Problem Set 3

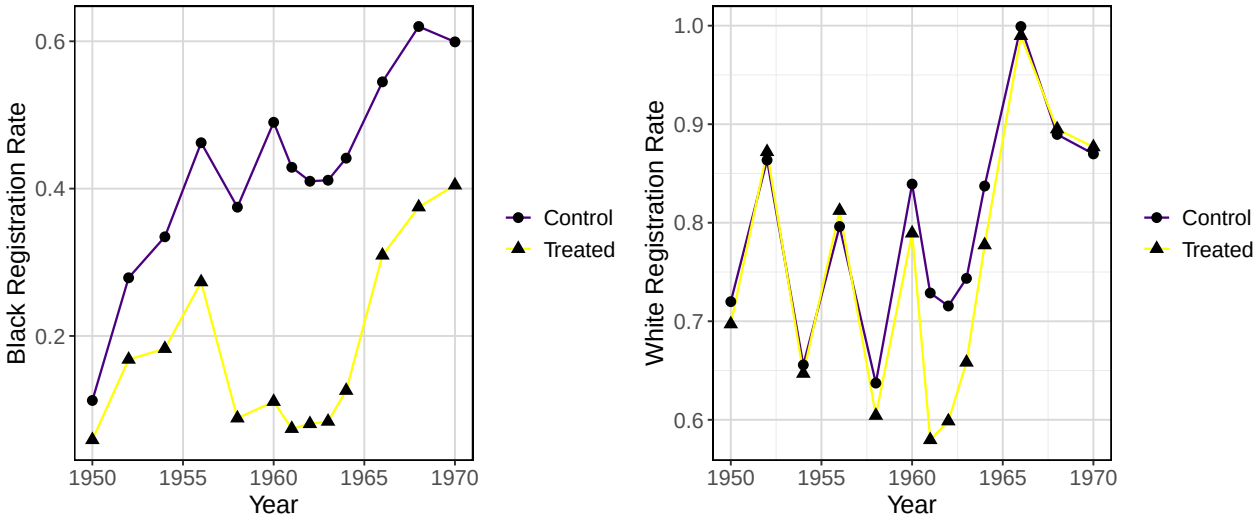
Edgar Aguilar

March 1

Replicate Figure 2

Below is a replication of Figure 2 from Keele, Cubbison, and White (2021), Suppressing Black Votes: A Historical Case Study of Voting Restrictions in Louisiana, published in the American Political Science Review.

FIGURE 2. Proportion of Registered Voters by Race and by Understanding Clause Status. Treated Parishes Enforced the Understanding Clause and Control Parishes Did Not



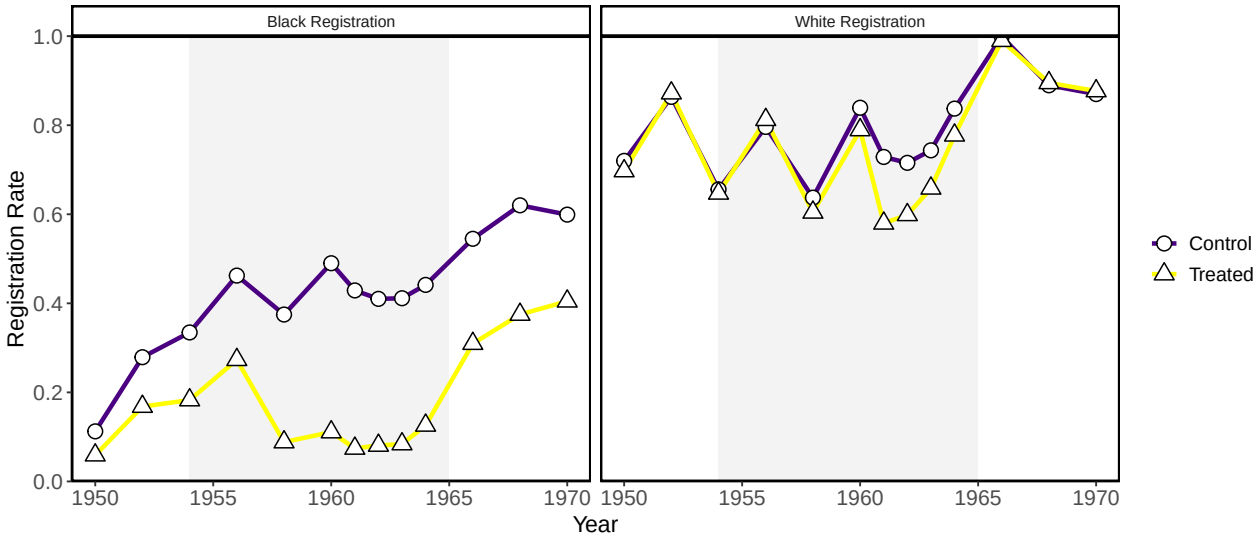
(a) Black Registration

(b) White Registration

Improve Figure 2

Below is an improved version of Figure 2 that sets the vertical axes from 0 to 1, removes gridlines, and adds a shaded area marking the Understanding Clause period (1954–1965). The plot now uses faceting to create two subplots instead of separate graphs, with enhanced line and shape visibility for clarity.

FIGURE 2. Proportion of Registered Voters by Race and by Understanding Clause Status. Treated Parishes Enforced the Understanding Clause and Control Parishes Did Not



Shaded area represents the period (1954−1965) where the Understanding Clause was in effect.

Appendix: Replication code

```
# This "setup" chunk specifies global options
# for handling code, plots, etc in your doc.
knitr::opts_chunk$set(
  eval = TRUE,
  echo = FALSE,
  warning = FALSE,
  message = FALSE,
  fig.align = 'center'
)

# Load necessary libraries
library(haven)
library(tidyverse)
library(patchwork)
# PART CERO OF ASSIGNMENT: UNDERSTANDING,CLEANING, AND TRASNFORMING THE DATA ----
# Understanding and Preparing the data
# Load dataset
data <- read_dta("la_turnout_basic.dta")

# Understanding the data ----
# Currently data is listed by parish number by year.
# Need to transform it to treatment group by year.

# Columms of interest: year, understandingclause2, whiteregrate, blackregrate

# Checking for missing values
cols_of_interest <- data %>% select(year, understandingclause2, whiteregrate, blackregrate)

# Summary of missing values
missing_summary <- cols_of_interest %>%
```

```r
  summarise(across(everything(), ~ sum(is.na(.)), .names = "Missing_{.col}"))

# Preparing the data ----

# Filter Data for Years 1950-1970
data_und <- data %>%
  filter(year >= 1950 & year <= 1970)

# Handling missing values using linear interpolation
data_und <- data_und %>%
  group_by(parishnumber) %>%
  mutate(
    brrate = ifelse(is.na(blackregrate), approx(year, blackregrate, year, rule = 2)$y, blackregrate),
    wrrate = ifelse(is.na(whiteregrate), approx(year, whiteregrate, year, rule = 2)$y, whiteregrate)
  ) %>%
  ungroup()

# Check for missing values again
missing_summary <- data_und %>%
  summarise(across(everything(), ~ sum(is.na(.)), .names = "Missing_{.col}"))

# Transform the Data ----
# Creating new columms: Year, Treatment Group, Black Registration Rate, White Registration Rate
# Treatment Group = "Treated" if understandingclause2 column = 1, "Control" if understandingclause2 col

transformed_data <- data_und %>%
  mutate(
    Treatment_Group = case_when(
      understandingclause2 == 1 ~ "Treated",
      understandingclause2 == 0 ~ "Control",
      TRUE ~ NA_character_
    )
  ) %>%
  group_by(year, Treatment_Group) %>%
  summarise(
    Black_Registration_Rate = mean(brrate, na.rm = TRUE),
    White_Registration_Rate = mean(wrrate, na.rm = TRUE),
    .groups = "drop"
  )
# PART ONE OF ASSIGNMENT: REPLICATING THE PLOT ----
# Black Registration Rate Plot
plot_black <- ggplot(transformed_data, aes(x = year, y = Black_Registration_Rate, group = Treatment_Grou
  geom_line(aes(colour = Treatment_Group), linewidth = 0.6) +
  geom_point(aes(shape = Treatment_Group), size = 2.5, fill = "black") +
  scale_color_manual(values = c("Control" = "#4B0082", "Treated" = "yellow")) +
  scale_shape_manual(values = c("Control" = 16, "Treated" = 17)) +
  labs(
    x = "Year",
    y = "Black Registration Rate",
    caption = "(a) Black Registration"
  ) +
  theme_bw() +
  theme(
```

```
    legend.position = "right",
    legend.title = element_blank(),
    legend.text = element_text(size = 12),
    panel.grid.major = element_line(color = "gray85"),
    panel.grid.minor = element_blank(),
    panel.border = element_rect(color = "black", fill = NA, linewidth = 0.7),
    axis.text = element_text(size = 12),
    axis.title = element_text(size = 14),
    plot.caption = element_text(size = 14, family = "serif", hjust = 0.5, margin = margin(t = 5))
  )

# White Registration Rate Plot
plot_white <- ggplot(transformed_data, aes(x = year, y = White_Registration_Rate, group = Treatment_Grou
  geom_line(aes(colour = Treatment_Group), linewidth = 0.6) +
  geom_point(aes(shape = Treatment_Group), size = 2.5, fill = "black") +
  scale_color_manual(values = c("Control" = "#4B0082", "Treated" = "yellow")) +
  scale_shape_manual(values = c("Control" = 16, "Treated" = 17)) +
  labs(
    x = "Year",
    y = "White Registration Rate",
    caption = "(b) White Registration"
  ) +
  theme_bw() +
  theme(
    legend.position = "right",
    legend.title = element_blank(),
    legend.text = element_text(size = 12),
    panel.grid.major = element_line(color = "gray85"),
    panel.minor = element_blank(),
    panel.border = element_rect(color = "black", fill = NA, linewidth = 0.7),
    axis.text = element_text(size = 12),
    axis.title = element_text(size = 14),
    plot.caption = element_text(size = 14, family = "serif", hjust = 0.5, margin = margin(t = 5))
  )

# Combine the two plots into a single figure
combined_plot <- (plot_black | plot_white) +
  plot_annotation(
    title = "FIGURE 2. Proportion of Registered Voters by Race and by Understanding Clause Status.\nTrea
    theme = theme(
      plot.title = element_text(
        size = 14,
        face = "bold",
        hjust = 0,
        margin = margin(b = 8)
      )
    )
  )

# Print the final combined plot
print(combined_plot)
# PART TWO OF ASSIGNMENT: IMPROVING THE PLOT ----
# Pivot data for faceting (Reshape from wide to long format)
```

```r
transformed_data_long <- transformed_data %>%
  pivot_longer(cols = c(Black_Registration_Rate, White_Registration_Rate),
               names_to = "Race",
               values_to = "Registration_Rate") %>%
  mutate(Race = recode(Race,
                       "Black_Registration_Rate" = "Black Registration",
                       "White_Registration_Rate" = "White Registration"))

# Create the plot with faceting
figure_2 <- ggplot(transformed_data_long, aes(x = year, y = Registration_Rate,
                                              group = Treatment_Group, color = Treatment_Group, shape =

  # Add a shaded area for Understanding Clause enforcement (1954-1965)
  annotate("rect", xmin = 1954, xmax = 1965, ymin = 0, ymax = 1, fill = "gray85", alpha = 0.3) +

  # Line plot (kept the same)
  geom_line(linewidth = 1.2) +

  # Points: TRUE hollow shapes with thin black border
  geom_point(size = 3.5, stroke = 0.6, fill = "white", color = "black") +

  # Custom colors and shapes (Now using 21 & 24 for hollow effect)
  scale_color_manual(values = c("Control" = "#4B0082", "Treated" = "yellow")) +
  scale_shape_manual(values = c("Control" = 21, "Treated" = 24)) +

  # Faceting to create two subplots
  facet_wrap(~Race) +

  # Adjust axes
  scale_x_continuous(breaks = seq(1950, 1970, by = 5), limits = c(1950, 1970)) +
  scale_y_continuous(breaks = seq(0, 1, by = 0.2), limits = c(0, 1), expand = expansion(mult = c(0, 0))

  # Labels and Title
  labs(
    x = "Year",
    y = "Registration Rate",
    title = "FIGURE 2. Proportion of Registered Voters by Race and by Understanding Clause Status. \nTr
    caption = "Shaded area represents the period (1954-1965) where the Understanding Clause was in effe
  ) +

  # Theme adjustments
  theme_classic() +  # Removes gridlines
  theme(
    legend.position = "right",
    legend.title = element_blank(),
    legend.text = element_text(size = 11),
    panel.border = element_rect(color = "black", fill = NA, linewidth = 1),
    axis.text = element_text(size = 12),
    axis.title = element_text(size = 13),
    plot.title = element_text(size = 14, face = "bold", hjust = 0),
    plot.caption = element_text(size = 11, hjust = 0.5, face = "italic")
  )
```

```r
# Print the improved figure
print(figure_2)
```