

Problem Set 5

Anna Weiner

2024-03-18

Problem 1: Replication of Figure 2

```
df =  
  read_dta("~/Downloads/Programming/Problem-Set-5/la_turnout_basic.dta")  
  
df1 =  
  df |>  
  mutate(  
    uc = FctWhen(  
      understandingclause2 == 0 ~ 'Control',  
      understandingclause2 == 1 ~ 'Treated'  
    ),  
  ) |>  
  filter(  
    year >= 1950 & year <= 1970  
  ) |>  
  filter(!is.na(whiteregrate)) |>  
  filter(!is.na(blackregrate))  
  
df2 = df1 |>  
  group_by(year, uc == 'Control') |>  
  mutate(  
    blackavg = mean(blackregrate, na.rm = T),  
    whiteavg = mean(whiteregrate, na.rm = T)  
  )  
  
palette_ugly = c('purple4', 'yellow')  
  
blackreg = df2 |>  
  ggplot(  
    aes(  
      x = year,
```

```

    y = blackavg,
    shape = uc,
    color = uc
  )
) +
geom_line() +
geom_point(color = "black") +
labs(
  caption = "(a) Black Registration",
  x = "Year",
  y = "Black Registration Rates"
) +
scale_x_continuous(
  breaks = seq(1950, 1970, by = 5)
) +
theme(legend.position = "right",
  legend.title = element_blank(),
  plot.caption = element_text(hjust = 0.5,
                              size = 13),
  panel.grid.major = element_line(color = 'grey90'),
  panel.grid.minor = element_line(color = 'grey90'),
  panel.background = element_blank(),
  legend.background = element_rect(fill = 'white', size = .5),
  axis.text.y = element_text(size=10),
  axis.text.x = element_text(size=10),
  panel.border = element_rect(color = 'black', fill = NA, size = 1)
) +
scale_color_manual(values = palette_ugly)

whitereg = df2 |>
ggplot(
  aes(
    x = year,
    y = whiteavg,
    shape = uc,
    color = uc
  )
) +
geom_line() +
geom_point(color = "black") +
labs(
  caption = "(b) White Registration",
  x = "Year",
  y = "White Registration Rates"
)

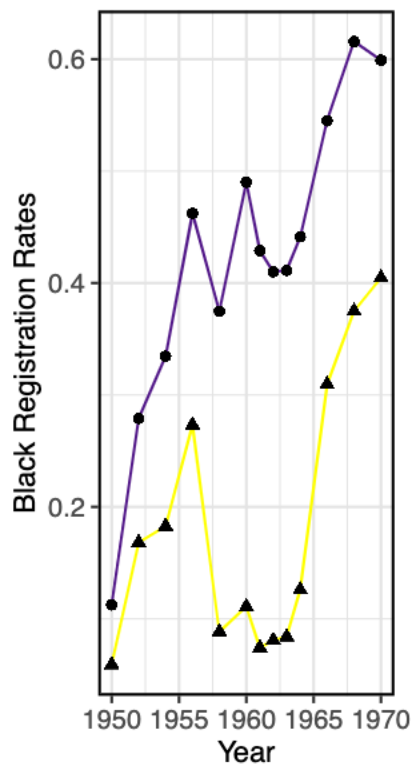
```

```

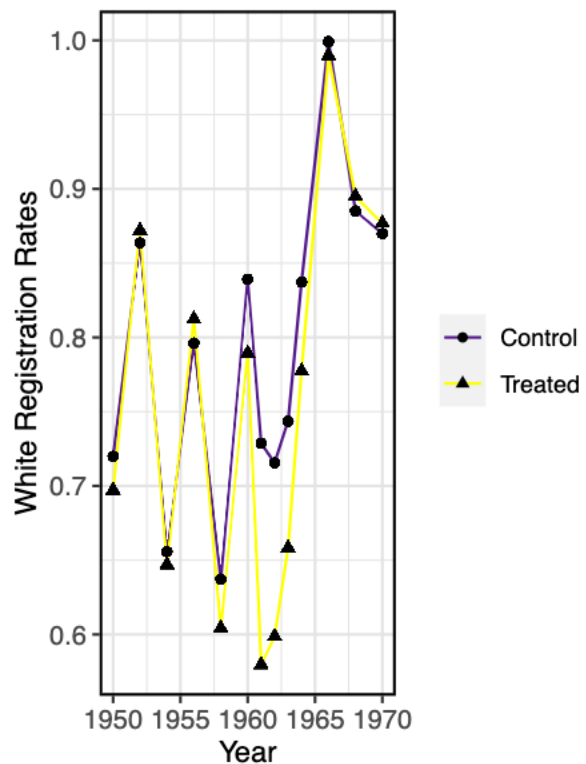
) +
scale_x_continuous(
  breaks = seq(1950, 1970, by = 5)
) +
theme(legend.position = "right",
      legend.title = element_blank(),
      plot.caption = element_text(hjust = 0.5,
                                   size = 13),
      panel.grid.major = element_line(color = 'grey90'),
      panel.grid.minor = element_line(color = 'grey90'),
      panel.background = element_blank(),
      legend.background = element_rect(fill = 'white', size = .5),
      axis.text.y = element_text(size=10),
      axis.text.x = element_text(size=10),
      panel.border = element_rect(color = 'black', fill = NA, size = 1)
) +
scale_color_manual(values = palette_ugly)

```

blackreg + whitereg



(a) Black Registration



(b) White Registration

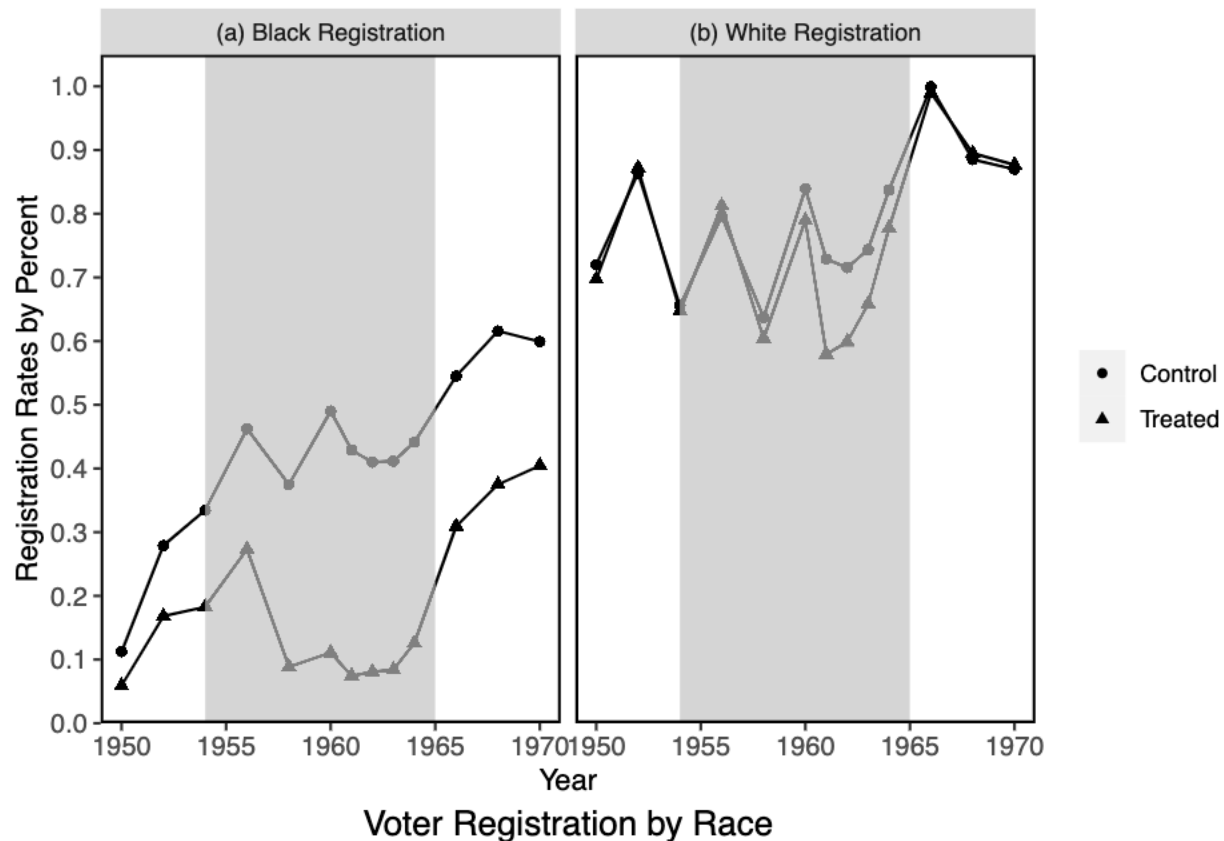
Problem 2: Improving Figure 2

```
df3 = df2 |>
  select(
    year,
    Parish,
    uc,
    blackavg,
    whiteavg
  ) |>
  pivot_longer(
    cols = blackavg:whiteavg,
    names_to = 'avg',
    values_drop_na = TRUE
  )

names = list(
  "blackavg" = "(a) Black Registration",
  "whiteavg" = "(b) White Registration"
)

tab = df3 |>
  ggplot(aes(x = year, y = value, shape = uc, fill = uc)) +
  facet_wrap(~ avg, ncol = 2, labeller = variable_labeller) +
  geom_line() +
  geom_point(color = "black") +
  geom_rect(aes(xmin = 1954,
                xmax = 1965,
                ymin = 0,
                ymax = Inf),
            alpha=0.01,
            fill="grey") +
  labs(
    caption = "Voter Registration by Race",
    x = "Year",
    y = "Registration Rates by Percent"
  ) +
  scale_x_continuous(
    breaks = seq(1950, 1970, by = 5)
  ) +
  scale_y_continuous(
    expand = expansion(mult = c(0, 0.05)),
    breaks = seq(0, 1, by = 0.1)
  ) +
  coord_cartesian(ylim = c(0, 1)) +
  theme(legend.position = "right",
```

```
    legend.title = element_blank(),
    plot.caption = element_text(hjust = 0.5,
                                size = 13),
    panel.grid.major = element_blank(),
    panel.grid.minor = element_blank(),
    panel.background = element_blank(),
    legend.background = element_rect(),
    axis.title.x = element_text(
      size = 11
    ),
    axis.title.y = element_text(
      size = 11
    ),
    axis.text.y = element_text(
      size = 10),
    axis.text.x = element_text(
      size = 10),
    panel.border = element_rect(color = 'black',
                                fill = NA,
                                size = 1)
  )
tab
```



Extra Credit: FiveThirtyEight Chart

```

topline = read_csv("approval_topline.csv")
poll_list = read_csv("approval_polllist.csv")

fivethirty = topline |>
  ggplot(
    aes(
      x = end_date
    )
  ) +
  geom_line(aes(y = approve_estimate), color = 'green') +
  geom_line(aes(y = disapprove_estimate), color = 'magenta') +
  labs(
    title = "How unpopular is Joe Biden?",
    x = " ",
    y = " "
  ) +
  scale_x_date(date_labels = "%m/%d",
               expand = expansion(mult = c(0, 0.05)),
               breaks = as.Date(c('2021-01-23', '2022-02-08', '2023-02-25', '2024-03-05'))
  )

```

```
    ) +  
    scale_y_continuous(  
      expand = expansion(mult = c(0, 0.05)),  
      breaks = seq(20, 80, by = 10),  
      limits = c(20, 80)  
    ) +  
    theme(  
      panel.grid.major = element_line(color = 'grey'),  
      panel.grid.minor = element_blank(),  
      panel.background = element_blank(),  
      axis.ticks = element_line(color = 'grey')  
    )  
  
fivethirty + geom_ribbon(  
  aes(ymin = topline$approve_lo,  
      ymax = topline$approve_hi,  
      fill = 'green',  
      linetype = NA),  
  alpha = .2) +  
  geom_ribbon(  
    aes(ymin = topline$disapprove_lo,  
        ymax = topline$disapprove_hi,  
        fill = 'magenta',  
        linetype = NA),  
    alpha = .2) +  
  scale_fill_manual(values=c("green", "magenta")) +  
  guides(fill = FALSE)
```

How unpopular is Joe Biden?

