Problem Set 5

Leah Shiferaw

2024-03-18

Instructions Generate the requested plots to the best of your ability. Be sure to knit the document to .pdf format. echo all code chunks. The chunks must show every step you took and option you specified to generate the plots. Submit your work in hard copy by March 20.

Suppressing Black Votes Keele et al (2021) evaluate the impact of efforts to disenfranchise African American voters in Louisiana in the 1950s and 1960s. They focus specifically on the Understanding Clause, which some parishes administered and others did not.

You need the following materials:

KEELE L, CUBBISON W, WHITE I. (2021). Suppressing Black Votes: A Historical Case Study of Voting Restrictions in Louisiana. American Political Science Review. https://doi.org/10.1017/S0003055421000034.

la_turnout_basic.dta replication data, available in this PS5 repository. Hint: this is a parish-by-year dataset while the plot you need to create is a treatment group-by-year dataset. You'll need to transform the data to get there. You may even need to clean the understandingclause2 variable which equals 1 for a parish that used the Understanding Clause (Treated) and 0 for a parish that did not (Control).

Your task is to replicate, and then improve on, Figure 2. Be sure to choose an appropriate size for the visuals in your final output.

```
df = read_dta('la_turnout_basic.dta')

df2 = df |>
  select(year, understandingclause2, blackregrate, whiteregrate)

df3 = df2 |>
  mutate(
    understandingclause2 = case_when(
      understandingclause2 == 0 ~ "Treated",
      understandingclause2 == 1 ~ "Control",
    )
  )

df4 = df3 |>
  filter(year >= 1950 & year <= 1970 ) |>
  group_by(year) |>
  summarize(avg_blackregrate = mean(blackregrate, na.rm = TRUE),
            avg_whiteregrate = mean(whiteregrate, na.rm = TRUE))

black_reg_rate = df3 |>
  filter(year >= 1950 & year <= 1970 ) |>
  group_by(year, understandingclause2) |>
  summarize(avg_blackregrate = mean(blackregrate, na.rm = TRUE))
```

```
## `summarise()` has grouped output by 'year'. You can override using the
```

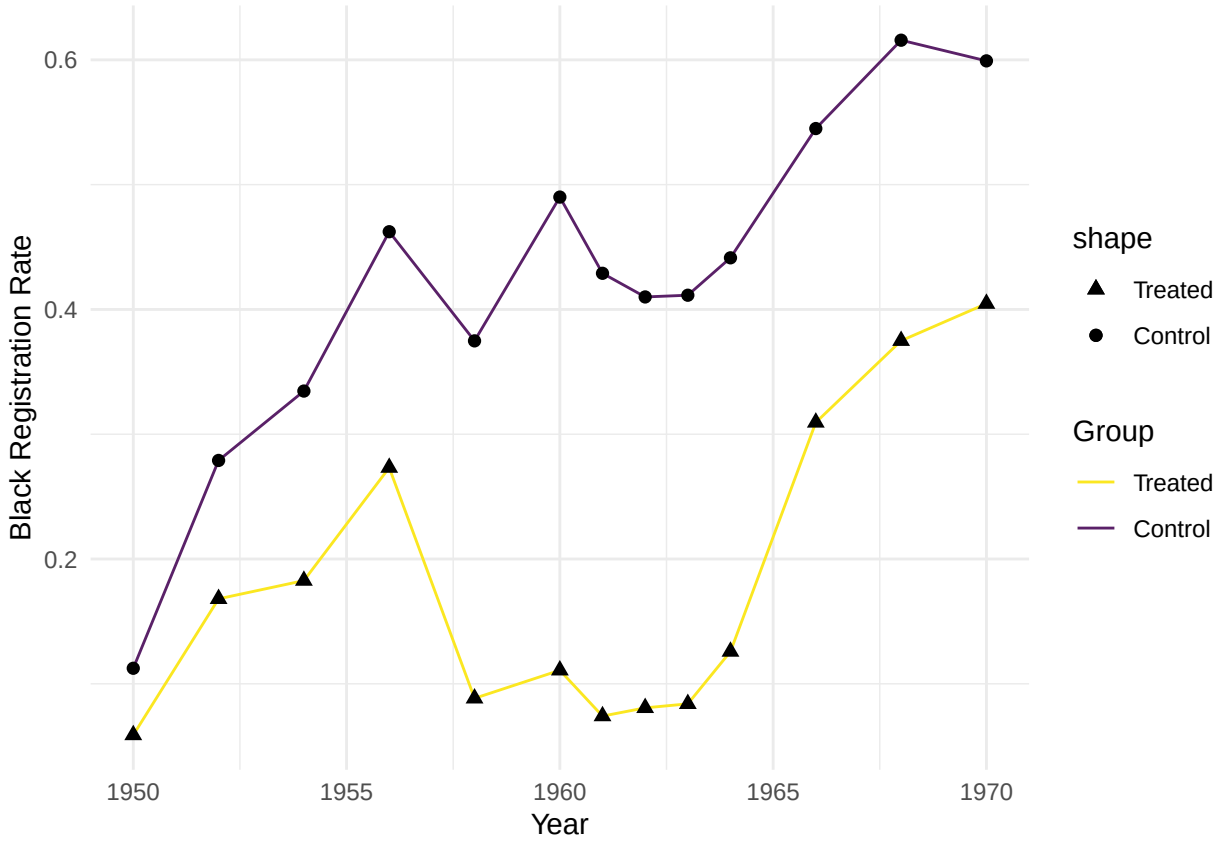
```
## `.groups` argument.
```

```
white_reg_rate = df3 |>
  filter(year >= 1950 & year <= 1970 ) |>
  group_by(year, understandingclause2) |>
  summarize(avg_whiteregrate = mean(whiteregrate, na.rm = TRUE))
```

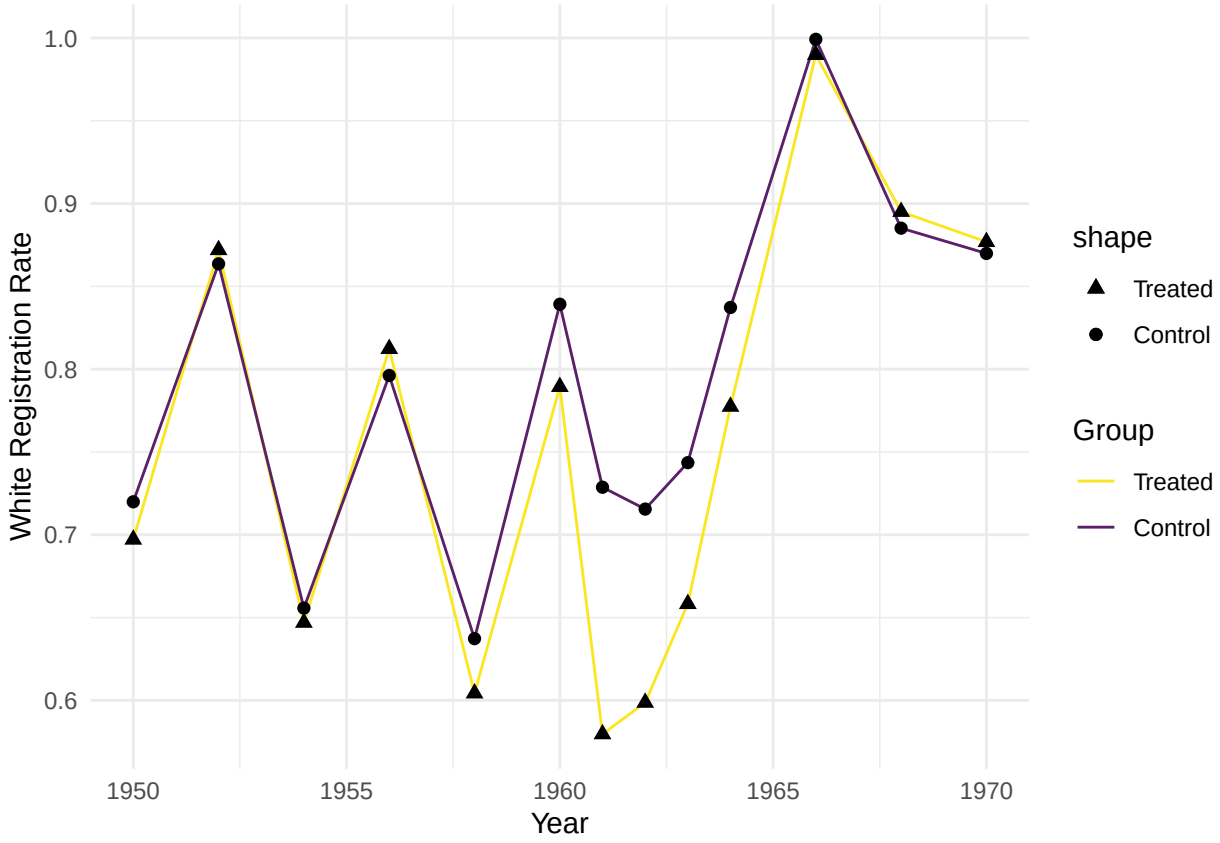
```
## `summarise()` has grouped output by 'year'. You can override using the
## `.groups` argument.
```

Replicate Figure 2 from the article. Match every element of the plot. Note that for this figure, they create two separate graphs and paste them together. You will need to do the same with patchwork.

```
ggplot(black_reg_rate, aes(x = year, y = avg_blackregrate,
                           color = factor(understandingclause2))) +
  geom_line() +
  geom_point(data = subset(black_reg_rate, understandingclause2 == "Treated"),
             aes(shape = "Treated"), color = "black", fill = "black", size = 2) +
  geom_point(data = subset(black_reg_rate, understandingclause2 == "Control"),
             aes(shape = "Control"), color = "black", fill = "black", size = 2) +
  labs(x = "Year", y = "Black Registration Rate") +
  scale_color_manual(values = c("#FBE721", "#5a2168"),
                     labels = c("Treated", "Control")) +
  scale_shape_manual(values = c("Treated" = 16, "Control" = 17),
                     labels = c("Treated", "Control")) +
  guides(color = guide_legend(title = "Group")) +
  theme_minimal()
```

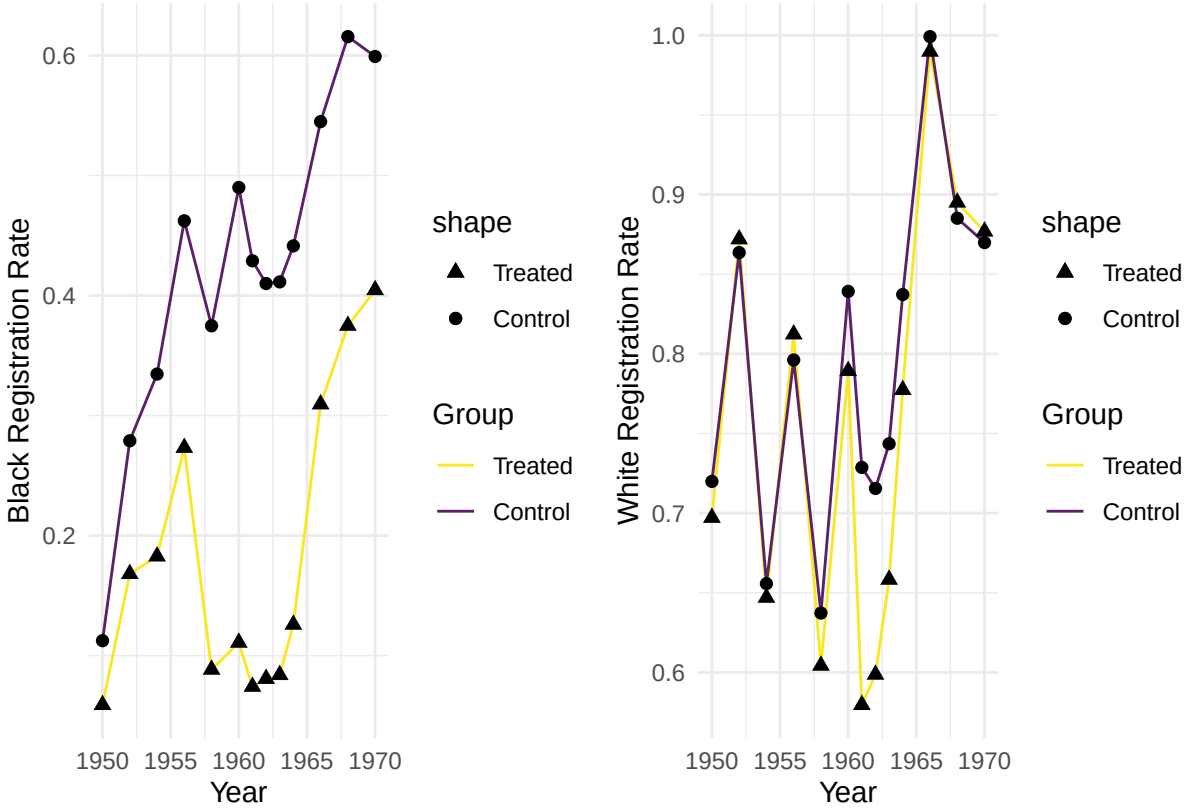


```
  ggplot(white_reg_rate, aes(x = year, y = avg_whiteregrate,
                             color = factor(understandingclause2))) +
  geom_line() +
  geom_point(data = subset(white_reg_rate, understandingclause2 == "Treated"),
             aes(shape = "Treated"), color = "black", fill = "black", size = 2) +
  geom_point(data = subset(white_reg_rate, understandingclause2 == "Control"),
             aes(shape = "Control"), color = "black", fill = "black", size = 2) +
  labs(x = "Year", y = "White Registration Rate") +
  scale_color_manual(values = c("#FBE721", "#5a2168"),
                     labels = c("Treated", "Control")) +
  scale_shape_manual(values = c("Treated" = 16, "Control" = 17),
                     labels = c("Treated", "Control")) +
  guides(color = guide_legend(title = "Group")) +
  theme_minimal()
```




```
p1 <- black_reg_rate |>
  ggplot(aes(x = year, y = avg_blackregrate, color = factor(understandingclause2))) +
  geom_line() +
  geom_point(data = subset(black_reg_rate, understandingclause2 == "Treated"),
             aes(shape = "Treated"), color = "black", fill = "black", size = 2) +
  geom_point(data = subset(black_reg_rate, understandingclause2 == "Control"),
             aes(shape = "Control"), color = "black", fill = "black", size = 2) +
  labs(x = "Year", y = "Black Registration Rate") +
  scale_color_manual(values = c("#FBE721", "#5a2168"),
                     labels = c("Treated", "Control")) +
  scale_shape_manual(values = c("Treated" = 16, "Control" = 17),
                     labels = c("Treated", "Control")) +
```

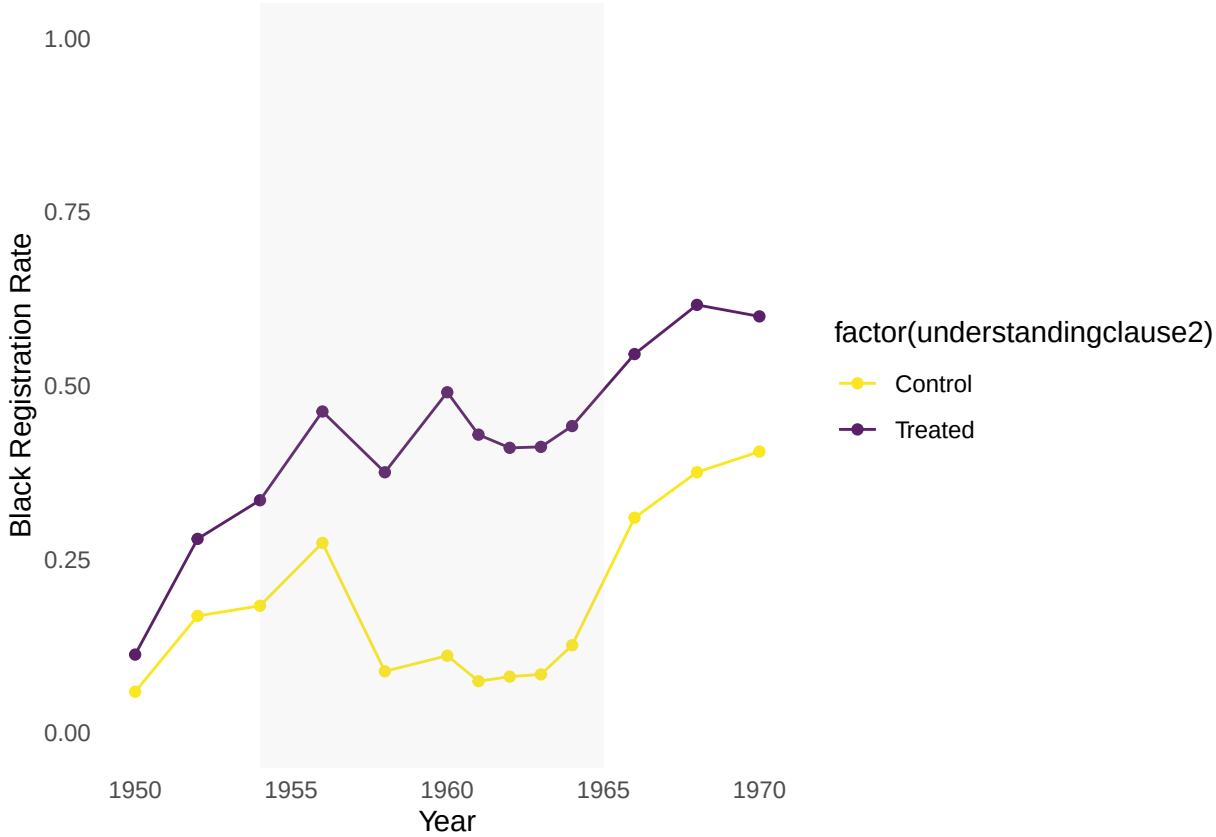
```
  guides(color = guide_legend(title = "Group")) +
  theme_minimal()


p2 <- white_reg_rate |>
  ggplot(aes(x = year, y = avg_whiteregrate,
             color = factor(understandingclause2))) +
  geom_line() +
  geom_point(data = subset(white_reg_rate,
                           understandingclause2 == "Treated"),
             aes(shape = "Treated"), color = "black", fill = "black",
             size = 2) +
  geom_point(data = subset(white_reg_rate,
                           understandingclause2 == "Control"),
             aes(shape = "Control"), color = "black", fill = "black",
             size = 2) +
  labs(x = "Year", y = "White Registration Rate") +
  scale_color_manual(values = c("#FBE721", "#5a2168"),
                     labels = c("Treated", "Control")) +
  scale_shape_manual(values = c("Treated" = 16, "Control" = 17),
                     labels = c("Treated", "Control")) +
  guides(color = guide_legend(title = "Group")) +
  theme_minimal()

p1 + p2 + plot_layout(widths = c(8, 8), heights = c(2))
```



```
ggsave("combined_plot.png", width = 8, height = 2)
```

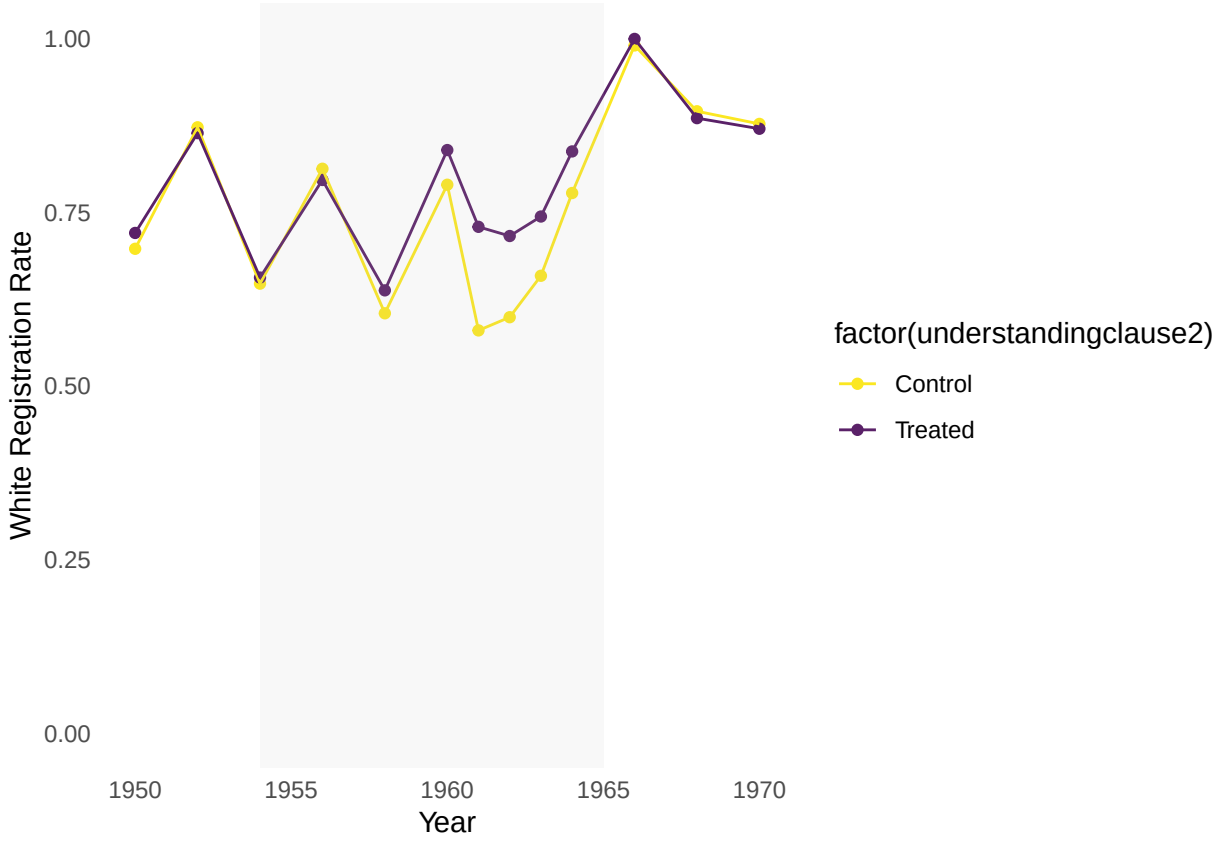
Improve Figure 2. I'm asking for several changes: - Set the vertical axes to range from 0-1. - Remove the gridlines from the plot. - Add a lightly shaded area denoting the use of the Understanding Clause (1954 to 1965 according to Justice Department investigations). - Use faceting to create two subplots instead of two separate plots. Hint: you will need to pivot to make this work. - Improve the visibility of the lines/shapes as you see fit.

```
black_reg_rate <- black_reg_rate
ggplot(black_reg_rate, aes(x = year, y = avg_blackregrate,
                           color = factor(understandingclause2))) +
  geom_point() +
  geom_line() +
  labs(x = "Year", y = "Black Registration Rate") +
  scale_color_manual(values = c("#FBE721", "#5a2168"),
                     labels = c("Control", "Treated")) +
  scale_shape_manual(values = c("Treated" = 16, "Control" = 17),
                     labels = c("Treated", "Control")) +
  ylim(0, 1) +
  theme_minimal() +
  theme(panel.grid = element_blank()) +
  annotate("rect", xmin = 1954, xmax = 1965, ymin = -Inf, ymax = Inf,
           alpha = 0.1, fill = "gray")
```



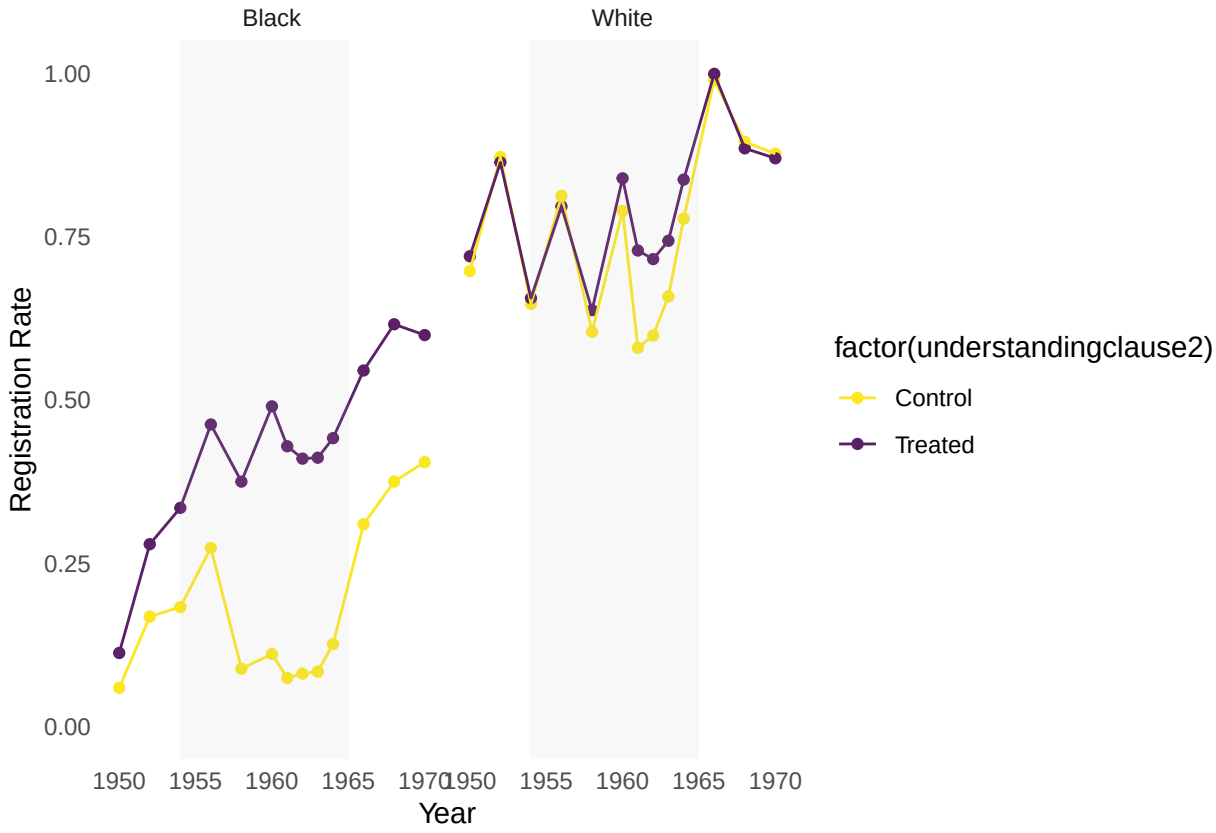

```
white_reg_rate <- white_reg_rate
ggplot(white_reg_rate, aes(x = year, y = avg_whiteregrate,
                           color = factor(understandingclause2))) +
  geom_point() +
```

```
  geom_line() +
  labs(x = "Year", y = "White Registration Rate") +
  scale_color_manual(values = c("#FBE721", "#5a2168"),
                     labels = c("Control", "Treated")) +
  scale_shape_manual(values = c(16, 17),
                     labels = c("Control", "Treated")) +
  ylim(0, 1) +
  theme_minimal() +
  theme(panel.grid = element_blank()) +
  annotate("rect", xmin = 1954, xmax = 1965, ymin = -Inf, ymax = Inf,
           alpha = 0.1, fill = "gray")
```




```
black_long <- pivot_longer(black_reg_rate,
                           cols = avg_blackregrate,
                           names_to = "race",
                           values_to = "avg_regrate")

white_long <- pivot_longer(white_reg_rate,
                           cols = avg_whiteregrate,
                           names_to = "race",
                           values_to = "avg_regrate")


combined_data <- rbind(mutate(black_long, race = "Black"),
                       mutate(white_long, race = "White"))

ggplot(combined_data, aes(x = year, y = avg_regrate,
```

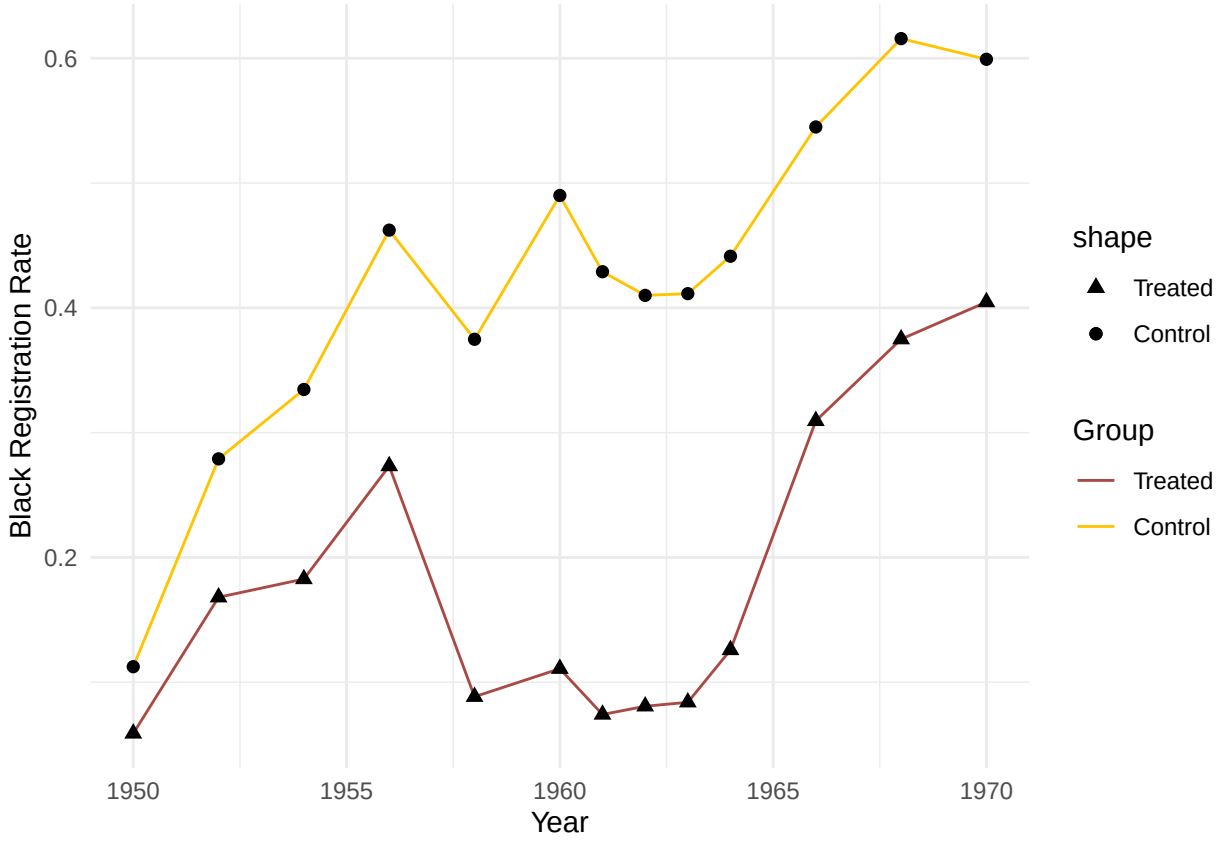
```
                          color = factor(understandingclause2))) +
  geom_point() +
  geom_line() +
  labs(x = "Year", y = "Registration Rate") +
  scale_color_manual(values = c("#FBE721", "#5a2168"),
                     labels = c("Control", "Treated")) +
  scale_shape_manual(values = c("Treated" = 16, "Control" = 17),
                     labels = c("Treated", "Control")) +
  ylim(0, 1) +
  theme_minimal() +
  theme(panel.grid = element_blank()) +
  annotate("rect", xmin = 1954, xmax = 1965, ymin = -Inf, ymax = Inf,
           alpha = 0.1, fill = "gray") +
  facet_wrap(~ race)
```



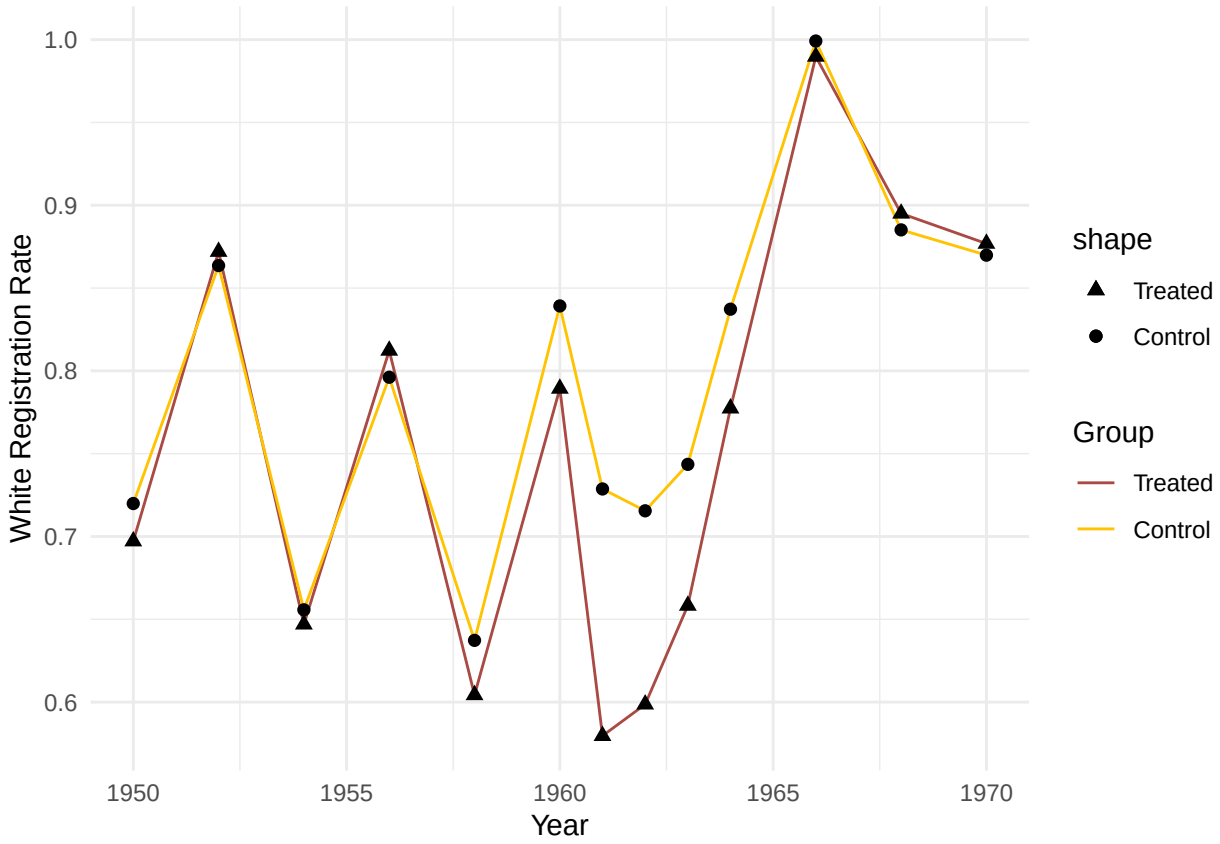
Improved graph

```
ggplot(black_reg_rate, aes(x = year, y = avg_blackregrate,
                           color = factor(understandingclause2))) +
  geom_line() +
  geom_point(data = subset(black_reg_rate, understandingclause2 == "Treated"),
             aes(shape = "Treated"), color = "black", fill = "black", size = 2) +
  geom_point(data = subset(black_reg_rate, understandingclause2 == "Control"),
             aes(shape = "Control"), color = "black", fill = "black", size = 2) +
  labs(x = "Year", y = "Black Registration Rate") +
  scale_color_manual(values = c("#AA4A44", "#FFC300"),
                     labels = c("Treated", "Control")) +
  scale_shape_manual(values = c("Treated" = 16, "Control" = 17),
```

```
                    labels = c("Treated", "Control")) +
guides(color = guide_legend(title = "Group")) +
theme_minimal()
```

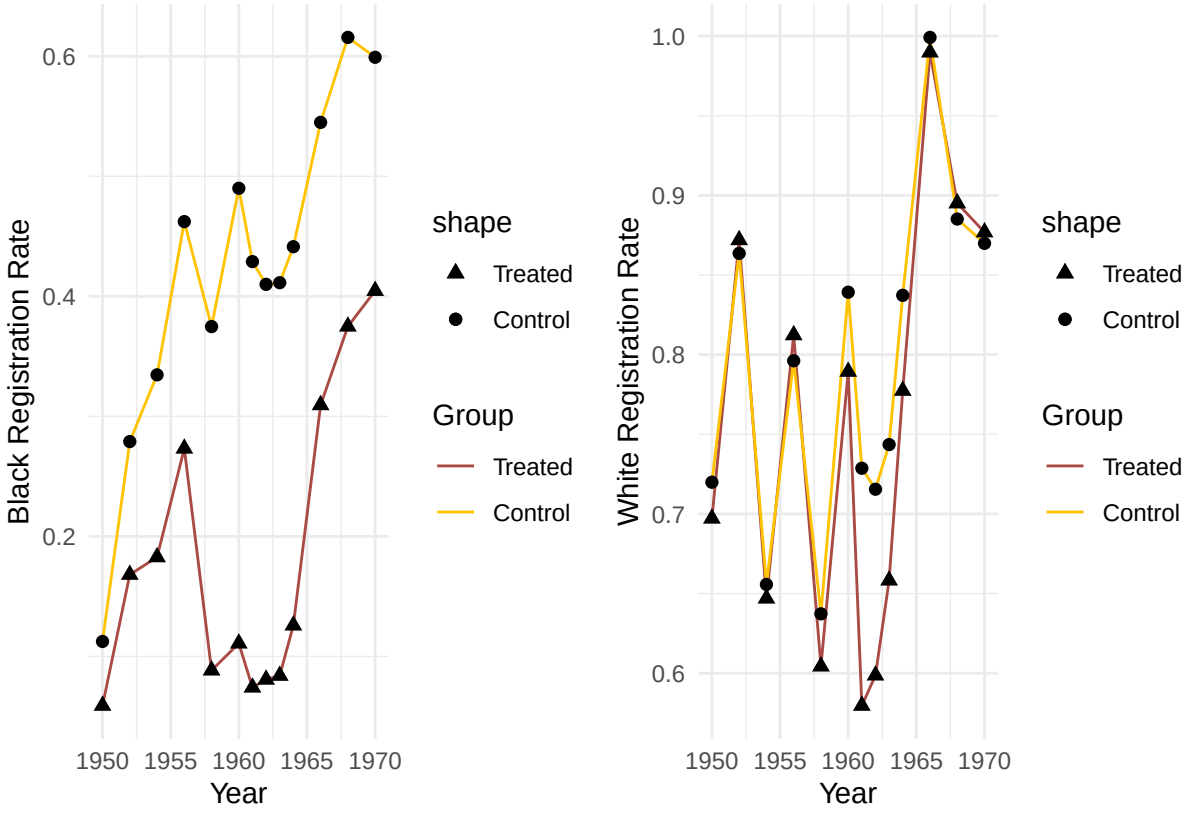



```
ggplot(white_reg_rate, aes(x = year, y = avg_whiteregrate,
                           color = factor(understandingclause2))) +
geom_line() +
geom_point(data = subset(white_reg_rate, understandingclause2 == "Treated"),
           aes(shape = "Treated"), color = "black", fill = "black", size = 2) +
geom_point(data = subset(white_reg_rate, understandingclause2 == "Control"),
           aes(shape = "Control"), color = "black", fill = "black", size = 2) +
labs(x = "Year", y = "White Registration Rate") +
scale_color_manual(values = c("#AA4A44", "#FFC300"),
                   labels = c("Treated", "Control")) +
scale_shape_manual(values = c("Treated" = 16, "Control" = 17),
                   labels = c("Treated", "Control")) +
guides(color = guide_legend(title = "Group")) +
theme_minimal()
```




```
p1 <- black_reg_rate |>
  ggplot(aes(x = year, y = avg_blackregrate, color = factor(understandingclause2))) +
  geom_line() +
  geom_point(data = subset(black_reg_rate, understandingclause2 == "Treated"),
             aes(shape = "Treated"), color = "black", fill = "black", size = 2) +
  geom_point(data = subset(black_reg_rate, understandingclause2 == "Control"),
             aes(shape = "Control"), color = "black", fill = "black", size = 2) +
  labs(x = "Year", y = "Black Registration Rate") +
  scale_color_manual(values = c("#AA4A44", "#FFC300"),
                     labels = c("Treated", "Control")) +
  scale_shape_manual(values = c("Treated" = 16, "Control" = 17),
                     labels = c("Treated", "Control")) +
  guides(color = guide_legend(title = "Group")) +
  theme_minimal()


p2 <- white_reg_rate |>
  ggplot(aes(x = year, y = avg_whiteregrate,
             color = factor(understandingclause2))) +
  geom_line() +
  geom_point(data = subset(white_reg_rate,
                           understandingclause2 == "Treated"),
             aes(shape = "Treated"), color = "black", fill = "black",
             size = 2) +
  geom_point(data = subset(white_reg_rate,
                           understandingclause2 == "Control"),
             aes(shape = "Control"), color = "black", fill = "black",
```

```
            size = 2) +
  labs(x = "Year", y = "White Registration Rate") +
  scale_color_manual(values = c("#AA4A44", "#FFC300"),
                     labels = c("Treated", "Control")) +
  scale_shape_manual(values = c("Treated" = 16, "Control" = 17),
                     labels = c("Treated", "Control")) +
  guides(color = guide_legend(title = "Group")) +
  theme_minimal()

p1 + p2 + plot_layout(widths = c(8, 8), heights = c(2))
```




```
ggsave("combined_plot.png", width = 8, height = 2)
```