Assignment 3 Notebook

Shu Wang

Please read this code before starting. There are several ways to run regressions in R. I'd like us to use the function 'feols' with the option se = 'hetero'. We will discuss this in class.

```
# Code that shows how different regression functions work:
this_reg_lm <- lm(racism.scores.post.2mon ~ any_treatment,
                  data = tweets_data)

this_reg_feols <- feols(racism.scores.post.2mon ~ any_treatment,
                        data = tweets_data)

this_reg_feols_robust <- feols(racism.scores.post.2mon ~ any_treatment,
                               data = tweets_data, se = 'hetero')

# this_reg_feols_inter <- feols(racism.scores.post.2mon ~ any_treatment*anonymity,
#                                 data = tweets_data)

#
# modelsummary(list(this_reg_lm, this_reg_feols, this_reg_feols_robust),
#               stars = TRUE, gof_omit = "AIC|BIC|Log|F|RMSE|Adj")
#
# modelsummary(list(this_reg_lm, this_reg_feols, this_reg_feols_robust),
#               coef_map = c('any_treatment' = 'Treatment',
#                 'anonymity' = 'Anonymity',
#                 'any_treatment:anonymity' =
#.                 'Treatment * Anon.',
#                 '(Intercept)' = 'Constant'),
#                 stars = TRUE,
#                 gof_omit = "AIC|BIC|Log|F|RMSE|Adj")
```

1.a: Of the above variables, please identify all that may be 'good' control variables in a regression where the outcome is 'racism.scores.post.2mon' and the regressor is 'any_treatment'?

Note: By good control variables, I mean those which do not prevent a causal interpretation of the coefficient on treatment.

Based on the given information,'anonymity' and 'log_followers' are possible control variables that is related to the regression. 'racism.score.pre.2mon' may also affect the racism score after the treatment since it records the original case of racism. So these 3 variables are related to the regression.

1.b: Run a regression of 'racism.scores.post.2mon' on 'any_treatment'.

```
reg_2 <- feols(racism.scores.post.2mon ~ any_treatment,
                              data = tweets_data, se = 'hetero')
etable(reg_2)
```

```
                                 reg_2
Dependent Var.: racism.scores.post.2mon

Constant             0.2522*** (0.0634)
any_treatment          -0.0828 (0.0687)
_______________ _______________________
S.E. type       Heteroskedasticity-rob.
Observations                        243
R2                              0.00773
Adj. R2                         0.00361
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

What do we learn from this regression about the effect of the treatment? Please explain in words in addition to just returning the number.

Based on our settings,we set the variance term to be heteroskedasticity,which means the standard error is biased instead of constant.

The model give a coefficient of -0.0827.For 1 unit increase in any_treatment,we are expected to see 0.0827 unit decrease in racism_score_2mo. However,based on the summary of this model,the regressor (any_treatment) is not significant at 5% significance level. Hence, we fail to reject the null hypothesis that the coefficient of (any_treatment) is different from 0 at 5% significance level.The intercept is significant at 5% significance level.

The adj-R^2 value is relatively low,indicating that only 0.3% of the variation in racism_score is explained by the predictor variable any_treatment.The R^2 value is low too.

Below is an example regression and how you should output it. In this regression, our outcome variable is 'anonymity' and our explanatory variable is 'log.followers'. This regression tells us whether there is correlation between anonymous twitter accounts and those who have a lot of followers (in our dataset).

```
reg_anon_followers <- feols(anonymity ~ log.followers, data = tweets_data, se = 'hetero')
etable(reg_anon_followers)
```

Repeat the above exercise, but to answer question 1.b.

1.c Add the variables from a) as controls into the regression from b). What happens to our estimate of the effect of the treatment and its standard error? Why does this happen in words?

```
reg_c <- feols(racism.scores.post.2mon ~ anonymity + log.followers  + racism.scores.pre.2m
                               data = tweets_data, se = 'hetero')
```

```
NOTE: 1 observation removed because of NA values (RHS: 1).
```

```
etable(reg_c)
```

```
                                         reg_c
Dependent Var.:        racism.scores.post.2mon

Constant                       0.0568 (0.1002)
anonymity                      0.0181 (0.0225)
log.followers                 -0.0014 (0.0163)
racism.scores.pre.2mon      0.7659*** (0.1287)
any_treatment                  0.0221 (0.0457)
______________________ _______________________
S.E. type              Heteroskedasticity-rob.
Observations                               242
R2                                     0.27703
```

```
Adj. R2                                   0.26482
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The coefficient of the treatment changes from negative to positive.This indicates that any_treatment would increase the racism_score.The standard error of the any_treatment decreases.

Possible explanation is that by adding variables to the regression model,they are interacting with each other hence the regression model would adjust the coefficient accordingly. The standard error would also change.

1.d Use a regression to check for differences between the treatment and control for one of the variables identified in a). Also use the prop.test function to check whether the randomization proportion was intended. Based on these results, should we be concerned that the randomization was done improperly?

```r
reg_d <- feols(racism.scores.pre.2mon ~ any_treatment + anonymity,
                                 data = tweets_data, se = 'hetero')
```

```
NOTE: 1 observation removed because of NA values (LHS: 1).
```

```r
etable(reg_d)
```

```
                                  reg_d
Dependent Var.: racism.scores.pre.2mon

Constant              0.1729** (0.0531)
any_treatment        -0.1328. (0.0704)
anonymity             0.0343* (0.0161)
_______________ ______________________
S.E. type       Heteroskedastici.-rob.
Observations                       242
R2                             0.05110
Adj. R2                        0.04317
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Based on the results,the coefficient of any_treatment is significant at 10% significance level and the coefficient of anonymity is significant at 5% significant level. This suggests that there's difference between pre-treatment score on racism for different treatment groups.On average,people who receive treatment has a 0.1328 racism_score lower than those who don't receive treatment before the treatment.

```
reg_d2 <- feols(racism.scores.pre.2mon ~ any_treatment,
                                 data = tweets_data, se = 'hetero')
```

```
NOTE: 1 observation removed because of NA values (LHS: 1).
```

```
etable(reg_d2)
```

```
                               reg_d2
Dependent Var.: racism.scores.pre.2mon

Constant              0.2277** (0.0707)
any_treatment        -0.1349. (0.0713)
_______________ ______________________
S.E. type       Heteroskedastici.-rob.
Observations                       242
R2                             0.04342
Adj. R2                        0.03943
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

For simple regression,the coefficient is significant at 10% significance level.

(Hint: we can do this by comparing the pre-treatment outcomes between the treatment and control group. If there are significant differences, then there may be a problem with the experiment.Note: Each arm of the experiment was assigned with equal probability and there are 4 treatment arms and one control.)

```
library(dplyr)
tweets_data%>% count(any_treatment)
```

```
   any_treatment   n
1:             0  52
2:             1 191
```

```r
prop.test(191,243,0.8)
```

```
	1-sample proportions test with continuity correction

data:  191 out of 243, null probability 0.8
X-squared = 0.21631, df = 1, p-value = 0.6419
alternative hypothesis: true p is not equal to 0.8
95 percent confidence interval:
 0.7280013 0.8347620
sample estimates:
        p 
0.7860082 
```

Our proportion test shows that 95% confidence interval includes our proportion,0.786. Hence,at 5% significance level,we fail to reject the null hypothesis that the treatment group is randomized with a proportion of 0.8.

Combine these results together,we can say that there's no significant differences in treatment group and control group before the experiment and randomization is appropriate.

1.e We would like to know whether treatment arm 2 or treatment arm 3 is statistically significantly better at reducing racist behavior. Perform a t.test or regression and test for the null hypothesis that treatment arm 2 has the same effect as treatment arm 3.

```r
arm2_treatment <- tweets_data %>%
  filter(treatment_arm == 2) %>%
  pull(racism.scores.post.2mon)
arm3_treatment <- tweets_data %>%
  filter(treatment_arm == 3) %>%
  pull(racism.scores.post.2mon)
# Compute t-test
res <- t.test(arm2_treatment, arm3_treatment) # assume equal variance
res
```

```
	Welch Two Sample t-test

data:  arm2_treatment and arm3_treatment
```

```
t = 1.6246, df = 64.539, p-value = 0.1091
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.0151519  0.1472047
sample estimates:
 mean of x  mean of y
0.13892962 0.07290323
```

Let μ_1 denote the mean of racism_score_2mon for treatment 2. Let μ_2 denote the mean of racism_score_2mon for treatment 3.

Null hypothesis: $\mu_1 = \mu_2$ Alternative hypothesis: $\mu_1 \neq \mu 2$.

Based on our above results, the p-value is 0.1. The result is not significant at 5% significance level,hence we fail to reject the null hypothesis that $\mu_1 = \mu_2$.

2.a Describe the treatment in the first experiment and the unit of randomization. What share was randomized to the treatment?

(This refers to the experiment conducted in August 2015, the first experiment described in the introduction of the paper.)

The unit of randomization is consumer/user on Stubhub.The experiment was designed on a browser/cookie level. The treatment is that treated consumers were initially shown Back-End fees(ticket prices without fees while fees were added in check-out page).Consumers assigned to control group were shown UP-Front fees(conspicuous onsite announcements confirming that the prices they saw upfront included all charges and fees)

The odds of assignment to the treatment group are 50.11% initially.

Based on the table,after we perform two glitches,50.06% of consumers is assigned to treatment group(showing Black-End fees).

2.b Table II displays a randomization / balance check. A randomization check is a regression where the dependent variable occurs before the experiment. It should be very unlikely that there are substantial differences in before experiment variables if the experiment was done properly. Suggest a variable not used by the authors that would be appropriate to include in a balance check.

Geographic location of users. Users in different zones(West,middle,East) may have different shopping habits. This can be done by tracking the IP address of users.

Also,type of device they are using may also affect their behaviors.

2.c What is the effect of the treatment on the Propensity to Purchase at least one product? Calculate the 95% confidence interval for this estimate.

Table III shows that the treatment group shows a 14.1% difference in Propensity to Purchase at least one product compared to control group.

The 95% CI is constructed as: $[0.141 - z_{0.975} * se, 0.141 + z_{0.975} * se]$

```
z_score = 1.96
mean_difference = 0.141
sd_error = 0.0009

Ci_lower = mean_difference - z_score*sd_error
Ci_upper = mean_difference + z_score*sd_error

Ci_lower
```

```
[1] 0.139236
```

```
Ci_upper
```

```
[1] 0.142764
```

The 95% confidence interval is:

$[0.139236, 0.142764]$

2.d Suppose the authors randomized by city of the event. Name one benefit that may occur as a result of this randomization strategy and one harm.

Benefit: Clustered randomization is useful when evaluating interventions delivered at city level.Also,clustered randomization has lower implementation costs/administrative convenience.

Harm: There is recruitment bias in clustered randomization. If the city has different market situation from others, then the treatment might be affected by this result. This may cause unequal proportion of users in the randomization step as the population of users in each city may vary from each other.

2.e Suppose that you are the product manager for the monetization team at Stubhub. Based on the evidence presented above, would you launch the treatment to the entire site? The answer should be 1 paragraph. It should consist of an answer (Yes, no), and two pieces of evidence relating to that recommendation. Case participation will also constitute part of this grade.

Np,I would not recommend launch the treatment to the entire site. Based on the result, although there's 14.1% increase in propensity to purchase at least once, the result is not ideal conditioning on purchasing. The total seat sold,12-month churn rate,transaction within 10 days all indicates that in those who make purchases,Back-End users have a lower rate compared to Upfront users.This is not a good sign as churn rate and attendance rate also plays an important role when evaluating customers. In order to make profit,it's not enough to just look at revenue.We should also care about long-term profit.

How long did this assignment take you to do (hours)? How hard was it (easy, reasonable, hard, too hard)?

reasonable.