Homework 9

Gergel Anastasia

7/1/2021

Exercise 1

(a) Probit model estimation

```
library(haven)
library(dplyr)
library(stargazer)
library(clubSandwich)

data <- read_dta("/Users/herrhellana/Dropbox/_NYU studies/Quant I/home assignments/HW9/cces.dta")

# recode race to white
data <- data %>% mutate(white = as.numeric(race==1))

# recode educ
data <- data %>% mutate(educf = as.factor(educ))
data <- data %>% mutate(urbancity = as.factor(urbancity))

# recode CC18_320d
data <- data %>% mutate(gun = case_when(CC18_320d == 2 ~ 0, # against
                                        CC18_320d == 1 ~ 1)) # for

unique(data$gender) # 8 and 9 values are not in the dataset
```

```
<labelled<double>[2]>: Gender
[1] 2 1

Labels:
 value     label
     1      Male
     2    Female
     8   skipped
     9 not asked
```

```
data <- data %>% mutate(female = case_when(gender == 1 ~ 0,
                                           gender == 2 ~ 1))


probit_1 <- glm(data=data, gun ~ white + educf
                + white:educf + urbancity + female,
                family = binomial(link = 'probit'))
vp_i <- vcovCR(probit_1, cluster = data$inputstate, type = 'CR1')
```

```r
stargazer(probit_1, type = 'text',
          covariate.labels = c("white", "high school", "some college",
                               "2-year college", "4-year college", "post-grad",
                               "suburb", "town", "rural", "other urban",
                               "female", "white:high school", "white:some college",
                               "white:2-year college", "white:4-year college",
                               "white:post-grad", "intercept"),
          se = list(sqrt(diag(vp_i))))
```

```
================================================
                         Dependent variable:
                     ---------------------------
                                 gun
------------------------------------------------
white                          0.119*
                              (0.061)

high school                   -0.143**
                              (0.061)

some college                  -0.268***
                              (0.067)

2-year college                -0.212***
                              (0.081)

4-year college                -0.326***
                              (0.065)

post-grad                     -0.375***
                              (0.067)

suburb                        0.084***
                              (0.021)

town                          0.222***
                              (0.025)

rural                         0.411***
                              (0.024)

other urban                    0.143*
                              (0.075)

female                        -0.465***
                              (0.009)

white:high school               0.024
                              (0.057)

white:some college             0.169**
                              (0.072)
```

```
white:2-year college           0.154*
                               (0.082)

white:4-year college            0.032
                               (0.081)

white:post-grad                -0.050
                               (0.073)

intercept                     -0.180***
                               (0.059)

------------------------------------------------
Observations                   59,545
Log Likelihood               -36,664.960
Akaike Inf. Crit.            73,363.920
================================================
Note:                *p<0.1; **p<0.05; ***p<0.01
```

(b) **Interpretation of the results**

There is a strong negative and significant association between the predictors of support for concealed carry permits and education (and strength of the association increases with an increase in education years), between the predictors and being a woman. The significant positive association exists between the predictors and being a white person, as well between the predictors and not living in a city. Two interaction terms are significant: having some college (`some college` and 2 years of college) increases the association between the predictors and being white.

(c) $Pr(gun = 1|x_1) = \Phi(-0.180+0.084\cdot1-0.375\cdot1) = \Phi(-0.471) = 0.319$ (for an x_1, who is a non-white man w a post-grad degree living in a suburb). $Pr(gun = 1|x_2) = \Phi(-0.180+0.119-0.375+0.084-0.050) = \Phi(-0.402) = 0.344$ (for an x_2, who is a white man otherwise similar to the x_1).

(d)

```r
library(margins)
data_fem_0 <- data %>% mutate(female = 0)
data_fem_1 <- data %>% mutate(female = 1)

y_hat_0 <- predict(probit_1, newdata = data_fem_0, type = 'response')
y_hat_1 <- predict(probit_1, newdata = data_fem_1, type = 'response')

mean(y_hat_0, na.rm = T) # 0.44 for men
```

```
[1] 0.4403996
```

```r
mean(y_hat_1, na.rm = T) # 0.27 for women
```

```
[1] 0.2732113
```

```r
(mean(y_hat_0, na.rm = T) - mean(y_hat_1, na.rm = T)) # difference
```

```
[1] 0.1671882
```

```r
# using the package 'margins'
prediction(probit_1, at = list(female = unique(data$female)), vcov = vp_i) %>% summary()
```

```
 at(female) Prediction       SE     z p  lower  upper
          0     0.4404 0.007789 56.54 0 0.4251 0.4556
          1     0.2732 0.006540 41.77 0 0.2604 0.2860
```

Estimated difference in probability	estimate
white	0.062
size of place	0.120
education for whites	0.062
education for non-whites	0.061

And here we also see the CIs for both values of the average predictive margins of the female variable.

(e)

```
p_white <- prediction(probit_1, at = list(white = unique(data$white)), vcov = vp_i) %>% summary()
p_place <- prediction(probit_1, at = list(urbancity = c('2', '4')), vcov = vp_i) %>% summary()

p_educ <- prediction(probit_1, at = list(white = unique(data$white),
                                         educf = c('2', '5')), vcov = vp_i) %>% summary()

# differences:
p_white$Prediction[2] - p_white$Prediction[1] # white
```

```
[1] 0.06218036
```

```
p_place$Prediction[2] - p_place$Prediction[1] # place
```

```
[1] 0.1202043
```

```
p_educ$Prediction[4] - p_educ$Prediction[2] # white educ
```

```
[1] -0.06253926
```

```
p_educ$Prediction[3] - p_educ$Prediction[1] # non-white educ
```

```
[1] -0.06147606
```

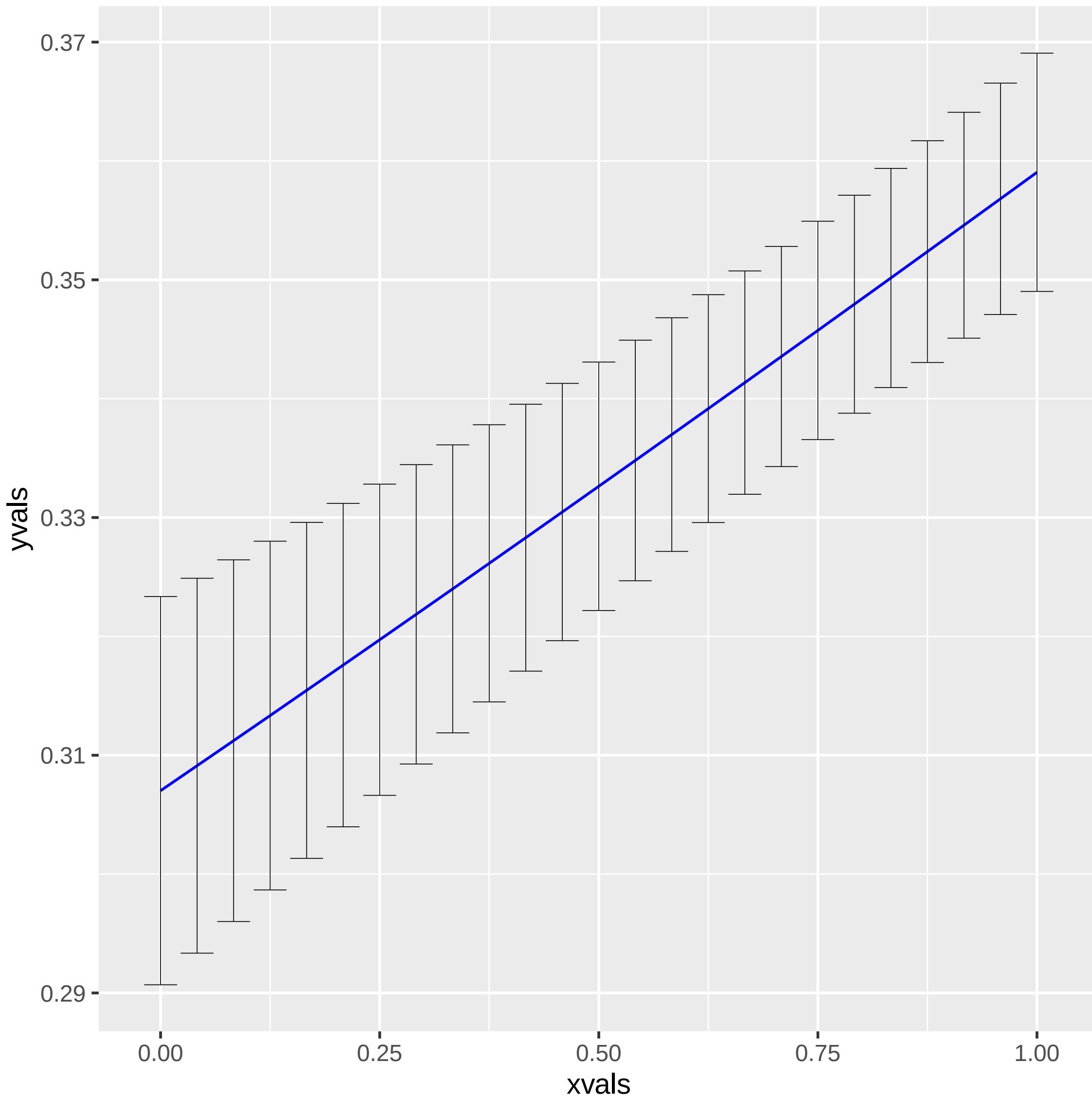
Interestingly, the average magnitude in the shifts in predicted probabilities are similar across comparable shifts in race and education, as well as their interaction. In a sense, shifts from whites to non-whites produce a similar change of the magnitude in predicted probabilities as shifts from high school to college education within whites and within non-whites. But the story is different for shifts in the size of place: on average, they produce a bigger shift in predicted probabilities than shifts in race and/or education variables.

This means that the probability that a person would support conealed carry of a gun is more strongly associated with whether a person lives in a rural or an urban area than with their education or race.

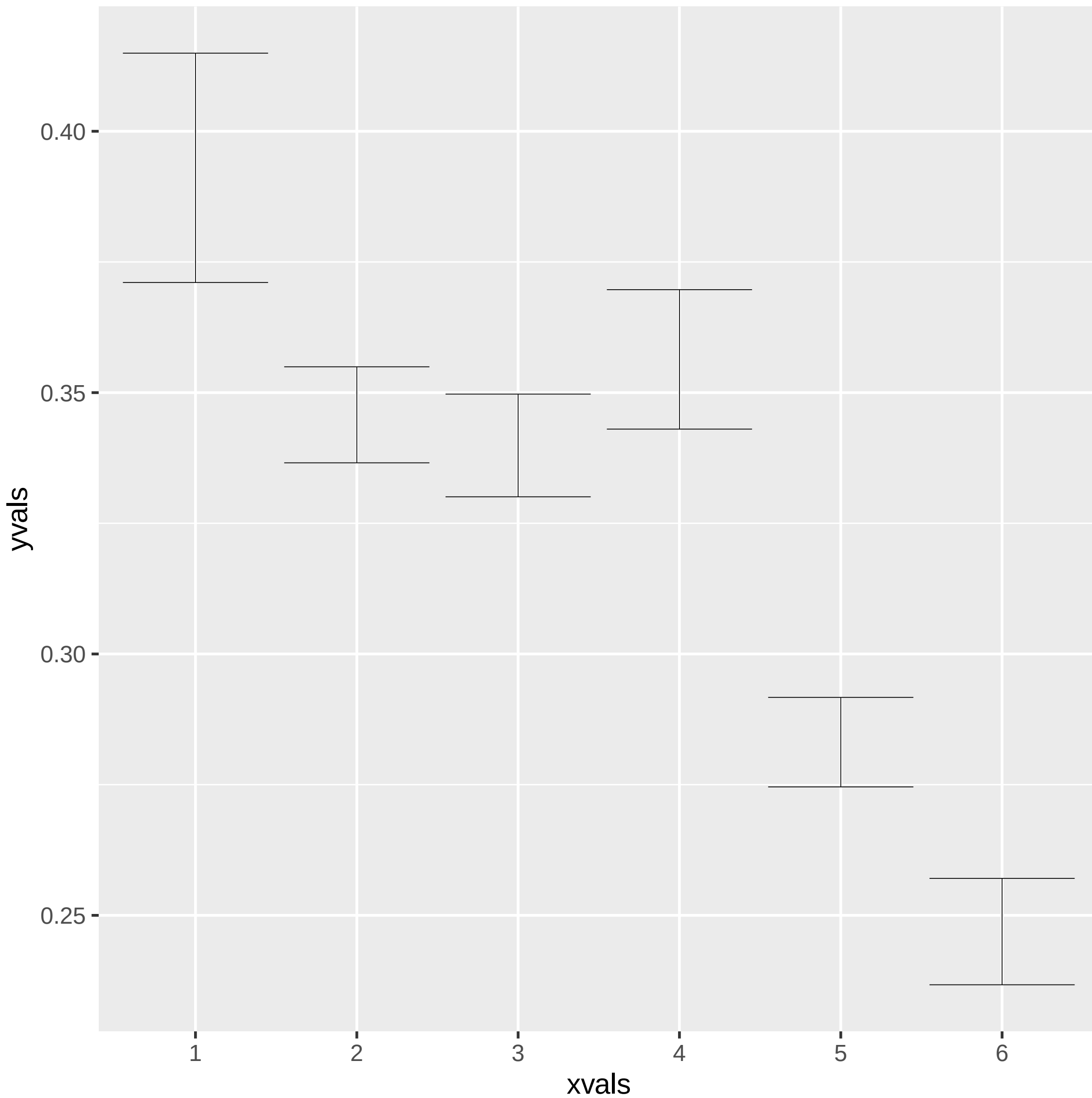
(f)

```
g_white <- cplot(probit_1, 'white', data = data, draw = F, vcov = vp_i)
g_educ <- cplot(probit_1, 'educf', data = data, draw = F, vcov = vp_i)

library(ggplot2)
# for whites
ggplot(data = g_white, aes(x = xvals, y = yvals)) + geom_line(color = 'blue') + geom_errorbar(aes(ymin =
```




```
# for education
ggplot(data = g_educ, aes(x = xvals, y = yvals)) + geom_line(color = 'blue') + geom_errorbar(aes(ymin =
```



Exercise 2

(a) Replication.

```
data <- data %>% mutate(approval = case_when(CC18_308a == 4 ~ 0, # strongly disapprove
                                             CC18_308a == 3 ~ 1, # somewhat disapprove
                                             CC18_308a == 2 ~ 2, # somewhat approve
                                             CC18_308a == 1 ~ 3)) # strongly approve
data <- data %>% mutate(approval = as.factor(approval))
data <- data %>% mutate(inputstate = as.factor(inputstate))


library(MASS)
library(sandwich)
probit_2 <- polr(data=data, approval ~ female + white + educf
                + white:educf + inputstate,
```

```r
                method = 'probit')

vp_2 <- vcovCL(probit_2, cluster = data$inputstate, type = 'HC1')

stargazer(probit_2, type = 'text',
          omit = 'inputstate', digits = 2,
          covariate.labels = c("female", "white", "high school", "some college",
                               "2-year college", "4-year college", "post-grad",
                               "white:high school", "white:some college",
                               "white:2-year college", "white:4-year college",
                               "white:post-grad"),
          add.lines = c("state fixed effects", "yes"),
          se = list(sqrt(diag(vp_2))))
```

```
================================================
                         Dependent variable:
                     ---------------------------
                              approval
------------------------------------------------
female                        -0.33***
                               (0.01)

white                          0.77***
                               (0.06)

high school                    -0.08
                               (0.05)

some college                  -0.14***
                               (0.05)

2-year college                 -0.05
                               (0.05)

4-year college                 -0.02
                               (0.05)

post-grad                      -0.05
                               (0.05)

white:high school               0.10*
                               (0.06)

white:some college             -0.03
                               (0.05)

white:2-year college           -0.03
                               (0.07)

white:4-year college          -0.39***
                               (0.07)

white:post-grad               -0.55***
```

```
                                    (0.06)

-----------------------------------------------
state fixed effects
yes
Observations                        58,202
===============================================
Note:                  *p<0.1; **p<0.05; ***p<0.01
```

(b) $(\widehat{Pr = 0}|x_1) = \Phi(-0.04 - 0.33 + 0.77 - 0.02 - 0.39 - 0.16) = 0.43$, where x_1 is a white woman w a 4-year college who lives in Minnesota (`inputstate`= 32, it's coef in the `summary` is -0.16), $\tau_1 = -0.04$

Exercise 3

$$cov(\hat{X}, Y) = \hat{\pi}_1 cov(Z, Y)$$

The first stage estimated equation is $\hat{X}_i = \hat{\pi}_0 + \hat{\pi}_1 Z_i$. Substitute the estimated X in its covariance with Y:

$$cov(\hat{X}, Y) = cov(\hat{\pi}_0 + \hat{\pi}_1 Z, Y) = cov(\hat{\pi}_1 Z, Y) = \hat{\pi}_1 cov(Z, Y),$$

by properties of the covariance. Do the same for the variance, using its properties:

$$var(\hat{X}) = var(\hat{\pi}_0 + \hat{\pi}_1 Z) = var(\hat{\pi}_1 Z) = \hat{\pi}^2 var(Z).$$

Hence, it is evident that

$$\hat{\beta}_1^{TSLS} = \frac{cov(\hat{X}, Y)}{var(\hat{X})} = \frac{\hat{\pi}_1 cov(Z, Y)}{\hat{\pi}^2 var(Z)} = \frac{cov(Z, Y)}{\frac{cov(X,Z)}{var(Z)} var(Z)} = \frac{cov(Z, Y)}{cov(Z, X)},$$

since $\hat{\pi}_1 = \dfrac{Cov(X, Z)}{Var(Z)}$.