Problem Set 2

Anita Ye - yy3557 - Section 003

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 yourNetID_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.)

```r
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")

tidy(difference_in_means(RUSTAND ~ TRCKNOW, data=TRC_data))
```

```
##      term   estimate  std.error statistic    p.value  conf.low  conf.high
## 1 TRCKNOW -0.2177317 0.04433111 -4.911488 9.533946e-07 -0.304655 -0.1308085
##        df outcome
## 1 2934.442 RUSTAND
```

```r
model = lm_robust(RUSTAND ~ TRCKNOW, data=TRC_data)
tidy(model)
```

```
##          term   estimate  std.error statistic    p.value  conf.low  conf.high
## 1 (Intercept)  2.5311438 0.02805761 90.212369 0.000000e+00  2.476131  2.5861565
## 2     TRCKNOW -0.2177317 0.04433111 -4.911488 9.491614e-07 -0.304652 -0.1308115
##      df outcome
## 1 3203 RUSTAND
## 2 3203 RUSTAND
```

```
#summary(model)
#print(model)
coef(model)["TRCKNOW"]
```

```
##     TRCKNOW
## -0.2177317
```

```
conf_interval <- confint(model, level = 0.95)["TRCKNOW", ]
```

The confidence interval (-0.304652, -0.1308115) does not include zero, which indicates that the treatment effect is statistically significant. The estimated ATE of TRC exposure on respondents' racial attitudes is -0.2177, and this negative ATE suggests that on average exposure to the TRC is associated with a statistically significant 0.2177 decrease in respondents' agreement with the statement 'I find it difficult to understand the customs and ways of the opposite racial group.

A negative result indicate that exposure to TRC is associated with more positive racial attitudes, and so people with the knowledge of TRC are more likely to understand opposite racial groups. The lower the score, the less difficulty.

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?

```
TRC_data <- TRC_data %>%
  mutate(age_std = age/sd(age),
         female_std = female/sd(female),
         wealth_std = wealth/sd(wealth),
         religiosity_std = religiosity/sd(religiosity),
         ethsalience_std = ethsalience/sd(ethsalience),
         rcwhite_std = rcwhite/sd(rcwhite),
         rccol_std = rccol/sd(rccol),
         EDUC_std = EDUC/sd(EDUC),
         rcblack_std = rcblack/sd(rcblack))


balance_table <- TRC_data %>% group_by(TRCKNOW) %>% summarize(age_std = mean(age_std),
         wealth_std = mean(wealth_std),
         female_std = mean(female_std),
         religiosity_std = mean(religiosity_std),
         ethsalience_std = mean(ethsalience_std),
         rcwhite_std = mean(rcwhite_std),
         rccol_std = mean(rccol_std),
         EDUC_std = mean(EDUC_std),
         rcblack_std = mean(rcblack_std),
         .groups = "keep")
print(balance_table)
```

```
## # A tibble: 2 x 10
## # Groups:   TRCKNOW [2]
##   TRCKNOW age_std wealth_std female_std religiosity_std ethsalience_std
##     <dbl>   <dbl>      <dbl>      <dbl>           <dbl>           <dbl>
## 1       0    2.62      0.774      0.866            2.15            4.69
## 2       1    2.52      0.928      1.08             2.11            4.73
## # i 4 more variables: rcwhite_std <dbl>, rccol_std <dbl>, EDUC_std <dbl>,
## #   rcblack_std <dbl>
```

```r
abs_balance_diff <- abs(balance_table[1, 2:ncol(balance_table)] -
balance_table[2, 2:ncol(balance_table)])
abs_balance_diff
```

```
##     age_std wealth_std female_std religiosity_std ethsalience_std rcwhite_std
## 1 0.0980385  0.1539068  0.2106527      0.04139381      0.03641136  0.03762548
##   rccol_std  EDUC_std rcblack_std
## 1 0.1367213 0.3840434  0.07762829
```

Exposed respondents tend to be more female, have higher wealth, higher ethsalience, and a higher percentage of rcblack and rcwhite compared to nonexposed respondents. On the other hand, nonexposed respondents tend to be older and less religious compared to exposed respondents.

Minimal absolute differences, approaching zero, suggest that the standardized covariates exhibit a good balance between the treatment and control groups. For this scenario, there isn't a significant disparity in religiosity and ethsalience between the exposed and unexposed respondents. However, notable variations are observed in aspects like age, gender, wealth, race, and education level.

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

```r
# Fit a logistic regression model to estimate the propensity scores
propensity_model <- glm(TRCKNOW ~ age + female + wealth + religiosity +
                          ethsalience + rcblack + rcwhite + rccol + EDUC,
                        data = TRC_data,
                        family = "binomial"(link="logit"))

model_summary <- summary(propensity_model)

#tidy(propensity_model)

# Extract coefficient estimates, standard errors, z-values, and p-values
coefficients <- model_summary$coefficients[, "Estimate"]
standard_errors <- model_summary$coefficients[, "Std. Error"]
z_values <- model_summary$coefficients[, "z value"]
p_values <- model_summary$coefficients[, "Pr(>|z|)"]

# Create a data frame with the results
coef_summary <- data.frame(Coefficient = coefficients, Standard_Error = standard_errors,
                           Z_Value = z_values, P_Value = p_values)

# Print the summary
print(coef_summary)
```

```
##               Coefficient Standard_Error    Z_Value      P_Value
## (Intercept) -2.515274e+00   3.132652e-01 -8.0292164 9.809722e-16
## age          3.706434e-04   2.538394e-03  0.1460149 8.839096e-01
## female       3.881911e-01   7.505673e-02  5.1719700 2.316388e-07
```

```
## wealth       2.435941e-05   6.848309e-06  3.5569967 3.751188e-04
## religiosity  1.128748e-02   2.088595e-02  0.5404342 5.888976e-01
## ethsalience  6.014118e-02   6.502776e-02  0.9248539 3.550419e-01
## rcblack      4.721256e-01   1.522705e-01  3.1005708 1.931480e-03
## rcwhite     -2.795784e-01   1.631277e-01 -1.7138623 8.655402e-02
## rccol       -2.154211e-01   1.711909e-01 -1.2583677 2.082588e-01
## EDUC         3.924675e-01   3.958923e-02  9.9134940 3.637000e-23
```

```r
# Get the estimated propensity scores
TRC_data$propensity_scores <- predict(propensity_model, type = "response")
head(TRC_data$propensity_scores)
```

```
##         1         2         3         4         5         6
## 0.5388899 0.4457546 0.4331449 0.4339534 0.2674423 0.4512316
```

```r
# Generate the weights (unstabilized weights)
TRC_data$wt <- NA
TRC_data$wt[TRC_data$TRCKNOW == 1] <- 1/TRC_data$propensity_scores[TRC_data$TRCKNOW==1]
TRC_data$wt[TRC_data$TRCKNOW == 0] <- 1/(1 - TRC_data$propensity_scores[TRC_data$TRCKNOW==0])
point_wtd <- mean(TRC_data$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 indicates TRC is associated with on average a 0.1623 decrease in hostile racial attitudes.

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.)

```r
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){
  TRC_data_boot <- TRC_data[sample(1:nrow(TRC_data), nrow(TRC_data), replace = TRUE), ]

  # Fit logistic regression model on the bootstrap sample
  logistic_boot <- glm(TRCKNOW ~ age + female + wealth + religiosity + ethsalience + rcblack +
                         rcwhite + rccol + EDUC,
                       data = TRC_data_boot,
                       family = binomial(link = "logit"))

  # Calculate propensity scores on the bootstrap sample
  TRC_data_boot$predict_score <- predict(logistic_boot, type = "response")

  # Calculate the weights
  TRC_data_boot$w <- NA
  TRC_data_boot$wt[TRC_data_boot$TRCKNOW == 1] <-
    1/TRC_data_boot$predict_score[TRC_data_boot$TRCKNOW==1]
  TRC_data_boot$wt[TRC_data_boot$TRCKNOW == 0] <-
    1/(1 - TRC_data_boot$predict_score[TRC_data_boot$TRCKNOW==0])
  # Compute and store the ATE
```

```
  ate_boot[boot] <- mean(TRC_data_boot$wt * TRC_data_boot$RUSTAND * TRC_data_boot$TRCKNOW -
  TRC_data_boot$wt * TRC_data_boot$RUSTAND * (1-TRC_data_boot$TRCKNOW))
}
cat("Bootstrap Estimate of ATE:", mean(ate_boot), "\n")
```

```
## Bootstrap Estimate of ATE: -0.1596519
```

```
# Take the SD of the ate_boot to get our estimated SE - can do asymptotic inference
cat("Bootstrap Standard Error: \n", sd(ate_boot), "\n")
```

```
## Bootstrap Standard Error:
##  0.04534277
```

```
# Asymptotic 95\% CI
ci = c(point_wtd - qnorm(.975)*sd(ate_boot),
point_wtd + qnorm(.975)*sd(ate_boot))
cat("Bootstrap 95% Confidence Interval: \n", ci, "\n")
```

```
## Bootstrap 95% Confidence Interval:
##  -0.2511992 -0.07345878
```

```
# Take quantiles to get CIs directly from the bootstrapped distribution (esp. if skewed)
quantile(ate_boot, c(.025, .975))
```

```
##        2.5%       97.5%
## -0.24860000 -0.06970394
```

```
#quantile is NOT CI
```

Using the bootstrap method, we estimated the standard error of the IPTW estimator for the ATE to be approximately 0.0453, and this value quantifies the variability in the ATE estimate due to sampling.

The average of the ATE estimates from the bootstrap samples is -0.1596519. This value suggests that, on average, exposure to the TRC is associated with a decrease in respondents' agreement with the statement "I find it difficult to understand the customs and ways of the opposite racial group."

Comparing the result from part d to the ATE estimated in part a (-0.2177), this decrease is slightly less in magnitude compared to the -0.2177 estimate from Part A. We observe a slight rightward shift in the point estimate, which suggests that the IPTW estimator may provide a more accurate representation of the treatment effect. In conclusion, there is a small difference between the values from the two parts that may not be significant enough, but both estimates are negative and significant, indicating a consistent finding that exposure to the TRC is associated with more positive racial attitudes.

The 95% confidence interval for the ATE is (-0.2512, -0.0735). Since this interval does not include zero, it indicates that the treatment effect of TRC exposure remains statistically significant even after considering the uncertainty introduced by the bootstrap resampling. The confidence interval overlaps, indicating a statistical consistency between the estimates. The CI from part D compared to part A is narrower, which indicates a more precise estimate due to bootstrapping. Through bootstrapping, we can get a more precise estimate and a more robust result as there are fewer assumptions about distribution of the estimator.

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.

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

m.out1 = matchit(TRCKNOW ~ age + female + wealth + religiosity + ethsalience + rcblack +
                   rcwhite + rccol+ EDUC,
                 data = TRC_data,
                 method = "nearest",
                 distance = "glm")
#summary(m.out1)
m.data1 <- match.data(m.out1)
#head(m.data1)
model = lm_robust(RUSTAND ~ age + female + wealth + religiosity + ethsalience + rcblack +
                    rcwhite + rccol+ EDUC + TRCKNOW,
                  data = m.data1)
tidy(model)
```

```
##          term       estimate    std.error  statistic      p.value       conf.low
## 1  (Intercept)  2.863798e+00 1.924724e-01 14.8790020 2.730161e-48  2.486399e+00
## 2          age  2.124821e-04 1.565603e-03  0.1357190 8.920530e-01 -2.857340e-03
## 3       female -1.224397e-01 4.632683e-02 -2.6429544 8.263332e-03 -2.132769e-01
## 4        wealth -1.228282e-05 4.332807e-06 -2.8348408 4.616964e-03 -2.077855e-05
## 5  religiosity -1.812679e-02 1.267233e-02 -1.4304225 1.527048e-01 -4.297459e-02
## 6  ethsalience  1.799750e-02 3.994206e-02  0.4505901 6.523191e-01 -6.032057e-02
## 7       rcblack  4.221736e-01 9.376444e-02  4.5024918 6.984141e-06  2.383211e-01
## 8       rcwhite  1.093332e-01 9.596039e-02  1.1393574 2.546493e-01 -7.882516e-02
## 9        rccol -1.079063e-01 1.038701e-01 -1.0388580 2.989584e-01 -3.115740e-01
## 10        EDUC -1.065639e-01 2.402210e-02 -4.4360766 9.506390e-06 -1.536662e-01
## 11     TRCKNOW -1.802434e-01 4.545871e-02 -3.9649914 7.519278e-05 -2.693784e-01
##      conf.high   df outcome
## 1   3.241196e+00 2867 RUSTAND
## 2   3.282305e-03 2867 RUSTAND
## 3  -3.160243e-02 2867 RUSTAND
## 4  -3.787086e-06 2867 RUSTAND
## 5   6.721014e-03 2867 RUSTAND
## 6   9.631556e-02 2867 RUSTAND
## 7   6.060262e-01 2867 RUSTAND
## 8   2.974915e-01 2867 RUSTAND
## 9   9.576137e-02 2867 RUSTAND
## 10 -5.946155e-02 2867 RUSTAND
## 11 -9.110832e-02 2867 RUSTAND
```

```r
att <- coef(model)["TRCKNOW"]
cat("ATT: ", att)
```

```
## ATT:  -0.1802434
```

ATT: -0.1802434

95% CI (-2.693784e-01 -9.110832e-02)

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.

```
m.out2 = matchit(TRCKNOW ~ age + female + wealth + religiosity + ethsalience + rcblack +
                    rcwhite + rccol+ EDUC,
                 data = TRC_data,
                 method = "exact",
                 distance = "glm")
#summary(m.out2)
m.data2 <- match.data(m.out2)
#head(m.data2)
model = lm_robust(RUSTAND ~ age + female + wealth + religiosity + ethsalience + rcblack +
                    rcwhite + rccol+ EDUC + TRCKNOW,
                 data = m.data2)
tidy(model)
```

```
##          term      estimate    std.error  statistic      p.value       conf.low
## 1  (Intercept)  5.339649e+00 9.933369e-01  5.3754662 2.505836e-07   3.3787019714
## 2          age  1.951547e-03 9.940822e-03  0.1963165 8.445983e-01  -0.0176726334
## 3       female  1.644171e-01 2.057038e-01  0.7992906 4.252439e-01  -0.2416628433
## 4       wealth  1.227178e-05 2.953197e-05  0.4155422 6.782722e-01  -0.0000460273
## 5   religiosity -9.708998e-02 6.069108e-02 -1.5997406 1.115239e-01  -0.2169002704
## 6   ethsalience -4.504668e-01 3.575883e-01 -1.2597358 2.095016e-01  -1.1563821048
## 7       rcblack -1.213618e-01 4.595932e-01 -0.2640635 7.920527e-01  -1.0286448919
## 8       rcwhite -6.978282e-01 3.429611e-01 -2.0347156 4.344219e-02  -1.3748677894
## 9         rccol -1.380816e-01 7.976226e-01 -0.1731164 8.627671e-01  -1.7126687326
## 10         EDUC -2.433734e-01 1.423939e-01 -1.7091555 8.925777e-02  -0.5244733483
## 11      TRCKNOW  1.027187e-01 1.787245e-01  0.5747322 5.662367e-01  -0.2501013981
##       conf.high  df outcome
## 1   7.300595e+00 169 RUSTAND
## 2   2.157573e-02 169 RUSTAND
## 3   5.704971e-01 169 RUSTAND
## 4   7.057087e-05 169 RUSTAND
## 5   2.272030e-02 169 RUSTAND
## 6   2.554484e-01 169 RUSTAND
## 7   7.859213e-01 169 RUSTAND
## 8  -2.078863e-02 169 RUSTAND
## 9   1.436506e+00 169 RUSTAND
## 10  3.772655e-02 169 RUSTAND
## 11  4.555389e-01 169 RUSTAND
```

```
att <- coef(model)["TRCKNOW"]
cat("ATT: ", att)
```

```
## ATT:  0.1027187
```

ATT: 0.1027187

95% CI: (-0.2501013981 4.555389e-01)

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.

```
m.out3 = matchit(TRCKNOW ~ age + female + wealth + religiosity + ethsalience + rcblack
                    + rcwhite + rccol+ EDUC,
                 data = TRC_data,
```

```
                method = "cem",
                distance = "glm")
#summary(m.out3)
m.data3 <- match.data(m.out3)
#head(m.data3)
model = lm_robust(RUSTAND ~ age + female + wealth + religiosity + ethsalience + rcblack +
                  rcwhite + rccol+ EDUC + TRCKNOW,
                data = m.data3)
tidy(model)
```

```
##          term      estimate    std.error  statistic      p.value       conf.low
## 1  (Intercept)  3.170812e+00 4.004629e-01  7.9178660 4.630727e-15   2.3852920172
## 2          age  1.604415e-03 2.560838e-03  0.6265196 5.310684e-01  -0.0034187421
## 3       female -1.077344e-01 6.444329e-02 -1.6717705 9.477582e-02  -0.2341417805
## 4       wealth -1.660956e-05 9.169960e-06 -1.8113011 7.029196e-02  -0.0000345967
## 5  religiosity -1.693928e-02 1.968193e-02 -0.8606517 3.895659e-01  -0.0555459540
## 6  ethsalience  2.672914e-02 8.404640e-02  0.3180283 7.505072e-01  -0.1381303092
## 7      rcblack -3.070173e-02 2.349849e-01 -0.1306541 8.960663e-01  -0.4916313644
## 8      rcwhite -3.336013e-01 2.241358e-01 -1.4883893 1.368561e-01  -0.7732502536
## 9        rccol -7.255778e-01 2.411758e-01 -3.0085013 2.668642e-03  -1.1986512005
## 10        EDUC -8.750641e-02 4.634598e-02 -1.8881122 5.920157e-02  -0.1784153933
## 11     TRCKNOW -1.261722e-01 6.176064e-02 -2.0429225 4.123263e-02  -0.2473174864
##         conf.high   df outcome
## 1   3.956331e+00 1517 RUSTAND
## 2   6.627572e-03 1517 RUSTAND
## 3   1.867299e-02 1517 RUSTAND
## 4   1.377583e-06 1517 RUSTAND
## 5   2.166739e-02 1517 RUSTAND
## 6   1.915886e-01 1517 RUSTAND
## 7   4.302279e-01 1517 RUSTAND
## 8   1.060476e-01 1517 RUSTAND
## 9  -2.525044e-01 1517 RUSTAND
## 10  3.402573e-03 1517 RUSTAND
## 11 -5.026912e-03 1517 RUSTAND
```

```
att <- coef(model)["TRCKNOW"]
cat("ATT: ", att)
```

```
## ATT:  -0.1261722
```

ATT: -0.1261722

95% CI: (-0.2473174864 -5.026912e-03)

part d (10 points)

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

K Nearest Neighbor (KNN) matching selects the K closest units to a specific unit i. It pairs each treated unit with the control unit having the nearest propensity score. The nearest neighbor matching yields an estimated average treatment effect on the treated (ATT) of approximately -0.1802, with a standard error of about 0.04546 and 95% CI : (-2.693784e-01 -9.110832e-02), indicating a statistically significant negative treatment effect. However, it leaves 327 control units unmatched.

Exact matching pairs units based on identical covariate values. The resulting distribution of covariates is identical for treated and matched controls. This method estimates an ATT of approximately 0.1027, with

a larger standard error of about 0.1787 and 95% CI: (-0.2501013981 4.555389e-01), suggesting a positive treatment effect but not statistical significant since 0 is in the range. It leaves a significant number of control (1674) and treated units (1351) unmatched.

Coarsened exact matching (CEM) involves binning covariates and then matching on these coarsened versions. It estimates an ATT of approximately -0.1262, with a standard error of about 0.06176 and 95% CI: (-0.2473174864 -5.026912e-03) , indicating a statistically significant negative treatment effect. This method leaves 964 control units and 713 treated units unmatched. This result indicates a consistently negative effect similar to nearest neighbor matching, but with a narrower range than exact matching.

Comparing the three, nearest neighbor and coarsened exact matching are preferable as they both show a statistically significant negative treatment effect with greater certainty and narrower 95% confidence intervals. The exact matching, while ensuring balance on matched covariates, often results in a smaller sample and thus wider confidence intervals, which make the result uncertain, and based on the CI also not statistically significant, meaning that we could not make any conclusion about the treatment effect. The choice between nearest neighbor and CEM depends on whether higher precision (nearest neighbor) or less biased samples (CEM) is prioritized.

Nearest neighbor and coarsened exact matching have more precise estimates than exact matching, because of their narrower confidence intervals. These methods also indicate a consistent negative effect of treatment on the outcome, while exact matching's wider interval includes the possibility of a positive effect. Therefore, nearest neighbor or coarsened exact matching are more effective for robust estimations, with the choice depending on specific analysis requirements, such as the balance between homogeneity and sample size.

If I have to pick one, I would choose nearest neighbor. Exact matching does not have large enough sample size and is statistically insignificant to provide evidence on the treatment effect, so it is not an option. The result of CEM is close to that of nearest neighbor and both are statistically significant, but nearest neighbor has a larger sample size, which makes the estimate result more precise. Meanwhile, cem has less biased sample, so it can also be an option if needed.

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.)

```r
# Load required packages
library(boot)

# Fit models for treatment and control groups
## Fit a model among nsw == 1 to get E[Y_i(1) | X]
model_treatment <- lm(RUSTAND ~ age + female + wealth + religiosity + ethsalience + rcblack +
                        rcwhite + rccol + EDUC,
                    data = TRC_data[TRC_data$TRCKNOW == 1, ])
## Fit a model among nsw == 0 to get E[Y_i(0) | X]
model_control <- lm(RUSTAND ~ age + female + wealth + religiosity + ethsalience + rcblack +
                        rcwhite + rccol + EDUC,
                    data = TRC_data[TRC_data$TRCKNOW == 0, ])

# Predict potential outcomes for all units
## Predict the potential outcome under treatment for all units
Y1_hat <- predict(model_treatment, newdata = TRC_data)
## Predict the potential outcome under control for all units
```

```r
Y0_hat <- predict(model_control, newdata = TRC_data)

# Calculate ATE
## Average of the differences
### Bootstrap for SEs
set.seed(123)
ATE <- mean(Y1_hat - Y0_hat)

# Bootstrap for SEs
set.seed(123)
n_iter <- 1000 # Suggested number of iterations
iter_result1 = rep(NA, 1000)
for (iter in 1:n_iter) {
  TRC_data_boot = TRC_data[sample(1:nrow(TRC_data), nrow(TRC_data), replace = T),]
  treatment_model_boot = lm_robust(RUSTAND ~ age + female + wealth + religiosity +
                                     ethsalience + rcblack + rcwhite + rccol + EDUC,
                                   data = subset(TRC_data_boot, TRCKNOW == 1))
  control_model_boot = lm_robust(RUSTAND ~ age + female + wealth + religiosity +
                                   ethsalience + rcblack + rcwhite + rccol + EDUC,
                                 data = subset(TRC_data_boot, TRCKNOW == 0))
  TRC_data_boot$treat_boot = predict(treatment_model_boot, newdata = TRC_data_boot)
  TRC_data_boot$control_boot = predict(control_model_boot, newdata = TRC_data_boot)
#store estimate
  iter_result1[iter] <- mean(TRC_data_boot$treat_boot - TRC_data_boot$control_boot)
}
# standard errors
ATE_se <- sd(iter_result1)

# Report and interpret ATE
cat("ATE:", ATE, "\n")
```

```
## ATE: -0.1743866
```

```r
cat("Bootstrap SE for ATE:", ATE_se, "\n")
```

```
## Bootstrap SE for ATE: 0.04359436
```

```r
CI1_95 = c(ATE - qnorm(0.975)*sd(iter_result1), ATE + qnorm(0.975)*sd(iter_result1))
cat("CI:", CI1_95)
```

```
## CI: -0.25983 -0.08894323
```

point estimate: -0.1743866 bootstrap standard error: 0.04359436 CI: -0.25983 -0.08894323

The bootstrap standard error of approximately 0.0436 represents the level of uncertainty in the ATE estimate. It suggests that there is some variability in the estimated treatment effect across different samples, but not very high.

The point estimate of the ATE suggests that exposure to TRC is associated with an estimated decrease of approximately 0.1744 in respondents' agreement with the statement "I find it difficult to understand the customs and ways of the opposite racial group." This decrease is statistically significant since the 95% confidence interval (-0.2598, -0.0889) does not include zero. It implies that the TRC exposure has a meaningful impact on reducing racial attitudes, with respondents becoming less likely to find it difficult to understand customs and ways of the opposite racial group.

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.

```r
# Fit models for treatment and control groups (similar to part a)
set.seed(123)
## Fit a model among nsw == 1 to get E[Y_i(1) | X]
model_treatment <- lm_robust(RUSTAND ~ age + female + wealth + religiosity + ethsalience +
                             rcblack + rcwhite + rccol + EDUC,
                             data = TRC_data[TRC_data$TRCKNOW == 1, ])
## Fit a model among nsw == 0 to get E[Y_i(0) | X]
model_control <- lm_robust(RUSTAND ~ age + female + wealth + religiosity + ethsalience +
                           rcblack + rcwhite + rccol + EDUC,
                           data = TRC_data[TRC_data$TRCKNOW == 0, ])

# Predict potential outcomes for all units (similar to part a)
## Predict the potential outcome under treatment for all units
Y1_hat <- predict(model_treatment, newdata = TRC_data)
## Predict the potential outcome under control for all units
Y0_hat <- predict(model_control, newdata = TRC_data)

# Calculate ATT
# ATT <- mean(Y1_hat - Y0_hat)
## Average of the differences
point_est_att <- mean(Y1_hat[TRC_data$TRCKNOW == 1] - Y0_hat[TRC_data$TRCKNOW == 1])


set.seed(123)
# Bootstrap for SEs (similar to part a)
n_iter <- 1000  # Suggested number of iterations
iter_result2 <- rep(NA, n_iter)

for (iter in 1:n_iter) {
  TRC_data_boot <- TRC_data[sample(1:nrow(TRC_data), nrow(TRC_data), replace = TRUE), ]

  # Fit treatment and control models on the bootstrapped data
  treatment_model_boot <- lm_robust(RUSTAND ~ age + female + wealth + religiosity +
                                     ethsalience + rcblack + rcwhite + rccol + EDUC,
                                     data = subset(TRC_data_boot, TRCKNOW == 1))
  control_model_boot <- lm_robust(RUSTAND ~ age + female + wealth + religiosity +
                                   ethsalience + rcblack + rcwhite + rccol + EDUC,
                                   data = subset(TRC_data_boot, TRCKNOW == 0))

  # Predict potential outcomes for the bootstrapped data
  TRC_data_boot$treat_boot <- predict(treatment_model_boot, newdata = TRC_data_boot)
  TRC_data_boot$control_boot <- predict(control_model_boot, newdata = TRC_data_boot)

  # Store estimate
  iter_result2[iter] <-
    mean(TRC_data_boot$treat_boot[TRC_data_boot$TRCKNOW == 1] -
         TRC_data_boot$control_boot[TRC_data_boot$TRCKNOW == 1])
}

# Calculate standard error for ATT
```

```
ATT_se <- sd(iter_result2)


# Report and interpret ATT
#cat("ATT:", ATT, "\n")
#cat("Estimated ATT:", mean(iter_result2), "\n")
cat("ATT:", point_est_att, "\n")
```

ATT: -0.2033737

```
cat("Bootstrap SE for ATT:", ATT_se, "\n")
```

Bootstrap SE for ATT: 0.04515208

```
CI1_95 = c(point_est_att - qnorm(0.975)*sd(iter_result2),
           point_est_att + qnorm(0.975)*sd(iter_result2))
cat("CI:", CI1_95)
```

CI: -0.2918702 -0.1148773

point estimate: -0.2033737 Bootstrap SE for ATT: 0.04515208 CI: -0.2918702 -0.1148773

The bootstrap standard error, approximately 0.0452, represents the level of uncertainty in the ATT estimate, considering variations in the data due to resampling.

The point estimate of the ATT indicates that among those exposed to the TRC, there is an estimated decrease of approximately 0.2034 in respondents' agreement with the statement "I find it difficult to understand the customs and ways of the opposite racial group." This decrease is statistically significant, as evidenced by the confidence interval (-0.2918, -0.1149) not including zero. The ATT provides insight into the specific impact of TRC exposure on respondents who have been treated, showing that they are less likely to find it difficult to understand customs and ways of the opposite racial group compared to those who have not been exposed to the TRC.

part c (+4 points)

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

ATE (Average Treatment Effect) measures the average impact of exposure to the TRC on racial attitudes for all individuals in the dataset, regardless of whether they were actually exposed to the TRC. It provides an estimate of the overall impact of the TRC on the entire population. It considers the entire population, providing insights into the overall impact of the TRC on racial attitudes, including both the treated and untreated individuals. However, it may be influenced by individuals who were not exposed to the TRC but are included in the analysis, potentially diluting the treatment effect if there are differences between the treated and control groups.

ATT (Average Treatment Effect on the Treated) measures the average impact of exposure to the TRC on racial attitudes only for individuals who were actually exposed to the TRC (those with TRCKNOW == 1). It focuses on the treatment group and estimates how the TRC affected their racial attitudes. It specifically focuses on the treated group, giving us an estimate of how the TRC influenced the racial attitudes of those who were exposed to it. On the other hand, ATT is limited to the treated group and may not provide information about the broader impact of the TRC on the entire population.

For the ATE in part a, the point estimate is -0.1744, which means that, on average, the treatment decreases the outcome measure by 0.1744 units for the entire population. For the ATT in part b, the point estimate is -0.2034, suggesting that for those who actually received the treatment, the treatment decreases the outcome measure by 0.2034 units on average. The negative values for ATE and ATT indicate a decrease in the outcome variable due to the treatment. The magnitude of the effect is slightly larger for ATT (-0.2034) than for ATE (-0.1744). This could imply that, based on the CI, the treatment has a stronger effect on those who received

the treatment than on the general population, which might suggest that the treatment was appropriately targeted or that those who received the treatment were more responsive to it.

The bootstrapped standard error for ATE is 0.0436, and for ATT, it is 0.0452. These values give us an idea of the variability of the estimates across different bootstrap samples. A higher standard error suggests greater variability and less certainty in the point estimate, while these two are not significantly different.

The confidence level of both do not include 0, so both ATE and ATT estimation can reject the null hypothesis that the exposure treatment has no effect on RUSTAND (outcome), and that both are statistically significant.