Econometrics II - Assignment 4

Uncensored sloths

27 Jan 2022

Question 1

a)

```
((0.7*0.4+0.3*0.6)-(0.4*0.2+0.6*0.5))/(0.7-0.4)
```

```
## [1] 0.2666667
```

b) Given that you were in prison, the probability to get arrested again increases by 30%. Assuming that there are no defiers (which is necessary to make any inferences on the estimation), it is the causal effect of compliers (people with on average the same characteristic but prisoned by Jones and not prisoned by Smith). The estimator concerns 30% of the population assuming again there are no defiers. Including defiers makes it impossible to estimate the fraction. We also assume that both judges get 50% of the cases per person and that due to randomization the types are equally distributed among the judges.

Think about slide 18 what happens if we include defiers!

c) People that are sentenced to prison by both judges. Hence, for Smith they make out 40% of the sample as compliers and never takers will not be sentenced to prison Smith. For Jones, they are part of the 70% together with the compliers as only the never takers are never sentenced to prison with Jones as a judge. They make out 40% of the sample assuming monotonicity (so there are no people who would not be sentenced to prison by Jones, while being sentenced to prison by Smith).

Question 2

a)

```
MDE_target <- 0.1
t_q = -0.524
t_alpha <- 1.960
p = 0.5
sigma = (1-0.5)*0.5
n_target = ((t_alpha - t_q)/MDE_target)^2*((sigma)/(p*(1-p)))
n_target
```

```
## [1] 617.0256
```

b)

```
rt <- 0.8
rc <- 0
n_complience = ((t_alpha - t_q)/(MDE_target*(rt - rc)))^2*((sigma)/(p*(1-p)))
n_complience
```

```
## [1] 964.1025
```

Question 3

```
# Load data
data <- read.csv("assignment4.csv")
```

a) Compute for the children assigned to the control group the variance in flu incidence. If the researcher aims at reducing flu incidence by 0.05, how many children should participate in the randomized experiment.

```
MDE_target <- -0.05
t_q <- -0.524
t_alpha <- 1.960
p <- 0.8
flu_control <- mean(data[data$treatgroup == 0, 'flu'])
sigma = (1-flu_control)*flu_control
n_flu = ((t_alpha - t_q)/MDE_target)^2*((sigma)/(p*(1-p)))
n_flu
```

```
## [1] 3631.72
```

b) Compute which fraction of the children in the treatment group actually received a flu shot. What is the implication for the power analysis of the experiment?

```
rt <- mean(data[data$treatgroup == 1, 'flushot'])
rt
```

```
## [1] 0.6679552
```

```
rc <- mean(data[data$treatgroup == 0, 'flushot'])
rc
```

```
## [1] 0
```

```
sigma = (1-flu_control)*flu_control
n_flu = ((t_alpha - t_q)/(MDE_target*(rt - rc)))^2*((sigma)/(p*(1-p)))
n_flu
```

```
## [1] 8139.875
```

c)

```
data$treatedT <- ifelse(c(data$treatgroup == 1 & data$flushot==1), 1, 0)
data$untreatedT <- ifelse(c(data$treatgroup == 1 & data$flushot==0), 1, 0)
data$treatment <- ifelse(data$treatedT== 1, 1, ifelse(data$untreatedT== 1, 2, 0))
```

```
balance <- balance_table(data[, !names(data) %in% c("treatgroup", "flushot", "treatedT", "untreatedT")]
balance
```

```
## # A tibble: 7 x 6
##   variables1  Media_control1 Media_trat1 Media_trat2 p_value1 p_value2
##   <chr>                <dbl>       <dbl>       <dbl>    <dbl>    <dbl>
## 1 agemother           26.1        26.6        24.9   1.22e-12 2.46e-52
## 2 educmother          12.3        12.5        11.8   5.86e- 6 1.31e-28
## 3 flu                  0.621       0.401       0.675 1.54e-79 1.83e- 5
## 4 genderchild          0.508       0.503       0.501 6.80e- 1 6.22e- 1
## 5 housincome        2270.       2374.       2111.    1.45e- 5 4.90e-10
## 6 married              0.957       0.977       0.939 1.83e- 5 1.24e- 3
## 7 nationality          0.278       0.239       0.341 1.33e- 4 3.00e- 7
```

d) Estimate this model using OLS. Next, include subsequently the individual characteristics. What do you learn from these regressions?

```
model1_robust <- rlm(flu ~ flushot, data = subset(data, data$treatgroup == 1))
model2_robust <- rlm(flu ~ flushot + genderchild + nationality + agemother + educmother + married + hous
stargazer(model1_robust, model2_robust)
```

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Do, Jan 27, 2022 - 13:04:49

Table 1:

	Dependent variable:	
	flu	
	(1)	(2)
flushot	−0.274***	−0.164***
	(0.010)	(0.010)
genderchild		0.016*
		(0.009)
nationality		0.092***
		(0.010)
agemother		−0.048***
		(0.002)
educmother		−0.030***
		(0.003)
married		−0.024
		(0.025)
housincome		0.00001
		(0.00000)
Constant	0.675***	2.200***
	(0.008)	(0.045)
Observations	10,089	10,089
Residual Std. Error	0.594 (df = 10087)	0.602 (df = 10081)
Note:		*p<0.1; **p<0.05; ***p<0.01

e) Use 2SLS to estimate β_1 and check the robustness with respect to adding individual characteristic

```
model3 <- feols(flu ~ 1 | flushot ~ treatgroup, data)
summary(model3)
```

```
## TSLS estimation, Dep. Var.: flu, Endo.: flushot, Instr.: treatgroup
## Second stage: Dep. Var.: flu
## Observations: 12,583
## Standard-errors: IID
##              Estimate Std. Error  t value  Pr(>|t|)
```

```
## (Intercept)  0.620690   0.009705  63.9530 < 2.2e-16 ***
## fit_flushot -0.192779   0.016227 -11.8802 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## RMSE: 0.484649   Adj. R2: 0.059242
## F-test (1st stage), flushot: stat = 5,016.2, p < 2.2e-16 , on 1 and 12,581 DoF.
##                  Wu-Hausman: stat =    17.1, p = 2.237e-5, on 1 and 12,580 DoF.
```

```
model4 <- feols(flu ~ genderchild + nationality + agemother + educmother + married + housincome | flushe
summary(model4)
```

```
## TSLS estimation, Dep. Var.: flu, Endo.: flushot, Instr.: treatgroup
## Second stage: Dep. Var.: flu
## Observations: 12,583
## Standard-errors: IID
##                Estimate Std. Error    t value  Pr(>|t|)
## (Intercept)  2.15266853 0.04010533  53.675370 < 2.2e-16 ***
## fit_flushot -0.19827462 0.01508642 -13.142589 < 2.2e-16 ***
## genderchild  0.01342733 0.00804823   1.668358   0.09527 .
## nationality  0.09024013 0.00913374   9.879864 < 2.2e-16 ***
## agemother   -0.04628152 0.00163287 -28.343613 < 2.2e-16 ***
## educmother  -0.02731668 0.00278890  -9.794798 < 2.2e-16 ***
## married     -0.02848121 0.02181911  -1.305333   0.19180
## housincome   0.00000357 0.00000423   0.844801   0.39824
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## RMSE: 0.451201   Adj. R2: 0.184224
## F-test (1st stage), flushot: stat = 5,400.9   , p < 2.2e-16 , on 1 and 12,575 DoF.
##                  Wu-Hausman: stat =     3.5763, p = 0.058633, on 1 and 12,574 DoF.
```

f) Estimate the first-stage regression using OLS. Are you afraid of a weak instruments problem?

```
first_stage <- rlm(flushot ~ treatgroup, data = data)
```

```
## Warning in rlm.default(x, y, weights, method = method, wt.method = wt.method, :
## 'rlm' failed to converge in 20 steps
```

```
stargazer(first_stage)
```

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Do, Jan 27, 2022 - 13:04:49

Also note that first stafe F test in the previous estimation is above 10!

Table 2:

	Dependent variable:
	flushot
treatgroup	0.706^{***}
	(0.011)
Constant	-0.000
	(0.010)
Observations	12,583
Residual Std. Error	0.439 (df = 12581)
Note:	*p<0.1; **p<0.05; ***p<0.01