

# Problem Set 2

Petra Ivanovic - pi2018 - Section 005

Due Nov 10, 2023

This homework must be turned in on Brightspace by Nov. 10, 2023. It must be your own work, and your own work only – you must not copy anyone’s work, or allow anyone to copy yours. This extends to writing code. You may consult with others, but when you write up, you must do so alone.

Your homework submission must be written and submitted using Rmarkdown. No handwritten solutions will be accepted. **No zip files will be accepted. Make sure we can read each line of code in the pdf document.** You should submit the following:

1. A compiled PDF file named yourNetID\_solutions.pdf containing your solutions to the problems.
2. A .Rmd file containing the code and text used to produce your compiled pdf named your-NetID\_solutions.Rmd.

Note that math can be typeset in Rmarkdown in the same way as Latex. Please make sure your answers are clearly structured in the Rmarkdown file:

1. Label each question part
2. Do not include written answers as code comments.
3. The code used to obtain the answer for each question part should accompany the written answer. Comment your code!

## Question 1 (Total: 50)

In new democracies and post-conflict settings, Truth and Reconciliation Commissions (TRCs) are often tasked with investigating and reporting about wrongdoing in previous governments. Depending on the context, institutions such as TRCs are expected to reduce hostilities (e.g. racial hostilities) and promote peace.

In 1995, South Africa's new government formed a national TRC in the aftermath of apartheid. [Gibson 2004] uses survey data collected from 2000-2001 to examine whether this TRC promoted inter-racial reconciliation. The outcome of interest is respondent racial attitudes (as measured by the level of agreement with the prompt: "I find it difficult to understand the customs and ways of [the opposite racial group]"). The treatment is "exposure to the TRC" as measured by the individual's level of self-reported knowledge about the TRC.

You will need to use the `trc_data.dta` file for this question. The relevant variables are:

- RUSTAND - Outcome: respondent's racial attitudes (higher values indicate greater agreement)
- TRCKNOW - Treatment dummy (1 = if knows about the TRC, 0 = otherwise)
- age - Respondent age (in 2001)
- female - Respondent gender
- wealth - Measure of wealth constructed based on asset ownership (assets are fridge, floor polisher, vacuum cleaner, microwave oven, hi-fi, washing machine, telephone, TV, car)
- religiosity - Self-reported religiosity (7 point scale)
- ethsalience - Self-reported ethnic identification (4 point scale)
- rcblack - Respondent is black
- rcwhite - Respondent is white
- rccol - Respondent is coloured (distinct multiracial ethnic group)
- EDUC - Level of education (9 point scale)

### Part a (15 points)

Estimate the average treatment effect of TRC exposure on respondents' racial attitudes under the assumption that TRC exposure is ignorable. Report a 95% confidence interval for your estimate and interpret your results. (Use robust standard errors throughout.)

```
library(tidyverse)
library(haven)
library(estimatr) # for lm with robust se : ?lm_robust()

# Load in the TRC data (it's a STATA .dta so we use the haven package)
TRC_data <- haven::read_dta("trc_data.dta")

# Estimate the model with robust standard errors
model <- lm_robust(RUSTAND ~ TRCKNOW, data = TRC_data)

# Get the summary of the model
summary(model)
```

```
##
## Call:
## lm_robust(formula = RUSTAND ~ TRCKNOW, data = TRC_data)
##
## Standard error type: HC2
##
```

```
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper  DF
## (Intercept)   2.5311    0.02806  90.212 0.000e+00   2.4761   2.5862 3203
## TRCKNOW       -0.2177    0.04433  -4.911 9.492e-07  -0.3047  -0.1308 3203
##
## Multiple R-squared:  0.007626 , Adjusted R-squared:  0.007316
## F-statistic: 24.12 on 1 and 3203 DF,  p-value: 9.492e-07

# Directly extract the confidence interval for the treatment effect (TRCKNOW)
confint(model, level = 0.95)

##           2.5 %      97.5 %
## (Intercept) 2.476131 2.5861565
## TRCKNOW     -0.304652 -0.1308115
```

We have built a model that estimates average effect (ATE) of TRC exposure on respondents' racial attitudes under the assumption that TRC exposure is ignorable. Our results show that on average people who have been exposed to TRC **reduced** respondent's racial attitudes by **0.2177**. As this is negative we can say that people with TRC exposure find it less difficult to understand the customs and ways of the opposite racial group. In the 95% confidence interval for this estimate is **[-0.3047, -0.1308]**. Evaluating our results we can say that both our intercept and TRC variable are statistically significant (not due to the random chance) due to the p-value being lower than  $\alpha = 0.05$  (lets choose this as our threshold). Further, we can see that 0 is not in our confidence interval for the TRC variable which indicates that we can be 95% confident that the true effect of TRC exposure on racial attitudes is negative, and thus, significant. In other words, the null hypothesis of no average treatment effect can be rejected at 5% ( $\alpha = 0.05$ ) significance level.

## Part b (15 points)

Examine whether exposed and nonexposed respondents differ on the full set of observed covariates using a series of balance tests. Briefly discuss, in which ways do exposed and nonexposed respondents differ?

```
library(dplyr)

# Calculate means for each group
balance_table <- TRC_data %>%
  group_by(TRCKNOW) %>%
  summarize(
    age_mean = mean(age, na.rm = TRUE),
    female_mean = mean(female, na.rm = TRUE),
    female_count = sum(female, na.rm = TRUE),
    wealth_mean = mean(wealth, na.rm = TRUE),
    religiosity_mean = mean(religiosity, na.rm = TRUE),
    ethsalience_mean = mean(ethsalience, na.rm = TRUE),
    black_mean = mean(rcblack, na.rm = TRUE),
    black_count = sum(rcblack, na.rm = TRUE),
    white_mean = mean(rcwhite, na.rm = TRUE),
    white_count = sum(rcwhite, na.rm = TRUE),
    color_mean = mean(rccol, na.rm = TRUE),
    color_count = sum(rccol, na.rm = TRUE),
    education_mean = mean(EDUC, na.rm = TRUE),
    .groups = "drop"
  )
```

```

# Calculate differences for each mean variable, excluding TRCKNOW
difference_row <- balance_table %>%
  summarize(across(ends_with("mean") | ends_with("count"), ~ .[2] - .[1])) %>%
  mutate(TRCKNOW = "Difference") # Add TRCKNOW as a new column with the value "Difference"

# Add the difference row to the balance table
# Ensure that TRCKNOW is a character in both data frames to match types
balance_table$TRCKNOW <- as.character(balance_table$TRCKNOW)
balance_table <- bind_rows(balance_table, difference_row)

# View the balance table
print(balance_table)

```

```

## # A tibble: 3 x 14
##   TRCKNOW    age_mean female_mean female_count wealth_mean religiosity_mean
##   <chr>      <dbl>      <dbl>         <dbl>      <dbl>          <dbl>
## 1 0         40.5        0.433         764        5793.          3.92
## 2 1         38.9        0.538         774        6945.          3.84
## 3 Difference -1.51        0.105         10        1152.         -0.0755
## # i 8 more variables: ethsalience_mean <dbl>, black_mean <dbl>,
## #   black_count <dbl>, white_mean <dbl>, white_count <dbl>, color_mean <dbl>,
## #   color_count <dbl>, education_mean <dbl>

```

```

# Print out number of rows in the dataset
cat("Number of rows in the dataset:", nrow(TRC_data), "\n")

```

```

## Number of rows in the dataset: 3205

```

```

library(tidyverse)
library(haven)
library(estimatr)

# Create a treatment group indicator
trc_exposed <- TRC_data$TRCKNOW == 1

# Initialize an empty list to store results
significant_covariates <- list()

# Compare the exposed and nonexposed groups on each covariate
for (covariate in c("age", "female", "wealth", "religiosity", "ethsalience",
  "rcblack", "rcwhite", "rccol", "EDUC")) {
  # Test for differences in means using a t-test
  t_test_result <- t.test(TRC_data[trc_exposed, covariate, drop = FALSE], TRC_data[!trc_exposed, covariate, drop = FALSE])

  # Check if the p-value is less than 0.05
  if (t_test_result$p.value < 0.05) {
    # Save the t-test result to the list
    significant_covariates[[covariate]] <- t_test_result
  }
}

# Print the results for statistically significant covariates

```

```

for (covariate in names(significant_covariates)) {
  cat("Covariate:", covariate, "\n")
  #cat("t-value:", significant_covariates[[covariate]]$statistic, "\n")
  cat("p-value:", significant_covariates[[covariate]]$p.value, "\n\n")
}

```

```

## Covariate: age
## p-value: 0.005398203
##
## Covariate: female
## p-value: 2.813896e-09
##
## Covariate: wealth
## p-value: 1.541194e-05
##
## Covariate: rcblack
## p-value: 0.02872185
##
## Covariate: rccol
## p-value: 9.144986e-05
##
## Covariate: EDUC
## p-value: 3.217942e-27

```

Above we have shown the differences in means (and counts for binary variables) between exposed and unexposed respondents. We can see that while there is difference in most variables in some it is significant (both statistically and ‘logically’ (in real world terms)) and for some it is not. Further, we have reported Covariates and corresponding p-values for those that have statistically significant differences between exposed and nonexposed respondents. We can see that those covariates are age, female (respondents gender), wealth, rcblack (respondent is black), rccol (respondent is coloured), and EDUC. Looking at the table above I would argue that the variables that are significantly different (in the real life applications not only statistically significant) are wealth as the mean in both groups is around 4 or 5 thousand to 1 thousand difference is a lot, count of black respondents as there are around 800 or 900 black respondents in each category, a difference of 112 is a lot, similarly for ‘coloured’ respondents the difference of 119 is extreme considering there is only around 280 respondents in one and 160 in other category. For 3 variables that are statistically significant, age, female, and education, I would argue that there is not really that significant of a difference as the difference is only about 1 year for age, 10 people for ‘female’, and 0.4 (on the 9 point scale) for an education.

However, since there are significant differences between groups we could say that the assumption that TRC exposure is completely ignorable is not entirely true.

### Part c (10 points)

Now assume that TRC exposure is conditionally ignorable given the set of observed covariates:

1. Use a logistic regression model to estimate the propensity score for each observation. (For purposes of this question, do not include any interactions.)
2. With this model, construct inverse propensity of treatment weights (IPTW) for each observation using the unstabilized weights.
3. Use the propensity score to construct an IPW estimator and report the point estimate for the ATE.

Use the following covariates: age, female, wealth, religiosity, ethsalience, rcblack, rcwhite, rccol, EDUC

1. Building a Logistic Regression model to estimate propensity scores for each observation.

```
library(tidyverse)
library(haven)
library(broom)

pscore_model <- glm(TRCKNOW ~ age + female + wealth + religiosity +
  ethsalience + rcblack + rcwhite + rccol + EDUC,
  data = TRC_data, family = binomial(link = "logit"))

tidy(pscore_model)
```

```
## # A tibble: 10 x 5
##   term      estimate std.error statistic  p.value
##   <chr>      <dbl>    <dbl>    <dbl>   <dbl>
## 1 (Intercept) -2.52      0.313     -8.03  9.81e-16
## 2 age          0.000371  0.00254      0.146  8.84e- 1
## 3 female       0.388     0.0751      5.17  2.32e- 7
## 4 wealth       0.0000244 0.00000685    3.56  3.75e- 4
## 5 religiosity  0.0113     0.0209      0.540  5.89e- 1
## 6 ethsalience 0.0601     0.0650      0.925  3.55e- 1
## 7 rcblack      0.472     0.152      3.10  1.93e- 3
## 8 rcwhite     -0.280     0.163     -1.71  8.66e- 2
## 9 rccol       -0.215     0.171     -1.26  2.08e- 1
## 10 EDUC        0.392     0.0396      9.91  3.64e-23
```

2. Constructing IPTW for each observation using the unstabilized weights

```
TRC_data$e <- predict(pscore_model, type = "response")
TRC_data$wt <- NA
TRC_data$wt[TRC_data$TRCKNOW == 1] <- 1/TRC_data$e[TRC_data$TRCKNOW==1]
TRC_data$wt[TRC_data$TRCKNOW == 0] <- 1/(1 - TRC_data$e[TRC_data$TRCKNOW==0])
```

3. Construct an IPW estimator to report point estimate for ATE

```
point_wtd <- mean(TRC_data$wt * TRC_data$RUSTAND * TRC_data$TRCKNOW -
  TRC_data$wt * TRC_data$RUSTAND * (1-TRC_data$TRCKNOW))

point_wtd
```

```
## [1] -0.162329
```

The point estimate for the ATE is -0.1623.

#### Part d (10 points)

Using the bootstrap method (resampling individual rows of the data with replacement), obtain an estimate for the standard error of your IPTW estimator for the ATE. Compute a 95% confidence interval and interpret your findings. (You should report estimate, standard error, 95% CI lower, 95% CI upper, for interpretation, compare your results in Part C/D to your estimate from Part A and briefly discuss your findings.)

```

# Set random seed
set.seed(123)

nBoot <- 1000 # Number of iterations
ate_boot <- rep(NA, nBoot) # Placeholder to store estimates
# For each iteration
for(boot in 1:nBoot){
  # Resample rows with replacement
  trc_boot <- TRC_data[sample(1:nrow(TRC_data), nrow(TRC_data), replace=T),]

  # Fit the propensity score model on the bootstrapped data
  pscore_model_boot <- glm(TRCKNOW ~ age + female + wealth + religiosity +
                           ethsalience + rcblack + rcwhite + rccol + EDUC,
                           data = trc_boot, family = binomial(link = "logit"))

  trc_boot$e <- predict(pscore_model_boot, type = "response")

  # Calculate the weights
  trc_boot$wt <- NA
  trc_boot$wt[trc_boot$TRCKNOW == 1] <- 1/trc_boot$e[trc_boot$TRCKNOW==1]
  trc_boot$wt[trc_boot$TRCKNOW == 0] <- 1/(1 - trc_boot$e[trc_boot$TRCKNOW==0])
  # Compute and store the ATE
  ate_boot[boot] <-
  mean(trc_boot$wt * trc_boot$RUSTAND * trc_boot$TRCKNOW -
        trc_boot$wt * trc_boot$RUSTAND * (1-trc_boot$TRCKNOW))
}
# Take the SD of the ate_boot to get our estimated SE - can do asymptotic inference
mean(ate_boot)

```

```
## [1] -0.1596519
```

```
sd(ate_boot)
```

```
## [1] 0.04534277
```

```

c(point_wtd - qnorm(.975)*sd(ate_boot),
  point_wtd + qnorm(.975)*sd(ate_boot))

```

```
## [1] -0.25119917 -0.07345878
```

Out point estimate for ATE from the bootstrap is -0.1597, standard error of ate is 0.0453 and the confidence interval for ATE is between -0.2512 and -0.0735. We can interpret this as being 95% confident that the true ATE lies within this interval, and since 0 is not in the interval we can say that the effect is statistically significant at the 5% level.

In part A our ATE was -0.2177 with a 95% CI between -0.3047 and -0.1308 and standard error of 0.0443. This result is very similar to the one from part D (the bootstrapping method) as they are both entirely negative and do not include 0 which indicates that in both cases we can conclude that TRC exposure most likely reduces negative racial attitudes.

In part C our ATE was -0.1623 and as such is very close to that of part D and falls in both confidence intervals from part D and part A. This is consistent with our previous claim that TRC exposure leads to a reduction in negative racial attitudes.

From this analysis we can see that ATE point estimates (and their confidence intervals) from each 3 cases showed us that TRC exposure has a statistically significant (at 0.05 level) negative effect on respondents' racial attitudes. This means that in general, those with TRC exposure find it less difficult to understand the customs and ways of the opposite racial group.

## Question 2 (Total: 50 points)

Use the same data set as in Question 1.

### Part a (15 points)

Estimate the ATT of TRC exposure on respondents' racial attitudes using the MatchIt approach. You can use the matchit function from MatchIt package in R. Implement the nearest neighbor matching algorithm and estimate the ATT. Report the 95% confidence interval of your estimate.

```
library(MatchIt)
```

```
## Warning: package 'MatchIt' was built under R version 4.0.5
```

```
# Read the help file first! Check out the default settings  
# ?matchit()
```

```
# Perform nearest neighbor matching  
matchit_model <- matchit(TRCKNOW ~ age + female + wealth + religiosity +  
  ethsalience + rcblack + rcwhite + rccol + EDUC,  
  data = TRC_data,  
  method = "nearest",  
  estimand = "ATT",  
  distance = "glm")
```

```
# Get the matched data  
matched_data <- match.data(matchit_model)
```

```
# Estimate the ATT on the matched data  
outcome_model <- lm_robust(RUSTAND ~ TRCKNOW + age + female + wealth +  
  religiosity + ethsalience + rcblack + rcwhite +  
  rccol + EDUC, data = matched_data)
```

```
# Extract the coefficient for TRCKNOW and its standard error  
att_estimate_nn <- coef(outcome_model)['TRCKNOW']  
se_nn <- summary(outcome_model)$coefficients['TRCKNOW', 'Std. Error']
```

```
# Calculate the 95% confidence interval  
ci_lower_nn <- att_estimate_nn - qnorm(0.975) * se_nn  
ci_upper_nn <- att_estimate_nn + qnorm(0.975) * se_nn
```

```
# Output the estimate, standard error, and confidence interval  
cat("ATT: ", att_estimate_nn, "\n",  
  "SE: ", se_nn, "\n",  
  "95% CI: [", ci_lower_nn, ", ", ci_upper_nn, "]", sep_nn = "")
```



```
## ATT: -0.1814553
## SE: 0.04544308
## 95% CI: [ -0.2705221 , -0.09238853 ]
```

Implementing the Nearest Neighbor Matching algorithm we have estimated ATT of TRC exposure on respondents' racial attitudes to be -0.1815. Further, the Standard Error of ATT is estimated at 0.0454 and Confidence interval is [-0.2705, -0.0924]. We can interpret this as being 95% confident that the true ATE lies within this interval, and since 0 is not in the interval we can say that the effect is statistically significant at the 5% level.

## Part b (15 points)

Estimate the ATT of TRC exposure on respondents' racial attitudes using the MatchIt approach. You can use the matchit function from MatchIt package in R. Implement the exact matching algorithm and estimate the ATT. Report the 95% confidence interval of your estimate.

```
library(MatchIt)
# Read the help file first! Check out the default settings
# ?matchit()

library(MatchIt)

# Perform nearest neighbor matching
matchit_model <- matchit(TRCKNOW ~ age + female + wealth + religiosity +
                        ethsalience + rcblack + rcwhite + rccol + EDUC,
                        data = TRC_data,
                        method = "exact",
                        estimand = "ATT",
                        distance = "glm")

# Obtain the matched data
matched_data <- match.data(matchit_model)

# Estimate the ATT on the matched data
outcome_model <- lm_robust(RUSTAND ~ TRCKNOW + age + female + wealth +
                        religiosity + ethsalience + rcblack + rcwhite +
                        rccol + EDUC, data = matched_data)

# Extract the coefficient for TRCKNOW and its standard error
att_estimate_ex <- coef(outcome_model)['TRCKNOW']
se_ex <- summary(outcome_model)$coefficients['TRCKNOW', 'Std. Error']

# Calculate the 95% confidence interval
ci_lower_ex <- att_estimate_ex - qnorm(0.975) * se_ex
ci_upper_ex <- att_estimate_ex + qnorm(0.975) * se_ex

# Output the estimate, standard error, and confidence interval
cat("ATT: ", att_estimate_ex, "\n",
    "SE: ", se_ex, "\n",
    "95% CI: [", ci_lower_ex, ", ", ci_upper_ex, "]", sep_nn = "")
```

```
## ATT: 0.1027187
## SE: 0.1787245
```

```
## 95% CI: [ -0.2475749 , 0.4530123 ]
```

Implementing the Exact Matching algorithm we have estimated ATT of TRC exposure on respondents' racial attitudes to be 0.1027. Further, the Standard Error of ATT is estimated at 0.1787 and Confidence interval is [-0.2478, 0.4530]. We can interpret this as being 95% confident that the true ATE lies within this interval, however, since 0 is included in our CI we cannot with say that the effect is statistically significant at the 5% level. It is worth noting that the CI is significantly broader than in part a and that the standard error is also much higher.

### Part c (10 points)

Estimate the ATT of TRC exposure on respondents' racial attitudes using the MatchIt approach. You can use the matchit function from MatchIt package in R. Implement the **coarsened exact matching** algorithm and estimate the ATT. Report the 95% confidence interval of your estimate.

```
library(MatchIt)
# Read the help file first! Check out the default settings
# ?matchit()

library(MatchIt)

# Perform nearest neighbor matching
matchit_model <- matchit(TRCKNOW ~ age + female + wealth + religiosity +
                        ethsalience + rcblack + rcwhite + rccol + EDUC,
                        data = TRC_data,
                        method = "cem", #coarsened exact matching
                        estimand = "ATT",
                        distance = "glm")

# Obtain the matched data
matched_data <- match.data(matchit_model)

# Estimate the ATT on the matched data
outcome_model <- lm_robust(RUSTAND ~ TRCKNOW + age + female + wealth +
                        religiosity + ethsalience + rcblack + rcwhite +
                        rccol + EDUC, data = matched_data)

# Extract the coefficient for TRCKNOW and its standard error
att_estimate_cem <- coef(outcome_model)['TRCKNOW']
se_cem <- summary(outcome_model)$coefficients['TRCKNOW', 'Std. Error']

# Calculate the 95% confidence interval
ci_lower_cem <- att_estimate_cem - qnorm(0.975) * se_cem
ci_upper_cem <- att_estimate_cem + qnorm(0.975) * se_cem

# Output the estimate, standard error, and confidence interval
cat("ATT: ", att_estimate_cem, "\n",
    "SE: ", se_cem, "\n",
    "95% CI: [", ci_lower_cem, ", ", ci_upper_cem, "]", sep_nn = "")

## ATT: -0.1261722
## SE: 0.06176064
## 95% CI: [ -0.2472208 , -0.005123568 ]
```

Implementing the Coarsened Exact Matching algorithm we have estimated ATT of TRC exposure on respondents' racial attitudes to be -0.1262. Further, the Standard Error of ATT is estimated at 0.0618 and Confidence interval is [-0.2472, -0.0051]. We can interpret this as being 95% confident that the true ATE lies within this interval, and since 0 is not in the interval we can say that the effect is statistically significant at the 5% level.

#### **part d (10 points)**

Compare and contrast the three different matching algorithms. Provide evidence and an argument about which one we should use.

The first method we used (part a), Nearest Neighbor Matching, works by pairing each treated unit with a control unit with the closest propensity score. This method is quite simple and efficient to use but because of that can have bad matches if there are not enough close control units for each treated unit. In our case, this method resulted in ATT of -0.1815, corresponding SE of 0.0454 and CI of [-0.2705, -0.0924]. This method did provide us with a statistically significant estimate of the ATT (as 0 is not in the interval) and the interval is relatively narrow. In other words we can be fairly certain of our estimated negative TRC (-0.2154) exposure on racial attitudes. This means that in general, those with TRC exposure find it less difficult to understand the customs and ways of the opposite racial group.

The second method we used (part b), Exact Matching, works by pairing treatment and control units that have identical values on all covariates. However, as we saw in the King Charles example in class, this might not always be the most optimal method. While there can be good matches due to same values on all covariates, there is also a large loss in data as very few units will actually match on all covariates (especially if we have a lot of them). This might be what happened in our case as we got very different results than in all previous (and following) tests. Exact Matching method resulted in point estimate of ATT being 0.1027, corresponding SE of 0.1787 and CI of [-0.2478, 0.4530]. We can see that this standard error is much higher than our previous one and the CI is much wider which means the result is much less precise.. Further, the result is non-significant estimate as 0 is in the CI. As mentioned this could be happening because Exact Matching only pairs values with all identical covariates and as we have quite a few covariates it is likely that there are few perfect matches thus leading to low accuracy.

The third method we used (part c), Coarsened Exact Matching, is a method that groups treatment and control units based on covariates and matches within these categories. In a way it is a compromise between previous two methods and it is used to reduce the model dependence and improve the balance of it. As such it usually has better results than Exact Matching but is much harder (and potentially more expensive - time and computational power wise) to implement. This approach resulted in point estimate of ATT of -0.1262, corresponding SE of 0.0618, and CI of [-0.2472, -0.0051]. We can notice that this result is statistically significant as 0 is not in the CI and our CI is much narrower than in previous part for Exact Matching and that the standard error is also lower than for Exact Matching. However, we can also notice that SE is somewhat larger for Coarsened Exact Matching and CI is wider in comparison to Nearest Neighbor Matching. This would indicate Nearest Neighbor Matching is a more precise method for our case.

In our case both Nearest Neighbor Matching and Coarsened Exact Matching would be good and acceptable choices as they produce statistically significant results with low SE. However, while Nearest Neighbor Matching does have more narrow CI and slightly lower SE, it is more unstable than Coarsened Exact Matching. On the other hand, Coarsened Exact Matching might be more balanced and might reduce model dependence, but it's also more complex which could bring problems on its own. As always we have a trade off between model quality and cost/complexity. I believe both of these would be a good choice in our case as they produce very similar results but if we are foreseeing more 'spread apart' units that Nearest Neighbor Matching would struggle with or need a more balanced model we should choose Coarsened Exact Matching even though it is more complex.

## BONUS ONLY: Question 3 (Total: Up to +12)

Question 3 is for bonus points. (See forthcoming lecture on Nov. 7th)

### part a (+4 points)

Using the regression method to predict potential outcomes for all individuals in the dataset and calculate the ATE with bootstrapped standard errors. Report and interpret your results. (Hint: Start by fitting the treatment and control model with subsets of the data.)

```
## Fit a model among TRCKNOW == 1 to get  $E[Y_i(1) | X]$ 
treatment_model <- lm_robust(RUSTAND ~ age + female + wealth + religiosity +
  ethsalience + rcblack + rcwhite + rccol + EDUC,
  data = subset(TRC_data, TRCKNOW==1))

## Fit a model among TRCKNOW == 0 to get  $E[Y_i(0) | X]$ 
control_model <- lm_robust(RUSTAND ~ age + female + wealth + religiosity +
  ethsalience + rcblack + rcwhite + rccol + EDUC,
  data = subset(TRC_data, TRCKNOW==0))

## Predict the potential outcome under treatment for all units
TRC_data$trc_treated <- predict(treatment_model, newdata = TRC_data )

## Predict the potential outcome under control for all units
TRC_data$trc_control <- predict(control_model, newdata = TRC_data )

## Average of the differences
point <- mean( TRC_data$trc_treated - TRC_data$trc_control)

### Bootstrap for SEs
set.seed(123)

nBoot <- 2000 # Number of iterations
boot_results <- rep(NA, 2000)

for (iter in 1:nBoot){
  # Resample w/ replacement
  trc_boot <- TRC_data[sample(1:nrow(TRC_data), nrow(TRC_data) , replace=T) ,]
  ## Fit a model among nsw == 1 to get  $E[Y_i(1) | X]$ 
  treatment_model_boot <- lm_robust(RUSTAND ~ age + female + wealth + religiosity +
    ethsalience + rcblack + rcwhite + rccol + EDUC,
    data = subset(trc_boot, TRCKNOW==1))
  ## Fit a model among nsw == 0 to get  $E[Y_i(0) | X]$ 
  control_model_boot <- lm_robust(RUSTAND ~ age + female + wealth + religiosity +
    ethsalience + rcblack + rcwhite + rccol + EDUC,
    data = subset(trc_boot, TRCKNOW==0))

  ## Predict the potential outcome under treatment for all units
  trc_boot$trc_treated_boot <- predict(treatment_model_boot, newdata = trc_boot )

  ## Predict the potential outcome under control for all units
  trc_boot$trc_control_boot <- predict(control_model_boot, newdata = trc_boot )
}
```

```

## Store bootstrapped estimate
boot_results[iter]<- mean(trc_boot$trc_treated_boot - trc_boot$trc_control_boot)
}

### Standard error
se_hat = sd(boot_results)

#se_hat
cat("Point ATE Estimate: ", point, "\n",
    "Bootstrap SE: ", se_hat)

## Point ATE Estimate: -0.1743866
## Bootstrap SE: 0.04465872

### 95% confidence interval
c(point - 1.96*se_hat, point + 1.96*se_hat)

## [1] -0.26191769 -0.08685551

```

We can see that the point estimate for the ATE is -0.1744. This represents the average difference in the respondents' racial attitudes between those who are knowledgeable about the TRC (the treatment group) and those who are not (the control group). As in previous questions we can say that the negative value indicates that on average, respondents who know about the TRC have racial attitudes that are less difficult or more understanding towards the customs and ways of the opposite racial group than those who do not know about the TRC. Our bootstrapped standard error is about 0.0447. This value represents the variability/uncertainty around our ATE estimate. Our SE is relatively small in comparison to the ATE which would suggest that the estimated effect of the TRC on racial attitudes is statistically distinguishable from zero (assuming the usual critical values for statistical significance). Thus, our results suggest that there is a statistically significant association between knowledge of the TRC and more positive or understanding racial attitudes.

## part b (+4 points)

Using the regression method to predict potential outcomes for all individuals and calculate the ATT with bootstrapped standard errors. Report and interpret your results.

```

## Fit a model among TRCKNOW == 1 to get E[Y_i(1) | X]
treatment_model <- lm_robust(RUSTAND ~ age + female + wealth + religiosity +
    ethsalience + rcblack + rcwhite + rccol + EDUC,
    data = subset(TRC_data, TRCKNOW==1))

## Fit a model among TRCKNOW == 0 to get E[Y_i(0) | X]
control_model <- lm_robust(RUSTAND ~ age + female + wealth + religiosity +
    ethsalience + rcblack + rcwhite + rccol + EDUC,
    data = subset(TRC_data, TRCKNOW==0))

## Predict the potential outcome under treatment for all units
TRC_data$trc_treated <- predict(treatment_model, newdata = TRC_data )

## Predict the potential outcome under control for all units
TRC_data$trc_control <- predict(control_model, newdata = TRC_data )

```

```

## Average of the differences
point_ATT <- mean( TRC_data$trc_treated[TRC_data$TRCKNOW==1] -
                  TRC_data$trc_control[TRC_data$TRCKNOW==1])

### Bootstrap for SEs
set.seed(123)

nBoot <- 2000 # Number of iterations
boot_results_ATT <- rep(NA, 2000)

for (iter in 1:nBoot){
  # Resample w/ replacement
  treated_boot<- TRC_data[sample(1:nrow(TRC_data),
                                nrow(TRC_data), replace = T), ]

  ## Fit a model among nsw == 1 to get E[Y i(1) | X]
  treatment_model_boot <- lm_robust(RUSTAND ~ age + female + wealth + religiosity +
                                   ethsalience + rcblack + rcwhite + rccol + EDUC,
                                   data = subset(treated_boot, TRCKNOW==1))

  ## Fit a model among nsw == 0 to get E[Y i(0) | X]
  control_model_boot <- lm_robust(RUSTAND ~ age + female + wealth + religiosity +
                                   ethsalience + rcblack + rcwhite + rccol + EDUC,
                                   data = subset(treated_boot, TRCKNOW==0))

  ## Predict the potential outcome under treatment for all units
  treated_boot$trc_treated_boot <- predict(treatment_model_boot, newdata = treated_boot )

  ## Predict the potential outcome under control for all units
  treated_boot$trc_control_boot <- predict(control_model_boot, newdata = treated_boot )

  ## Store bootstrapped estimate
  boot_results_ATT[iter]<- mean(treated_boot$trc_treated_boot[treated_boot$TRCKNOW==1] -
                              treated_boot$trc_control_boot[treated_boot$TRCKNOW==1])
}

### Standard error
se_hat_ATT = sd(boot_results_ATT)

#se_hat
cat("Point ATT Estimate: ", point_ATT, "\n",
    "Bootstrap SE: ", se_hat_ATT)

## Point ATT Estimate: -0.2033737
## Bootstrap SE: 0.04641271

### 95% confidence interval
c(point_ATT - 1.96*se_hat_ATT, point_ATT + 1.96*se_hat_ATT)

## [1] -0.2943426 -0.1124048

```

We can see that the point estimate for the ATT is slightly smaller than ATE at -0.2034. This means that on average for individuals knowledgeable about the TRC their racial attitudes are reduced by about 0.2034,

compared to if they were unaware. As in previous questions we can say that the negative value indicates that on average, respondents who know about the TRC have racial attitudes that are less difficult or more understanding towards the customs and ways of the opposite racial group than those who do not know about the TRC. Our bootstrapped standard error is about 0.0464. This value represents the variability/uncertainty around our ATT estimate. Our SE is relatively small in comparison to the ATT but larger than that of ATE which would suggest that the estimated effect of the TRC on racial attitudes is statistically distinguishable from zero (assuming the usual critical values for statistical significance) but that our estimate is slightly less precise. Thus, our results suggest that there is a statistically significant association between knowledge of the TRC and more positive or understanding racial attitudes.

**part c (+4 points)**

Compare and contrast the ATE and ATT from the regression approach.

Both of these statistics measure the impact of our treatment (exposure to TRCs) on an outcome variables but they refer to effects on different groups within our study. ATE represents the Average Treatment Effect of the treatment on all individuals in the study, regardless of whether they received the treatment or not. Our ATE is estimated at -0.1744, with a standard error of 0.0447, this indicates that on average knowledge about TRCs leads to a decrease in the difficulty of understanding the customs and ways of the opposite racial group by 0.1744 points on the scale used to measure racial attitudes. As we can notice, our SE for ATE is slightly smaller than that of ATT which means that there is less variability in the estimated effect across the entire population and slightly more uncertainty around the estimate for the treated group. On the other hand, ATT, Average Treatment Effect on the Treated represents the average effect of the treatment (exposure to TRCs) on those individuals who actually received the treatment (individuals who have knowledge of TRCs). Our ATT is estimated to be -0.2034, with a standard error of 0.0464, which again indicates that those who are knowledgeable about the TRCs, on average have more understanding racial attitudes by 0.2034 points on the racial attitude scale compared to what their attitudes would have been without this knowledge. As mentioned, ATT has larger SE than ATE which indicates that the estimate of the effect among the treated individuals is less precise than the ATE. As ATT is smaller (but actually larger in absolute value since it is negative) than ATE it indicates that the effect of TRC knowledge on racial attitudes is stronger among those who are aware of the TRCs than it is across the entire population. Generally, ATE is used for understanding the overall policy/treatment impact, while the ATT is more specific to the subpopulation that is actually exposed to the treatment. Further, the difference in our estimates might be due to many different factors but we can definitely assume it is partially due to the fact that exposure to TRC may be affecting individuals differently and they might be exposed to different amounts. Additionally, as always there may be many more personal factors that affect the relationship between TRC knowledge and racial attitudes differently across the treated and untreated groups.