

# Unadjusted ATE Estimation

Amos Okutse

23 December, 2022

## Contents

0.0.1	LOAD DATA . . . . .	1
0.1	UNADJUSTED RESULTS . . . . .	1
0.2	EXTRACT THE RESULTS . . . . .	2

### 0.0.1 LOAD DATA

```
rm(list = ls())
## load the saved single data files
load("C:\\Users\\aokutse\\OneDrive - Brown
  ↳ University\\ThesisResults\\data\\df_one.RData")
load("C:\\Users\\aokutse\\OneDrive - Brown
  ↳ University\\ThesisResults\\data\\df_two.RData")
load("C:\\Users\\aokutse\\OneDrive - Brown
  ↳ University\\ThesisResults\\data\\df_three.RData")
load("C:\\Users\\aokutse\\OneDrive - Brown
  ↳ University\\ThesisResults\\data\\df_four.RData")

## load the saved list data files
load("C:\\Users\\aokutse\\OneDrive - Brown
  ↳ University\\ThesisResults\\data\\dssets1.RData")
load("C:\\Users\\aokutse\\OneDrive - Brown
  ↳ University\\ThesisResults\\data\\dssets2.RData")
load("C:\\Users\\aokutse\\OneDrive - Brown
  ↳ University\\ThesisResults\\data\\dssets3.RData")
load("C:\\Users\\aokutse\\OneDrive - Brown
  ↳ University\\ThesisResults\\data\\dssets4.RData")
```

### 0.1 UNADJUSTED RESULTS

```
## function to return the estimated treatment effects across n = 1000 data sets

## - full data
```

Table 1: Unadjusted estimates of the average treatment effect across  $n = 1000$  datasets under full and observed data analysis

Data generating values	n	ate	sd	bias
n = 500, SD = 1	500	50.10317	3.299761	0.1031665
n = 500, SD = 1	244	50.03152	4.349138	0.0315233
n = 500, SD = 45	500	50.17186	4.981188	0.1718594
n = 500, SD = 45	258	49.99792	7.325424	-0.0020821
n = 2000, SD = 1	2000	49.97075	1.660128	-0.0292450
n = 2000, SD = 1	997	50.00518	2.169228	0.0051760
n = 2000, SD = 45	2000	50.02245	2.662620	0.0224526
n = 2000, SD = 45	1003	50.17771	3.695738	0.1777148

```
##- observed data
## - observed modified

unadj <- function(df = NULL){
  ## since this is based on the full data set, then use the full data set
  full_unadjusted = mean(df$y[df$A == 1]) - mean(df$y[df$A == 0])
  full_bias_unadjusted = full_unadjusted - 50

  ## subset the data to only subjects with R == 1
  df2 <- dplyr::filter(df, R == 1)
  observed_unadjusted = mean(df2$y[df2$A == 1]) - mean(df2$y[df2$A == 0])
  observed_bias = observed_unadjusted - 50
  return(data.frame(full_unadjusted, full_bias_unadjusted, observed_unadjusted,
    ↪ observed_bias))
}
```

```
## analyze the results into a table
df_onea <- onea %>% map_dfr(data.frame) # n = 500; sd = 1
df_oneb <- oneb %>% map_dfr(data.frame) # n = 500; sd = 45
df_onec <- onec %>% map_dfr(data.frame) # n = 2000; sd = 1
df_oned <- oned %>% map_dfr(data.frame) # n = 2000; sd = 1
```

## 0.2 EXTRACT THE RESULTS

```
## create final table of results
unadjusted = bind_rows(list("n = 500, SD = 1" = full, "n = 500, SD = 1" = obs,
  "n = 500, SD = 45" = full2, "n = 500, SD = 45" = obs2,
  "n = 2000, SD = 1" = full3, "n = 2000, SD = 1" = obs3,
  "n = 2000, SD = 45" = full4, "n = 2000, SD = 45" = obs4),
  .id = "Data generating values")
kable(unadjusted, format = "latex", caption = "Unadjusted estimates of the average
  ↪ treatment effect across n = 1000 datasets under full and observed data analysis")
```

```
## the order of the rows starts with n = 500
write.csv(unadjusted, file = "C:\\Users\\aokutse\\OneDrive - Brown
  ↪ University\\ThesisResults\\[1]_unadjusted\\unadjusted_results.csv", row.names =
  ↪ FALSE)
```

---