Sofia Guo PSET 7 Econ 142

Sofia Guo

3/19/2019

1. Estimate fractions of C, AT, and NT for each month:

Given the first stage model:

$$FT_i = \pi_0 + \pi_1 T_i + \pi_x X_i + \eta_i$$

We first define each category:

1. Always takers (AT): $FT_i = 1$ and $T_i = 0$
2. Never takers (NT): $FT_i = 0$ and $T_i = 1$
3. NT + Compliers (C): $FT_i = 0$ and $T + i = 0$
4. AT + C: $FT_i = 1$ and $T_i = 1$

To get the fraction of compliers we subtract (1) from (4):

5. C: $[FT_i = 1, T_i = 1] - [FT_i = 1, T_i = 0]$

We calculate (1), (2), and (5) for each month:

```
#load packages
library(dplyr)
library(magrittr)
library(tidyr)
library(kableExtra)

#read in data
welfare <- read.csv("/Users/sofia/Box/Cal (sofiaguo@berkeley.edu)/2018-19/Spring 2019/Econ 142/PSETS/PS
```

Month 15:

```
#calculate [T = 0 & FT = 1] or AT
welfare_cont <- welfare %>%
  filter(treatment ==0)

AT_15 <- sum(welfare_cont$ft15)/nrow(welfare_cont)

#calculate [T = 0 & FT = ] or NT
welfare_treat <- welfare %>%
  filter(treatment ==1)

NTC_15 <- sum(welfare_treat$ft15)/nrow(welfare_treat)
NT_15 <- 1- NTC_15

#calculate compliers
C_15 <- NTC_15 - AT_15
```

Month 20:

```
#remove NA's
welfare_cont_na_20 <- welfare_cont %>%
  drop_na(ft20)
welfare_treat_na_20 <- welfare_treat %>%
  drop_na(ft20)

#calculate [T = 0 & FT = 1] or AT
AT_20 <- sum(welfare_cont_na_20$ft20)/nrow(welfare_cont_na_20)

#calculate [T = 0 & FT = ] or NT
NTC_20 <- sum(welfare_treat_na_20$ft20)/nrow(welfare_treat_na_20)
NT_20 <- 1- NTC_20

#calculate compliers
C_20 <- NTC_20 - AT_20
```

Month 24:

```
#remove NA's
welfare_cont_na_24 <- welfare_cont %>%
  drop_na(ft24)
welfare_treat_na_24 <- welfare_treat %>%
  drop_na(ft24)

#calculate [T = 0 & FT = 1] or AT
AT_24 <- sum(welfare_cont_na_24$ft24)/nrow(welfare_cont_na_24)

#calculate [T = 0 & FT = ] or NT
NTC_24 <- sum(welfare_treat_na_24$ft24)/nrow(welfare_treat_na_24)
NT_24 <- 1- NTC_24

#calculate compliers
C_24 <- NTC_24 - AT_24
```

Month 48:

```
#remove NA's
welfare_cont_na_48 <- welfare_cont %>%
  drop_na(ft48)
welfare_treat_na_48 <- welfare_treat %>%
  drop_na(ft48)

#calculate [T = 0 & FT = 1] or AT
AT_48 <- sum(welfare_cont_na_48$ft48)/nrow(welfare_cont_na_48)

#calculate [T = 0 & FT = ] or NT
NTC_48 <- sum(welfare_treat_na_48$ft48)/nrow(welfare_treat_na_48)
NT_48 <- 1- NTC_48
```

Table 1: Fraction of compliers, always takers and never takers by month

	Month 15	Month 20	Month 24	Month 48
Always Takers	0.1472	0.1609	0.1602	0.2337
Never Takers	0.7154	0.7245	0.7324	0.7158
Compliers	0.1374	0.1147	0.1074	0.0506

```
#calculate compliers
C_48 <- NTC_48 - AT_48
```

```
#create dataframe
fractions <- data.frame(Category = c("Always
                                     Takers", "Never Takers", "Compliers"),
                        Month15 = c(AT_15, NT_15, C_15),
                  Month20 = c(AT_20, NT_20, C_20),
                  Month24 = c(AT_24, NT_24, C_24),
                  Month48 = c(AT_48, NT_48, C_48))
```

```
#display table
kable(fractions,"latex", caption = "Fraction of compliers, always takers and never takers by month",  di
  kable_styling(position = "center")
```

2. Means of controls for the compliers

From lecture we are given:

$$E[X_i|C] = \frac{E[X_i|ATorC] * Pr(ATorC) - E[X_i|AT] * Pr(AT)}{Pr(C)}$$

We can calculate $E[X_i|C]$ using results from 1 and the means of controls $X_i = \{imm, hsgrad, agelt25, age35p, working_at_baseline$

Month 15:

```
#filter full time status
wt_15 <- welfare_treat %>%
  filter(ft15 == 1)
wc_15 <- welfare_cont %>%
  filter(ft15 ==1)
ATC_15 <- 1-NT_15

#hsgrad
ATC_hs_15 <- sum(wt_15$hsgrad)/nrow(wt_15)
AT_hs_15 <- sum(wc_15$hsgrad)/nrow(wc_15)
mean_hs_15 <- (ATC_hs_15 * ATC_15 - AT_hs_15 * AT_15)/C_15

#imm
ATC_imm_15 <- sum(wt_15$imm)/nrow(wt_15)
AT_imm_15 <- sum(wc_15$imm)/nrow(wc_15)
mean_imm_15 <- (ATC_imm_15 * ATC_15 - AT_imm_15 * AT_15)/C_15
```

```r
#agelt25
ATC_agelt25_15 <- sum(wt_15$agelt25)/nrow(wt_15)
AT_agelt25_15 <- sum(wc_15$agelt25)/nrow(wc_15)
mean_agelt25_15 <- (ATC_agelt25_15 * ATC_15 - AT_agelt25_15 * AT_15)/C_15

#age35p
ATC_age35p_15 <- sum(wt_15$age35p)/nrow(wt_15)
AT_age35p_15 <- sum(wc_15$age35p)/nrow(wc_15)
mean_age35p_15 <- (ATC_age35p_15 * ATC_15 - AT_age35p_15 * AT_15)/C_15

#baseline
ATC_baseline_15 <- sum(wt_15$working_at_baseline)/nrow(wt_15)
AT_baseline_15 <- sum(wc_15$working_at_baseline)/nrow(wc_15)
mean_baseline_15 <- (ATC_baseline_15 * ATC_15 - AT_baseline_15 * AT_15)/C_15

#anykidsu6
ATC_anykidsu6_15 <- sum(wt_15$anykidsu6)/nrow(wt_15)
AT_anykidsu6_15 <- sum(wc_15$anykidsu6)/nrow(wc_15)
mean_anykidsu6_15 <- (ATC_anykidsu6_15 * ATC_15 - AT_anykidsu6_15 * AT_15)/C_15

#nevermarried
ATC_nevermarried_15 <- sum(wt_15$nevermarried)/nrow(wt_15)
AT_nevermarried_15 <- sum(wc_15$nevermarried)/nrow(wc_15)
mean_nevermarried_15 <- (ATC_nevermarried_15 * ATC_15 - AT_nevermarried_15 * AT_15)/C_15

#x_15 means vector
X_15 <- c(mean_hs_15, mean_imm_15, mean_agelt25_15, mean_age35p_15, mean_baseline_15, mean_anykidsu6_15
```

Month 20:

```r
#filter full time status
wt_20 <- welfare_treat_na_20 %>%
  filter(ft20 == 1)
wc_20 <- welfare_cont_na_20 %>%
  filter(ft20 ==1)
ATC_20 <- 1-NT_20

#hsgrad
ATC_hs_20 <- sum(wt_20$hsgrad)/nrow(wt_20)
AT_hs_20 <- sum(wc_20$hsgrad)/nrow(wc_20)
mean_hs_20 <- (ATC_hs_20 * ATC_20 - AT_hs_20 * AT_20)/C_20

#imm
ATC_imm_20 <- sum(wt_20$imm)/nrow(wt_20)
AT_imm_20 <- sum(wc_20$imm)/nrow(wc_20)
mean_imm_20 <- (ATC_imm_20 * ATC_20 - AT_imm_20 * AT_20)/C_20

#agelt25
ATC_agelt25_20 <- sum(wt_20$agelt25)/nrow(wt_20)
AT_agelt25_20 <- sum(wc_20$agelt25)/nrow(wc_20)
mean_agelt25_20 <- (ATC_agelt25_20 * ATC_20 - AT_agelt25_20 * AT_20)/C_20
```

```r
#age35p
ATC_age35p_20 <- sum(wt_20$age35p)/nrow(wt_20)
AT_age35p_20 <- sum(wc_20$age35p)/nrow(wc_20)
mean_age35p_20 <- (ATC_age35p_20 * ATC_20 - AT_age35p_20 * AT_20)/C_20

#baseline
ATC_baseline_20 <- sum(wt_20$working_at_baseline)/nrow(wt_20)
AT_baseline_20 <- sum(wc_20$working_at_baseline)/nrow(wc_20)
mean_baseline_20 <- (ATC_baseline_20 * ATC_20 - AT_baseline_20 * AT_20)/C_20

#anykidsu6
ATC_anykidsu6_20 <- sum(wt_20$anykidsu6)/nrow(wt_20)
AT_anykidsu6_20 <- sum(wc_20$anykidsu6)/nrow(wc_20)
mean_anykidsu6_20 <- (ATC_anykidsu6_20 * ATC_20 - AT_anykidsu6_20 * AT_20)/C_20

#nevermarried
ATC_nevermarried_20 <- sum(wt_20$nevermarried)/nrow(wt_20)
AT_nevermarried_20 <- sum(wc_20$nevermarried)/nrow(wc_20)
mean_nevermarried_20 <- (ATC_nevermarried_20 * ATC_20 - AT_nevermarried_20 * AT_20)/C_20

#x_20 means vector
X_20 <- c(mean_hs_20, mean_imm_20, mean_agelt25_20, mean_age35p_20, mean_baseline_20, mean_anykidsu6_20
```

Month 24:

```r
#filter full time status
wt_24 <- welfare_treat_na_24 %>%
  filter(ft24 == 1)
wc_24 <- welfare_cont_na_24 %>%
  filter(ft24 ==1)
ATC_24 <- 1-NT_24

#hsgrad
ATC_hs_24 <- sum(wt_24$hsgrad)/nrow(wt_24)
AT_hs_24 <- sum(wc_24$hsgrad)/nrow(wc_24)
mean_hs_24 <- (ATC_hs_24 * ATC_24 - AT_hs_24 * AT_24)/C_24

#imm
ATC_imm_24 <- sum(wt_24$imm)/nrow(wt_24)
AT_imm_24 <- sum(wc_24$imm)/nrow(wc_24)
mean_imm_24 <- (ATC_imm_24 * ATC_24 - AT_imm_24 * AT_24)/C_24

#agelt25
ATC_agelt25_24 <- sum(wt_24$agelt25)/nrow(wt_24)
AT_agelt25_24 <- sum(wc_24$agelt25)/nrow(wc_24)
mean_agelt25_24 <- (ATC_agelt25_24 * ATC_24 - AT_agelt25_24 * AT_24)/C_24

#age35p
ATC_age35p_24 <- sum(wt_24$age35p)/nrow(wt_24)
AT_age35p_24 <- sum(wc_24$age35p)/nrow(wc_24)
mean_age35p_24 <- (ATC_age35p_24 * ATC_24 - AT_age35p_24 * AT_24)/C_24
```

```
ATC_age35p_24
```

```
## [1] 0.2810734
```

```
#baseline
ATC_baseline_24 <- sum(wt_24$working_at_baseline)/nrow(wt_24)
AT_baseline_24 <- sum(wc_24$working_at_baseline)/nrow(wc_24)
mean_baseline_24 <- (ATC_baseline_24 * ATC_24 - AT_baseline_24 * AT_24)/C_24

#anykidsu6
ATC_anykidsu6_24 <- sum(wt_24$anykidsu6)/nrow(wt_24)
AT_anykidsu6_24 <- sum(wc_24$anykidsu6)/nrow(wc_24)
mean_anykidsu6_24 <- (ATC_anykidsu6_24 * ATC_24 - AT_anykidsu6_24 * AT_24)/C_24

#nevermarried
ATC_nevermarried_24 <- sum(wt_24$nevermarried)/nrow(wt_24)
AT_nevermarried_24 <- sum(wc_24$nevermarried)/nrow(wc_24)
mean_nevermarried_24 <- (ATC_nevermarried_24 * ATC_24 - AT_nevermarried_24 * AT_24)/C_24

#x_24 means vector
X_24 <- c(mean_hs_24, mean_imm_24, mean_agelt25_24, mean_age35p_24, mean_baseline_24, mean_anykidsu6_24
```

Month 48:

```
#filter full time status
wt_48 <- welfare_treat_na_48 %>%
  filter(ft48 == 1)
wc_48 <- welfare_cont_na_48 %>%
  filter(ft48 ==1)
ATC_48 <- 1-NT_48

#hsgrad
ATC_hs_48 <- sum(wt_48$hsgrad)/nrow(wt_48)
AT_hs_48 <- sum(wc_48$hsgrad)/nrow(wc_48)
mean_hs_48 <- (ATC_hs_48 * ATC_48 - AT_hs_48 * AT_48)/C_48

#imm
ATC_imm_48 <- sum(wt_48$imm)/nrow(wt_48)
AT_imm_48 <- sum(wc_48$imm)/nrow(wc_48)
mean_imm_48 <- (ATC_imm_48 * ATC_48 - AT_imm_48 * AT_48)/C_48

#agelt25
ATC_agelt25_48 <- sum(wt_48$agelt25)/nrow(wt_48)
AT_agelt25_48 <- sum(wc_48$agelt25)/nrow(wc_48)
mean_agelt25_48 <- (ATC_agelt25_48 * ATC_48 - AT_agelt25_48 * AT_48)/C_48

#age35p
ATC_age35p_48 <- sum(wt_48$age35p)/nrow(wt_48)
AT_age35p_48 <- sum(wc_48$age35p)/nrow(wc_48)
mean_age35p_48 <- (ATC_age35p_48 * ATC_48 - AT_age35p_48 * AT_48)/C_48

#baseline
ATC_baseline_48 <- sum(wt_48$working_at_baseline)/nrow(wt_48)
```

Table 2: Mean of control variables for compliers

	Month 15	Month 20	Month 24	Month 48
HS Grad	0.5349	0.4390	0.5144	0.4325
Immigrant	0.1594	0.1276	0.1256	0.1440
Age < 25	0.1894	0.1848	0.2469	0.1646
Age > 35	0.2452	0.2697	0.2633	0.3382
Baseline Wage	0.2028	0.2199	0.1772	0.1024
Kids under 6	0.5632	0.5187	0.5643	0.4706
Never Married	0.4713	0.4442	0.4957	0.4090

```
AT_baseline_48 <- sum(wc_48$working_at_baseline)/nrow(wc_48)
mean_baseline_48 <- (ATC_baseline_48 * ATC_48 - AT_baseline_48 * AT_48)/C_48

#anykidsu6
ATC_anykidsu6_48 <- sum(wt_48$anykidsu6)/nrow(wt_48)
AT_anykidsu6_48 <- sum(wc_48$anykidsu6)/nrow(wc_48)
mean_anykidsu6_48 <- (ATC_anykidsu6_48 * ATC_48 - AT_anykidsu6_48 * AT_48)/C_48

#nevermarried
ATC_nevermarried_48 <- sum(wt_48$nevermarried)/nrow(wt_48)
AT_nevermarried_48 <- sum(wc_48$nevermarried)/nrow(wc_48)
mean_nevermarried_48 <- (ATC_nevermarried_48 * ATC_48 - AT_nevermarried_48 * AT_48)/C_48

#x_48 means vector
X_48 <- c(mean_hs_48, mean_imm_48, mean_agelt25_48, mean_age35p_48, mean_baseline_48, mean_anykidsu6_48
```

```
#make master df
df <- data.frame(Control = c("HS Grad",
                             "Immigrant",
                             "Age < 25",
                             "Age > 35",
                             "Baseline Wage",
                             "Kids under 6",
                             "Never Married"),
                 Month15 = X_15,
                 Month20 = X_20,
                 Month24 = X_24,
                 Month48 = X_48)
```

```
#display table
kable(df,"latex", caption = "Mean of control variables for compliers",  digits = 4, booktabs = T, align
  kable_styling(position = "center")
```