceptable unless the average 2006 Democratic support score in the treatment group was within 0.5 percentage points of the average 2006 Democratic support score in the control group. Do all subjects have the same probability of being assigned to the treatment group? If not, re-estimate the ATE, weighting the data as described in Box 4.5. Redo your hypothesis test in part (c) subject to this restriction on the randomization. Interpret the results.
Answer:

```
In [6]: clear
        clear matrix
        clear mata
        set matsize 11000
        set maxvar 32767
        set seed 67887975


        cap matrix drop z
        matrix z=J(65, 10000, .)

In [7]: // restircted RA loop
        qui forvalues i = 1/10000 {
                import delim ./data/chapter13/Middleton_Rogers_AI_2010, clear
                tempvar teststat Z
                gen `Z' = .
```

```
                gen `teststat' = 5
                while (abs(`teststat')>=0.5){
                        tempvar rannum ordering Zri
                    gen `rannum'=uniform()
                        egen `ordering' = rank(`rannum')
                        gen `Zri' = 1 if `ordering' <= 48
                        replace `Zri' = 0 if `ordering' > 48

                        qui reg dem_perf_06 `Zri'
                        replace `teststat' = _b[`Zri']
                }
                replace `Z' = `Zri'
                forvalues j = 1/65 {
                matrix z[`j', `i'] = `Z'[`j']
                }
                drop _all
        }
```

```
In [8]: import delim ./data/chapter13/Middleton_Rogers_AI_2010, clear
         rename relevant_measures_net Y
         gen int Z=.
         replace Z = 1 if treatment ==1
         replace Z = 0 if treatment ==0

In [9]: matrix rowm = z * J(colsof(z), 1, 1/colsof(z))
         matrix colnames rowm=probs
         svmat double rowm, names(col)

In [10]: // distribution of probabilities
         tabstat probs, stat(min p25 med mean p75 max)


    variable |       min       p25       p50      mean       p75       max
-------------+--------------------------------------------------------------
       probs |     .7263     .7341      .737  .7384615     .7425     .7619
------------------------------------------------------------------------------


In [11]: svmat z

In [12]: cap matrix drop tau_dis
         matrix tau_dis=J(10000, 1, .)


         // calculate estimate distribution
         forvalues i = 1/10000{
                 tempvar weight`i'
                 gen `weight`i'' = z`i'/probs + (1 - z`i')/(1 - probs)
```

```
                qui reg Y z`i' [pw=`weight`i'']
                matrix tau_dis[`i', 1] = _b[z`i']

        }

In [13]: set more off

In [14]: preserve
         svmat tau_dis
         qui count if tau_dis1 > $tau
         // one tailed p-value
         di r(N)/_N

         graph twoway (histogram tau_dis1,frequency bin(100)) ///
         (scatteri 0 $tau 300 $tau, c(l) lc(red) lw(thick) lp(dash) m(i)), legend(off) ///
         b1title("Estimated ATE") title("Distribution of the Estimated ATE") ///
         xtitle("")

         graph export ../results/chapter13/exercise_13_1_d_graph.pdf

         restore

.0238
```



Randomization inference applied to 10,000 simulated restricted randomizations shows that one-tailed p-value of the estimated ATE is 0.0238. This figure allows us to reject the null hypothesis at the conventional 0.05 threshold. The p-value here is lower than when we assume unrestricted randomization because re-randomization functions as a form of blocking..

Question 2

Question 3

Conduct your own randomized experiment, based on one of the suggested topics in Appendix B.

a) Compose a planning document.

b) Take an online research ethics course, and obtain your certification to conduct human subjects research. Obtain approval for your study from the institutional review board at your college or university.

c) Conduct a small pilot study to work out any problems in administering the treatment or measuring outcomes.

d) Conduct the experiment. Construct a data file and supporting metadata.

e) Compose a research report.

Answer:
Answers to this question will vary.