Final Project Code

Jonathan Bogen, Nathan Morgan, Jared Pugh, Momina Naveed, Jeremy Rodriguez

March 13, 2023

NOTE: Used chatgpt and stackoverflow for assistance

Read the data into R

```
library(sjPlot)
library(coefplot)
```

```
## Loading required package: ggplot2
```

```
library(sjmisc)
library(sjlabelled)
```

```
##
## Attaching package: 'sjlabelled'
```

```
## The following object is masked from 'package:ggplot2':
##
##     as_label
```

```
library(modelsummary)
library(ggplot2)
library(coefplot)
library(readr)


setwd("/Users/jeremyrodriguez/Dropbox (Dartmouth College)/GOVT 10 Final Project")

main_df <- read_csv("Govt10_Project.csv")
```

```
## Rows: 2199 Columns: 94
```

```
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (94): StartDate, EndDate, Status, IPAddress, Progress, Duration (in seco...
##
```

```
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

Racializing names

```
main_df$RaceTransgressor <- NA
main_df$RaceTransgressor[grep("Darnell", main_df$name)] <- "Black"
main_df$RaceTransgressor[grep("Jermaine", main_df$name)] <- "Black"
main_df$RaceTransgressor[grep("Roosevelt", main_df$name)] <- "Black"
main_df$RaceTransgressor[grep("David", main_df$name)] <- "White"
main_df$RaceTransgressor[grep("James", main_df$name)] <- "White"
main_df$RaceTransgressor[grep("John", main_df$name)] <- "White"
main_df$RaceTransgressor[grep("Michael", main_df$name)] <- "White"
main_df$RaceTransgressor[grep("Mark", main_df$name)] <- "White"
main_df$RaceTransgressor[grep("Richard", main_df$name)] <- "White"
main_df$RaceTransgressor[grep("Robert", main_df$name)] <- "White"
main_df$RaceTransgressor[grep("Thomas", main_df$name)] <- "White"
main_df$RaceTransgressor[grep("William", main_df$name)] <- "White"
table(main_df$RaceTransgressor)
```

```
##
## Black White
##   522  1661
```

```
main_df$RaceVictim <- NA
main_df$RaceVictim[grep("Darnell", main_df$name2)] <- "Black"
main_df$RaceVictim[grep("Jermaine", main_df$name2)] <- "Black"
main_df$RaceVictim[grep("Roosevelt", main_df$name2)] <- "Black"
main_df$RaceVictim[grep("David", main_df$name2)] <- "White"
main_df$RaceVictim[grep("James", main_df$name2)] <- "White"
main_df$RaceVictim[grep("John", main_df$name2)] <- "White"
main_df$RaceVictim[grep("Michael", main_df$name2)] <- "White"
main_df$RaceVictim[grep("Mark", main_df$name2)] <- "White"
main_df$RaceVictim[grep("Richard", main_df$name2)] <- "White"
main_df$RaceVictim[grep("Robert", main_df$name2)] <- "White"
main_df$RaceVictim[grep("Thomas", main_df$name2)] <- "White"
main_df$RaceVictim[grep("William", main_df$name)] <- "White"
table(main_df$RaceVictim)
```

```
##
## Black White
##   510  1502
```

Creating casetype variable

```
main_df$case_type <- ifelse(main_df$RaceTransgressor == "Black"
                            & main_df$RaceVictim == "White", "BlackWhite",
```

```
                        ifelse(main_df$RaceTransgressor == "Black" &
                                 main_df$RaceVictim == "Black", "BlackBlack",
                               ifelse(main_df$RaceTransgressor == "White" &
                                        main_df$RaceVictim == "Black",
                                      "WhiteBlack",
                                      "WhiteWhite")))
```

Subsetting This code needs to have run off page to work with removing this text. (Gives new line error)

```
main_df<-main_df[main_df$Q130 != "{\"ImportId\":\"QID210\"}", ]

main_df<-main_df[main_df$Q130 != "Imagine the following scenario.\n\n[Field-name], a 16
```

Displaying distribution of ounishments and their proportions

```
#Distrubution of Punishments
table<-table(main_df$Q130)
table
```

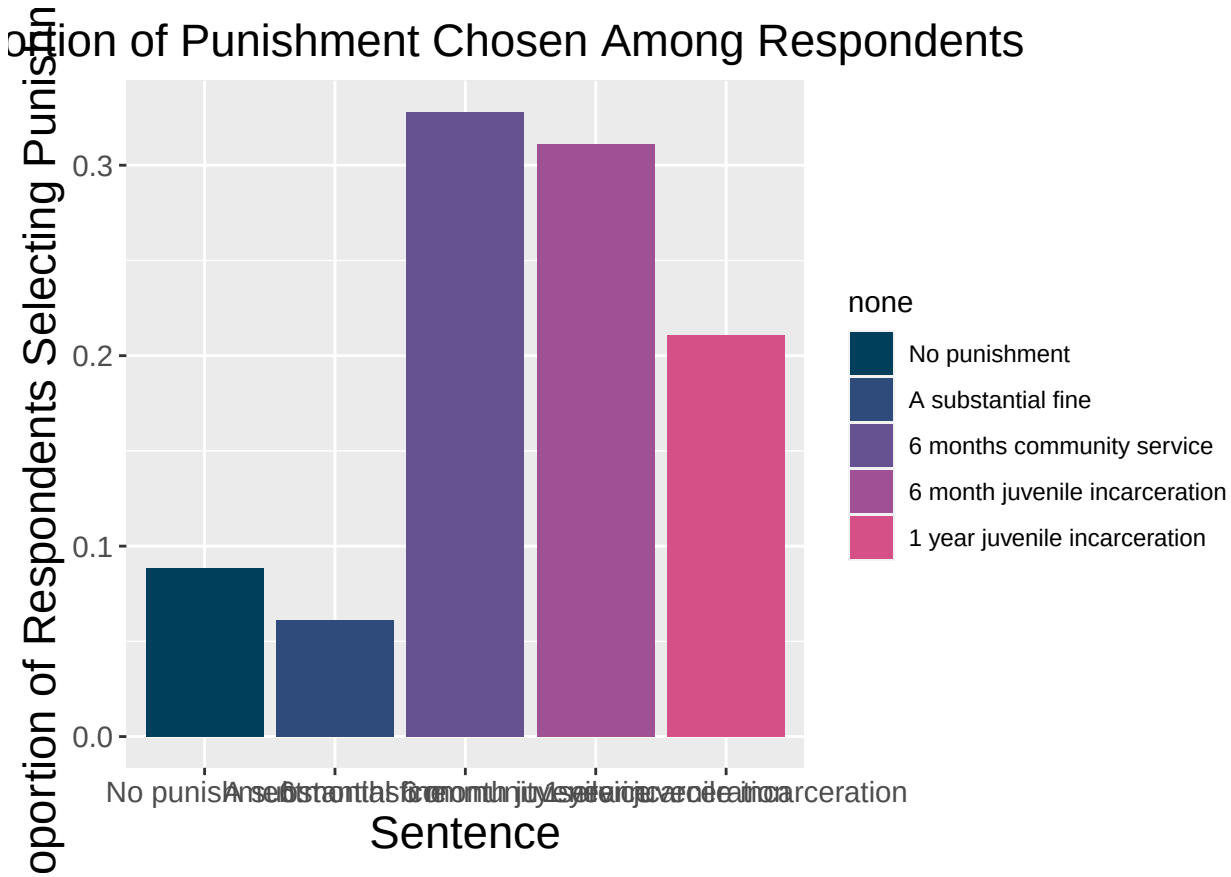
```
##
##  1 year juvenile incarceration 6 month juvenile incarceration
##                            450                             664
##     6 months community service            A substantial fine
##                            700                             131
##                  No punishment
##                            189
```

```
punishment_prop_table<-prop.table(table)

punishment_prop_df <- as.data.frame(punishment_prop_table)

names(punishment_prop_df) <- c("Sentence", "Proportion")


punishment_prop_df$Sentence<-
  factor(punishment_prop_df$Sentence,

         levels = c("No punishment",
                    "A substantial fine",
                    "6 months community service",
                    "6 month juvenile incarceration",
                    "1 year juvenile incarceration"))



g <- ggplot(punishment_prop_df,
```

```
          aes(x = Sentence, y = Proportion, fill = Sentence)) +
  geom_bar(stat = "identity") +
  xlab("Sentence") +
  ylab("Proportion of Respondents Selecting Punishment") +
  ggtitle("Proportion of Punishment Chosen Among Respondents") +
  theme(axis.text = element_text(size = 11),
        axis.title = element_text(size = 17, hjust = 0.5),
        plot.title = element_text(size=17, hjust = 0.5)) +
  scale_fill_manual(values = c("#003f5c", "#2f4b7c", "#665191",
                               "#a05195", "#d45087"),guides(fill = FALSE))
```

```
## Warning: The `<scale>` argument of `guides()` cannot be `FALSE`. Use "none" instead a
## of ggplot2 3.3.4.
```

```
g
```



Scaling Sentences

```
main_df$sentence_numeric[main_df$Q130
                         == "No punishment"] <- 1
```

```
## Warning: Unknown or uninitialised column: `sentence_numeric`.
```

```
main_df$sentence_numeric[main_df$Q130
                         == "A substantial fine"] <- 2
main_df$sentence_numeric[main_df$Q130
                         == "6 months community service"] <- 3
main_df$sentence_numeric[main_df$Q130
                         == "6 month juvenile incarceration"] <- 4
main_df$sentence_numeric[main_df$Q130
                         == "1 year juvenile incarceration"] <- 5


summary(main_df$sentence_numeric)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   1.000   3.000   4.000   3.494   4.000   5.000      63
```

Creating demographic dataframes and running bivariate regression

```
#Creating different models

white_df <- main_df[main_df$Q16 == "White", ]
black_df <- main_df[main_df$Q16 == "Black or African American", ]
male_df <- main_df[main_df$gender == 1, ]
female_df <- main_df[main_df$gender == 2, ]
dem_df <- main_df[main_df$Q6 == "Democrat", ]
rep_df <- main_df[main_df$Q6 == "Republican", ]
asian_df <- main_df[main_df$Q16 == "Asian or Pacific Islander", ]
hispanic_df <- main_df[main_df$Q16 == "Hispanic", ]

#Overall


model<-lm(sentence_numeric ~
            RaceVictim + RaceTransgressor, data = main_df)
summary(model)
```

```
##
## Call:
## lm(formula = sentence_numeric ~ RaceVictim + RaceTransgressor,
##     data = main_df)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.5234 -0.5234  0.4766  0.5564  1.6090
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept)            3.39096    0.06898  49.161   <2e-16 ***
## RaceVictimWhite        0.05268    0.05970   0.882    0.378
## RaceTransgressorWhite  0.07972    0.06158   1.295    0.196
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.144 on 1938 degrees of freedom
##   (256 observations deleted due to missingness)
## Multiple R-squared:  0.001305,   Adjusted R-squared:  0.0002743
## F-statistic: 1.266 on 2 and 1938 DF,  p-value: 0.2821
```

```
#White

model_white <- lm(sentence_numeric ~
            RaceVictim + RaceTransgressor, data = white_df)
summary(model_white)
```

```
##
## Call:
## lm(formula = sentence_numeric ~ RaceVictim + RaceTransgressor,
##     data = white_df)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.5432 -0.5432  0.4568  0.5996  1.6628
##
## Coefficients:
##                        Estimate Std. Error t value Pr(>|t|)
## (Intercept)            3.33716    0.09015  37.017   <2e-16 ***
## RaceVictimWhite        0.06327    0.07830   0.808   0.4193
## RaceTransgressorWhite  0.14280    0.08104   1.762   0.0784 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.091 on 996 degrees of freedom
##   (165 observations deleted due to missingness)
## Multiple R-squared:  0.003839,   Adjusted R-squared:  0.001838
## F-statistic: 1.919 on 2 and 996 DF,  p-value: 0.1473
```

```
#Black

model_black <- lm(sentence_numeric ~
            RaceVictim + RaceTransgressor, data = black_df)
summary(model_black)
```

```
##
```

```
## Call:
## lm(formula = sentence_numeric ~ RaceVictim + RaceTransgressor,
##     data = black_df)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.7302 -0.5610  0.2698  0.5717  1.5717
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)             3.7302     0.2094  17.812   <2e-16 ***
## RaceVictimWhite        -0.1326     0.1865  -0.711    0.478
## RaceTransgressorWhite  -0.1692     0.1914  -0.884    0.378
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.178 on 208 degrees of freedom
##   (82 observations deleted due to missingness)
## Multiple R-squared:  0.006684,   Adjusted R-squared:  -0.002867
## F-statistic: 0.6999 on 2 and 208 DF,  p-value: 0.4978
```

```
#Asian

model_asian <- lm(sentence_numeric ~
            RaceVictim + RaceTransgressor, data = asian_df)

summary(model_asian)
```

```
##
## Call:
## lm(formula = sentence_numeric ~ RaceVictim + RaceTransgressor,
##     data = asian_df)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.5464 -0.5464  0.4536  0.5678  1.6250
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)            3.37499    0.13575  24.861   <2e-16 ***
## RaceVictimWhite        0.05724    0.11615   0.493    0.622
## RaceTransgressorWhite  0.11419    0.12286   0.929    0.353
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 1.178 on 534 degrees of freedom
##   (120 observations deleted due to missingness)
## Multiple R-squared:  0.002151,   Adjusted R-squared:  -0.001586
## F-statistic: 0.5755 on 2 and 534 DF,  p-value: 0.5628
```

```
#Hispanic

model_hispanic <- lm(sentence_numeric ~
            RaceVictim + RaceTransgressor, data = hispanic_df)

summary(model_hispanic)
```

```
##
## Call:
## lm(formula = sentence_numeric ~ RaceVictim + RaceTransgressor,
##     data = hispanic_df)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.6281 -0.5161  0.4279  1.3567  1.5446
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)            3.68875    0.36336  10.152   <2e-16 ***
## RaceVictimWhite       -0.17269    0.30623  -0.564    0.574
## RaceTransgressorWhite -0.06066    0.27295  -0.222    0.824
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.297 on 123 degrees of freedom
##   (68 observations deleted due to missingness)
## Multiple R-squared:  0.002789,   Adjusted R-squared:  -0.01343
## F-statistic: 0.172 on 2 and 123 DF,  p-value: 0.8422
```

```
#Female

model_female <- lm(sentence_numeric ~
            RaceVictim + RaceTransgressor, data = female_df)
summary(model_female)
```

```
##
## Call:
## lm(formula = sentence_numeric ~ RaceVictim + RaceTransgressor,
##     data = female_df)
##
## Residuals:
```

```
##     Min      1Q  Median      3Q     Max
## -2.5082 -0.5081  0.4919  0.6257  1.6361
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)            3.36392    0.09474  35.506   <2e-16 ***
## RaceVictimWhite        0.13383    0.08198   1.633    0.103
## RaceTransgressorWhite  0.01040    0.08472   0.123    0.902
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.144 on 1027 degrees of freedom
##   (170 observations deleted due to missingness)
## Multiple R-squared:  0.002621,   Adjusted R-squared:  0.0006785
## F-statistic: 1.349 on 2 and 1027 DF,  p-value: 0.2599
```

```
#Male

model_male <- lm(sentence_numeric ~
            RaceVictim + RaceTransgressor, data = male_df)
summary(model_male)
```

```
##
## Call:
## lm(formula = sentence_numeric ~ RaceVictim + RaceTransgressor,
##     data = male_df)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.5776 -0.5411  0.4224  0.6214  1.6214
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)            3.41509    0.10112  33.774   <2e-16 ***
## RaceVictimWhite       -0.03650    0.08729  -0.418   0.6759
## RaceTransgressorWhite  0.16250    0.09002   1.805   0.0714 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.145 on 906 degrees of freedom
##   (154 observations deleted due to missingness)
## Multiple R-squared:  0.003735,   Adjusted R-squared:  0.001536
## F-statistic: 1.698 on 2 and 906 DF,  p-value: 0.1836
```

```
#Democrat
```

```r
model_dem<- lm(sentence_numeric ~
            RaceVictim + RaceTransgressor, data = dem_df)
summary(model_dem)
```

```
##
## Call:
## lm(formula = sentence_numeric ~ RaceVictim + RaceTransgressor,
##     data = dem_df)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.5324 -0.4738 -0.4273  0.5262  1.5727
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)            3.48592    0.09417  37.016   <2e-16 ***
## RaceVictimWhite       -0.05861    0.08437  -0.695    0.487
## RaceTransgressorWhite  0.04645    0.08503   0.546    0.585
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.062 on 829 degrees of freedom
##   (130 observations deleted due to missingness)
## Multiple R-squared:  0.0008907,  Adjusted R-squared:  -0.00152
## F-statistic: 0.3695 on 2 and 829 DF,  p-value: 0.6912
```

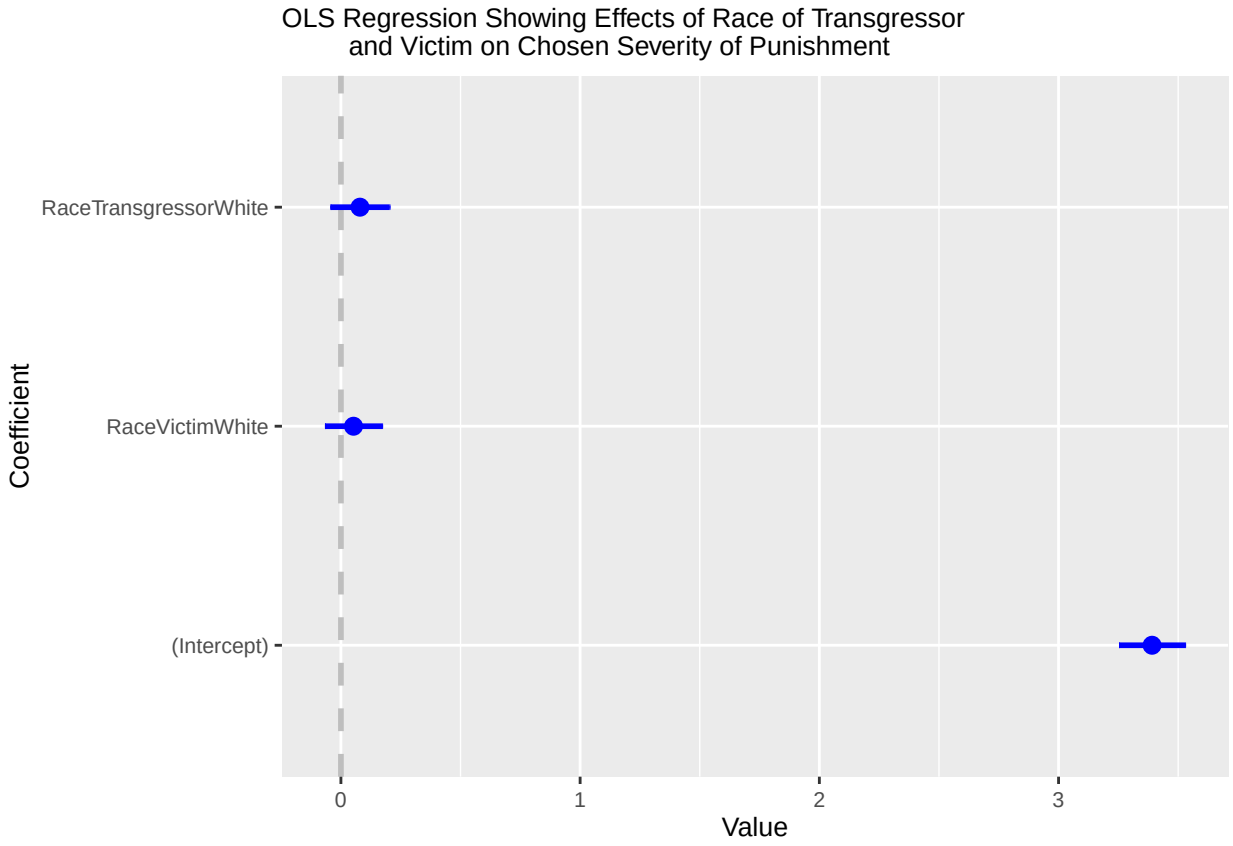
```r
#Republican

model_rep<- lm(sentence_numeric ~
            RaceVictim + RaceTransgressor, data = rep_df)
summary(model_rep)
```

```
##
## Call:
## lm(formula = sentence_numeric ~ RaceVictim + RaceTransgressor,
##     data = rep_df)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.7188 -0.6271  0.3729  0.5691  1.5691
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)            3.52268    0.13956  25.242   <2e-16 ***
## RaceVictimWhite       -0.09175    0.11717  -0.783    0.434
```

```
## RaceTransgressorWhite  0.19614    0.12650   1.550    0.122
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.144 on 473 degrees of freedom
##   (120 observations deleted due to missingness)
## Multiple R-squared:  0.006288,   Adjusted R-squared:  0.002087
## F-statistic: 1.497 on 2 and 473 DF,  p-value: 0.2249
```

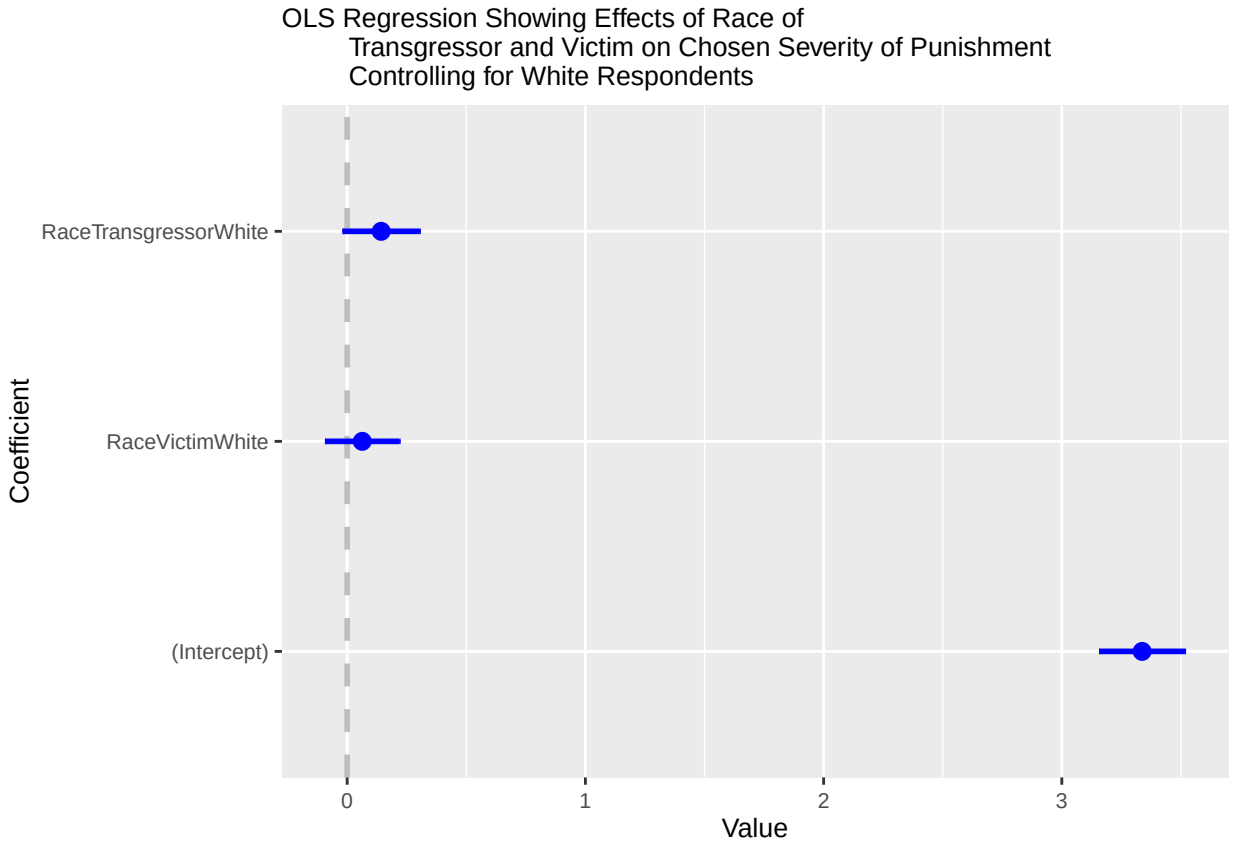
Creating demographic coefficient plots

```
#overall
new_names<-c( 'White Transgressor', 'White Victim')
coefplot(model,
         title='OLS Regression Showing Effects of Race of Transgressor
         and Victim on Chosen Severity of Punishment', innerCI=2)+
  theme(text = element_text(size = 10),
        plot.title = element_text(size = 10))
```

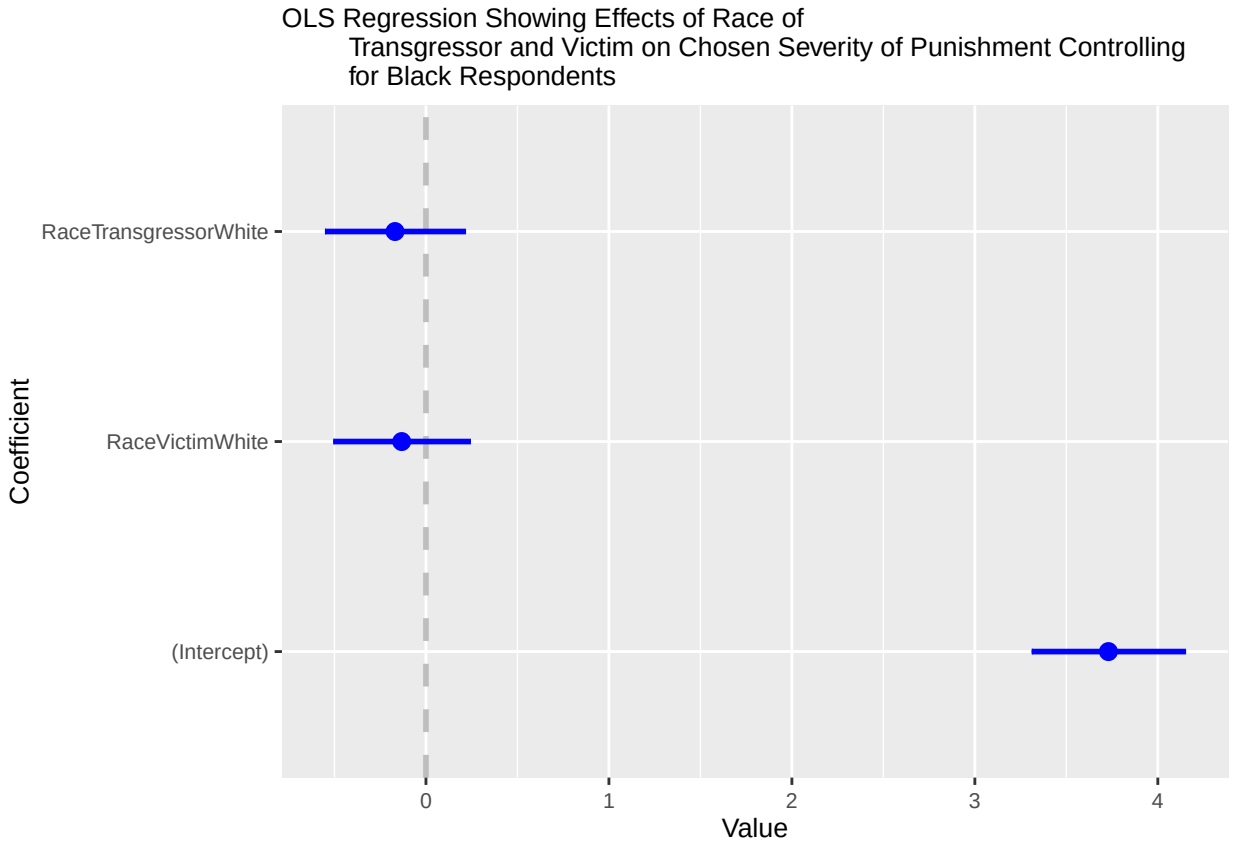



```
#White
coefplot(model_white, title='OLS Regression Showing Effects of Race of
         Transgressor and Victim on Chosen Severity of Punishment
         Controlling for White Respondents', innerCI=2)+
  theme(text = element_text(size = 10),
```

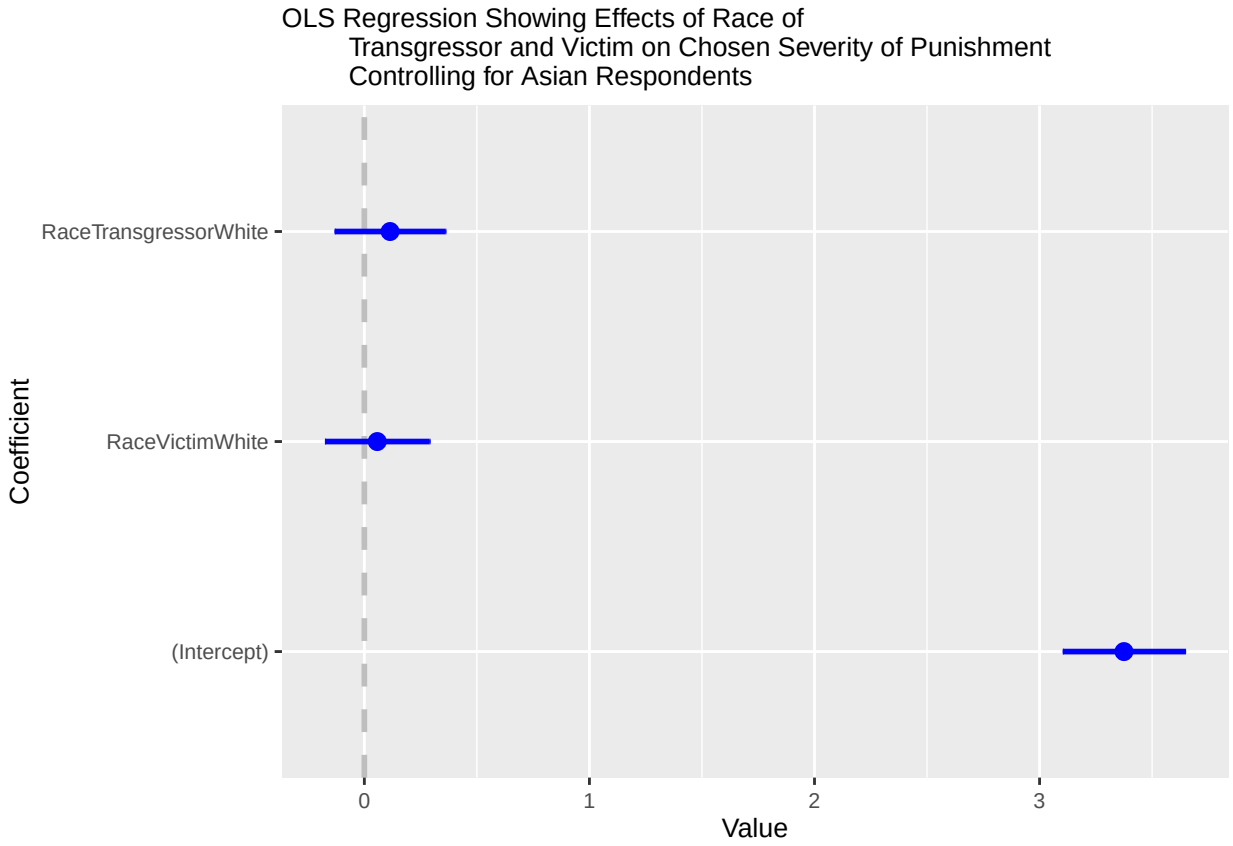
```
        plot.title = element_text(size = 10))
```



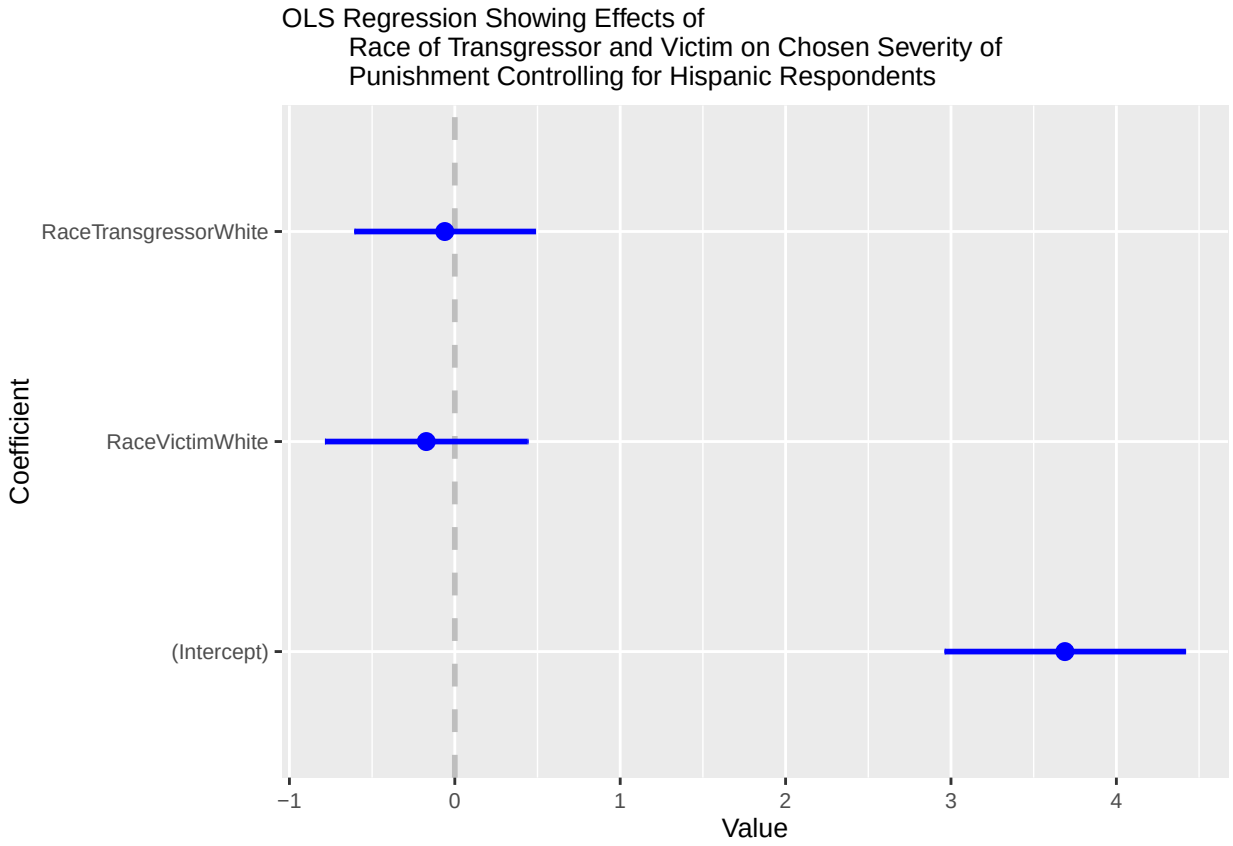

```
#Black
coefplot(model_black,title='OLS Regression Showing Effects of Race of
         Transgressor and Victim on Chosen Severity of Punishment Controlling
         for Black Respondents', innerCI=2)+
  theme(text = element_text(size = 10),
        plot.title = element_text(size = 10))
```



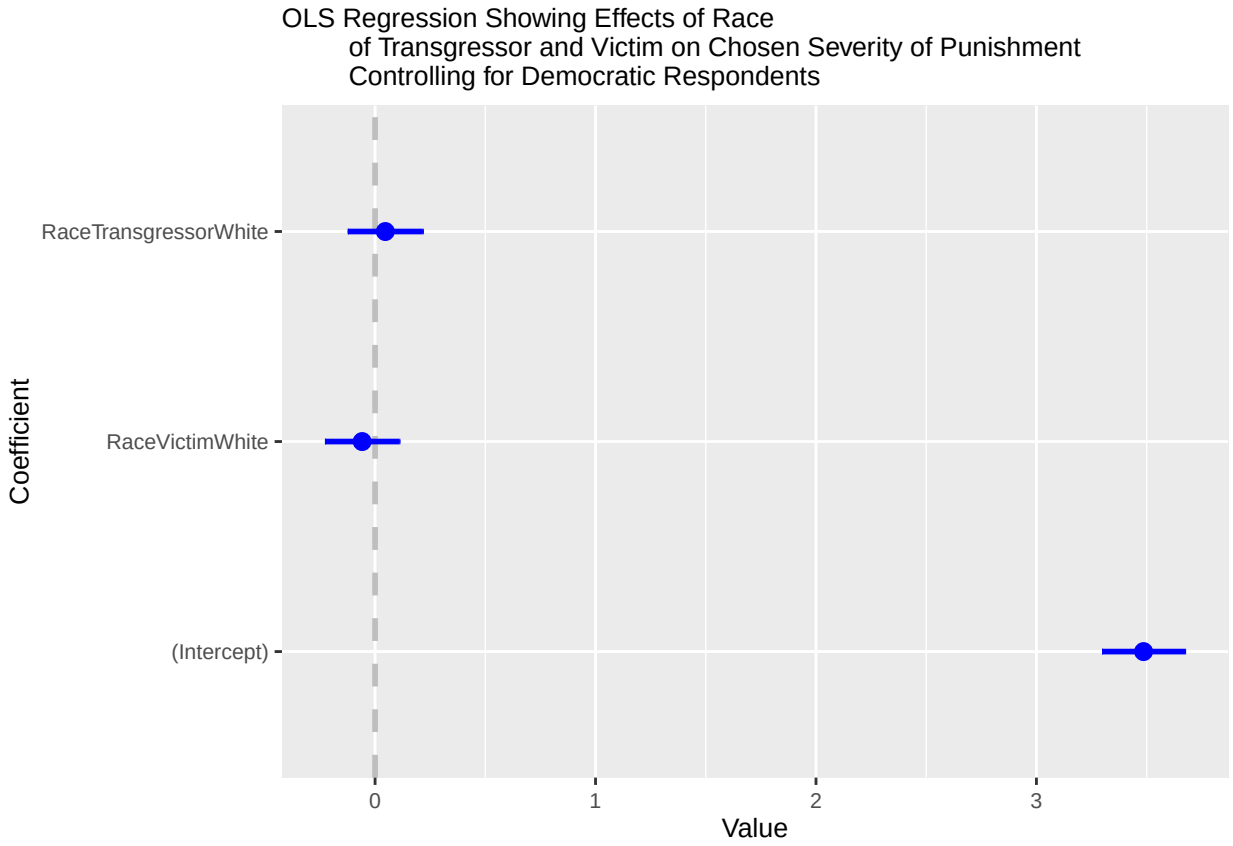

```
#Asian
coefplot(model_asian, title='OLS Regression Showing Effects of Race of
         Transgressor and Victim on Chosen Severity of Punishment
         Controlling for Asian Respondents', innerCI=2)+
  theme(text = element_text(size = 10),
        plot.title = element_text(size = 10))
```



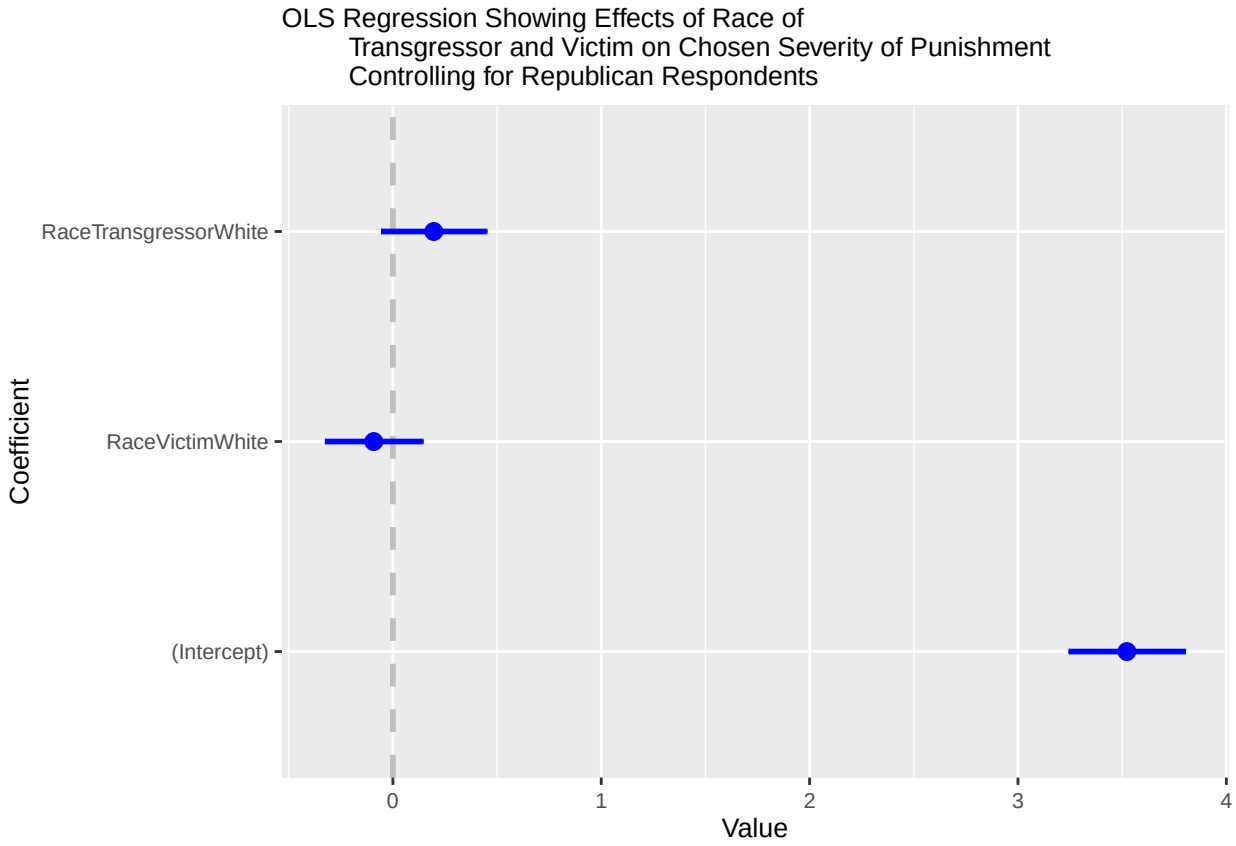

```
#Hispanic
coefplot(model_hispanic, title='OLS Regression Showing Effects of
         Race of Transgressor and Victim on Chosen Severity of
         Punishment Controlling for Hispanic Respondents', innerCI=2)+
  theme(text = element_text(size = 10),
        plot.title = element_text(size = 10))
```



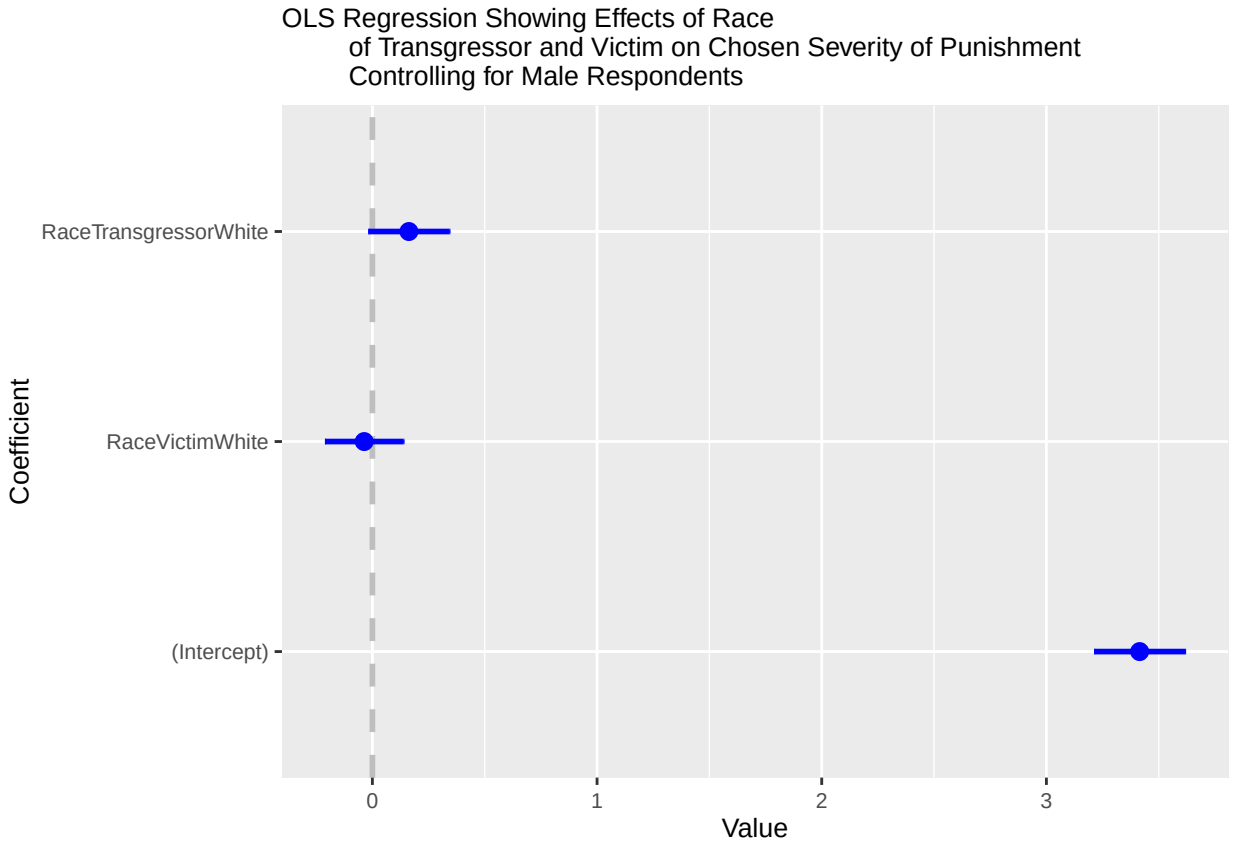

```
#Democrat
coefplot(model_dem,title='OLS Regression Showing Effects of Race
         of Transgressor and Victim on Chosen Severity of Punishment
         Controlling for Democratic Respondents', innerCI=2)+
  theme(text = element_text(size = 10),
        plot.title = element_text(size = 10))
```




```
#Republican
coefplot(model_rep, title='OLS Regression Showing Effects of Race of
         Transgressor and Victim on Chosen Severity of Punishment
         Controlling for Republican Respondents', innerCI=2)+
  theme(text = element_text(size = 10),
        plot.title = element_text(size = 10))
```




```
#Male
coefplot(model_male, title='OLS Regression Showing Effects of Race
         of Transgressor and Victim on Chosen Severity of Punishment
         Controlling for Male Respondents', innerCI=2)+
   theme(text = element_text(size = 10),
        plot.title = element_text(size = 10))
```




```
#Female
coefplot(model_female,
         title = 'OLS Regression Showing Effects of Race of
         Transgressor and Victim on Chosen Severity of Punishment
         Controlling for Female Respondents',
         innerCI = 2) +
  theme(text = element_text(size = 10),
        plot.title = element_text(size = 10))
```


OLS Regression Showing Effects of Race of
Transgressor and Victim on Chosen Severity of Punishment
Controlling for Female Respondents
RaceTransgressorWhite
RaceVictimWhite
(Intercept)
Coefficient
0
1
2
3
Value