Methods III: Replication Exercise 1, Field Experiments

March, 2020

The data for Gerber, Green and Larimer (GGL, 2008) is available on the course website. You should first read through the paper of GGL (2008) to understand the context of the field experiment.

Download the data from the course website and answer the following questions in either R or Stata. You should submit (i) your code, and (ii) a document or PDF containing your answers to jonnyphillips@gmail.com by midnight on Wednesday 25th March. If you get stuck, please feel free to email me, or in the worst case skip that question and continue to the next.

For the replication, don't worry about copying the specific details or formatting of the tables - as long as the results are clear. If you are using R, I encourage you to use R markdown to make it easy to combine your code and answers.

Here is a table of variable names and descriptions from the dataset:

Variable	Description
hh_id	Household Identifier
hh_size	Household Size
block	Block the household is a member of
treatment	Treatment status
g2000	Voted in 2000 General Election
g2002	Voted in 2002 General Election
p2000	Voted in 2000 Primary Election
p2002	Voted in 2002 Primary Election
p2004	Voted in 2004 Primary Election
sex	Sex
age	Age
voted	Voted in 2006 Primary Election

1. What hypotheses are GGL testing? Where did they get these hypotheses from?

2. What are the treatment and control conditions? What is the outcome variable?

3. What is the unit of analysis? What is the level at which treatment is applied?

Next we will evaluate the four assumptions of Field Experiments

4. Assumption (1): What evidence do they provide that random assignment was really implemented?

5. Assumption (2): What evidence do they provide that randomization produced balance on potential outcomes: Answer the questions below.

(a) Assess the balance of individuals' age between the Control and 'Self' treatment. What is the average difference in age between the two groups?

```
##
##  Welch Two Sample t-test
##
```

```
## data:  age by treatment
## t = 0.26042, df = 54632, p-value = 0.7945
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.1373111  0.1793911
## sample estimates:
## mean in group Control    mean in group Self
##              49.81355              49.79251
```

(b) Is this difference in (a) statistically significant using a t-test?

(c) To assess balance in their paper, GGL first aggregate the data to household level averages. Create a new household-level dataset by averaging the household size of each household.

(d) In the household-level dataset, calculate the average Household Size in each control/treatment condition to replicate the first line of Table 1 in GGL (2008).

treatment	hh_size
Civic Duty	1.910805
Control	1.912449
Hawthorne	1.910009
Neighbors	1.910050
Self	1.910900

(e) GGL do not bother to run a t-test, but let's run one between the 'Control' and 'Civic Duty' conditions. Interpret the result.

```
##
##  Welch Two Sample t-test
##
## data:  hh_size by treatment
## t = -0.29175, df = 28348, p-value = 0.7705
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.012693818  0.009404488
## sample estimates:
## mean in group Civic Duty    mean in group Control
##                 1.910804                 1.912449
```

6. Assumption (3): What is the risk of spillovers (violations of SUTVA) from treatment?

7. Assumption (4): Is there a risk of any 'parallel' treatments that violate the excludability assumption?

Now let's look at the results of the experiment.

8. Returning to the individual level data, perform a simple difference-in-means t-test for voter turnout between the 'Control' and 'Neighbors' groups in the individual data. Interpret the result and compare them to the findings in Table 2 of GGL.

```
##
##  Welch Two Sample t-test
##
## data:  voted by treatment
## t = -30.207, df = 52613, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.08658577 -0.07603405
```

```
## sample estimates:
##   mean in group Control mean in group Neighbors
##              0.2966383               0.3779482
```

9. GGL also run an OLS regression to understand the effect of each treatment on voter turnout. Table 3, column (a), shows the simple regression of treatment on turnout with no adjustments or controls. Run this regression and compare it to the findings in GGL. (Note: The authors include a series of dummies for each treatment condition, but this is equivalent to including the treatment variable as a factor variable with 'Control' as the first (baseline) level.) Interpret the results.

Table 2:

	Dependent variable:
	voted
treatmentCivic Duty	0.017899***
	(0.002600)
treatmentHawthorne	0.025736***
	(0.002601)
treatmentSelf	0.048513***
	(0.002600)
treatmentNeighbors	0.081310***
	(0.002601)
Constant	0.296638***
	(0.001061)
Observations	344,084
Note:	*p<0.1; **p<0.05; ***p<0.01

10. The experimental design included blocking: randomization was conducted *within* blocks defined by 'cells' of geographically close together houses.

a. What is the gain the authors obtain from 'blocking' their randomization on local neighbourhood?

b. The blocking also allows us to add the blocks as fixed effects in the regression so we are comparing only households that live within the same cell. However, there are many (10,000) blocks so your computer probably doesn't have enough memory to run this regression directly. An equivalent methodology is to remove the between-block variation in the treatment variable manually and then run the same regression as in Q9. To do this for the 'Neighbors' treatment:
1. Create a dummy variable (1/0) for individuals that received the 'Neighbors' treatment,
2. Remove the other treatments from your dataset (so you are left with just 'Neighbors' and 'Control' units),
3. For each 'block' calculate the mean value of the binary 'Neighbors' treatment variable,
4. Subtract the block mean from the individual values of the Neighbors treatment variable,
5. Run the same regression as in Q9 but using the block-mean-centered treatment variable you just created as the explanatory variable.

We can now replicate the results in Column (b) of Table 3. How do the results change compared to your answer to Q9? How does this change the comparisons we are making in the regression?

Table 3:

	Dependent variable:
	voted
Neighbors	0.082163^{***}
	(0.002599)
Constant	0.305652^{***}
	(0.000974)
Observations	229,444
Note:	*p<0.1; **p<0.05; ***p<0.01

11. In column (c) of Table 3, the authors add covariates (g2000,g2002,p2000,p2002,p2004) to the regression. How do the results change when we add covariates to your regression from Q10 (you may need to see lots of decimal places to see the difference)?

Table 4:

	Dependent variable:
	voted
Neighbors	0.081543^{***}
	(0.002504)
g2000	-0.004821^{*}
	(0.002822)
g2002	0.098813^{***}
	(0.002650)
p2000	0.098989^{***}
	(0.002180)
p2002	0.134248^{***}
	(0.001977)
p2004	0.155064^{***}
	(0.001906)
Constant	0.090243^{***}
	(0.002698)
Observations	229,444
Note:	*p<0.1; **p<0.05; ***p<0.01

12. Another feature of the randomized experiment is that treatment was randomized at the household, not the individual, level. So everyone in the same household has the same treatment status. This is clustering. Repeat your regression from Q9 (no blocking or control variables) for the comparison between the , but this time with standard errors clustered at the household

level. What difference does this make? Why do we do this?

term	estimate	std.error	p.value
(Intercept)	0.2966	0.0017	0
treatmentCivic Duty	0.0179	0.0031	0
treatmentHawthorne	0.0257	0.0031	0
treatmentSelf	0.0485	0.0032	0
treatmentNeighbors	0.0813	0.0033	0

13. GGL are being a bit lazy - their outcome variable is binary but they still use an OLS regression. Use a logit regression model to run the same model as in Q9. How would you interpret the results?

Table 5:

	Dependent variable:
	voted
treatmentCivic Duty	0.084^{***}
	(0.012)
treatmentHawthorne	0.120^{***}
	(0.012)
treatmentSelf	0.223^{***}
	(0.012)
treatmentNeighbors	0.365^{***}
	(0.012)
Constant	-0.863^{***}
	(0.005)
Observations	344,084
Note:	*p<0.1; **p<0.05; ***p<0.01

14. What is the population GGL estimate treatment effects for? (see p.36-37 under 'Study Population')

15. How generalizable are the findings of this study to the primary elections in Michigan four years later in 2010? To neighbouring Indiana in 2006? To elections in Brazil?