GOV2001 Replication of Figures and Tables

Replication of Gamm and Kousser (2021)

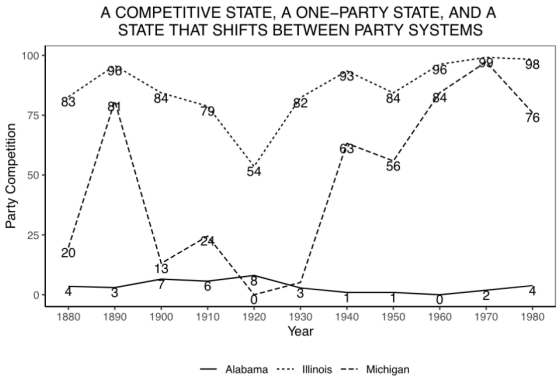
Miro Bergam, Yanxi Fang, and Dominic Skinnion

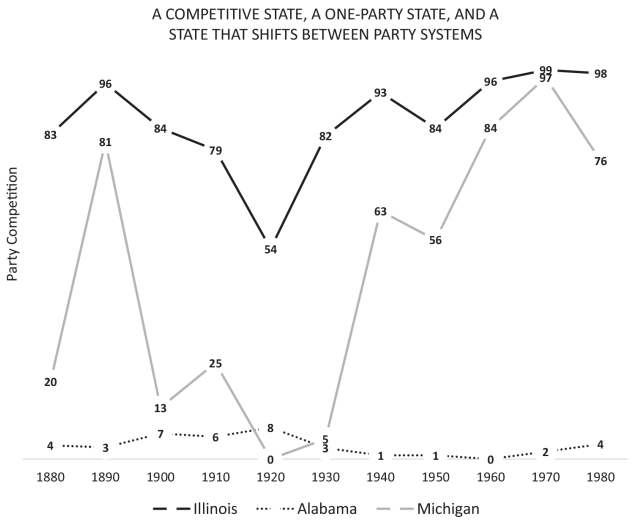
12/12/2021

Figure 1

```
df_fig1 <- df %>%
  filter(state == "Alabama" | state == "Illinois" | state == "Michigan") %>%
  filter(year > 1870 & year < 1990) %>%
  select(year, state, leg_party_competition)

df_fig1 %>%
  ggplot(aes(x = year, y = leg_party_competition,
             label = round(leg_party_competition, digits = 0))) +
  geom_line(aes(linetype = state)) +
  scale_x_continuous(limits = c(1880, 1980),
                     breaks = c(1880, 1890, 1900, 1910, 1920, 1930,
                                1940, 1950, 1960, 1970, 1980)) +
  geom_text(check_overlap = TRUE, vjust = 1) +
  xlab("") +
  ylab("Party Competition") +
  ggtitle("A COMPETITIVE STATE, A ONE-PARTY STATE, AND A \nSTATE THAT SHIFTS BETWEEN PARTY SYSTEMS") +
  theme_bw() +
  theme(panel.grid.major = element_blank(),
        panel.grid.minor = element_blank(),
        panel.background = element_blank(),
        axis.line = element_line(colour = "black")) +
  theme(plot.title = element_text(hjust = 0.5)) +
  theme(legend.position = 'bottom') +
  labs(linetype = '')

ggsave(path = "./figures", filename = "fig1.png")
```

```
knitr::include_graphics(path = c("./figures/fig1.png", "./orig_figures/fig1.png"))
```





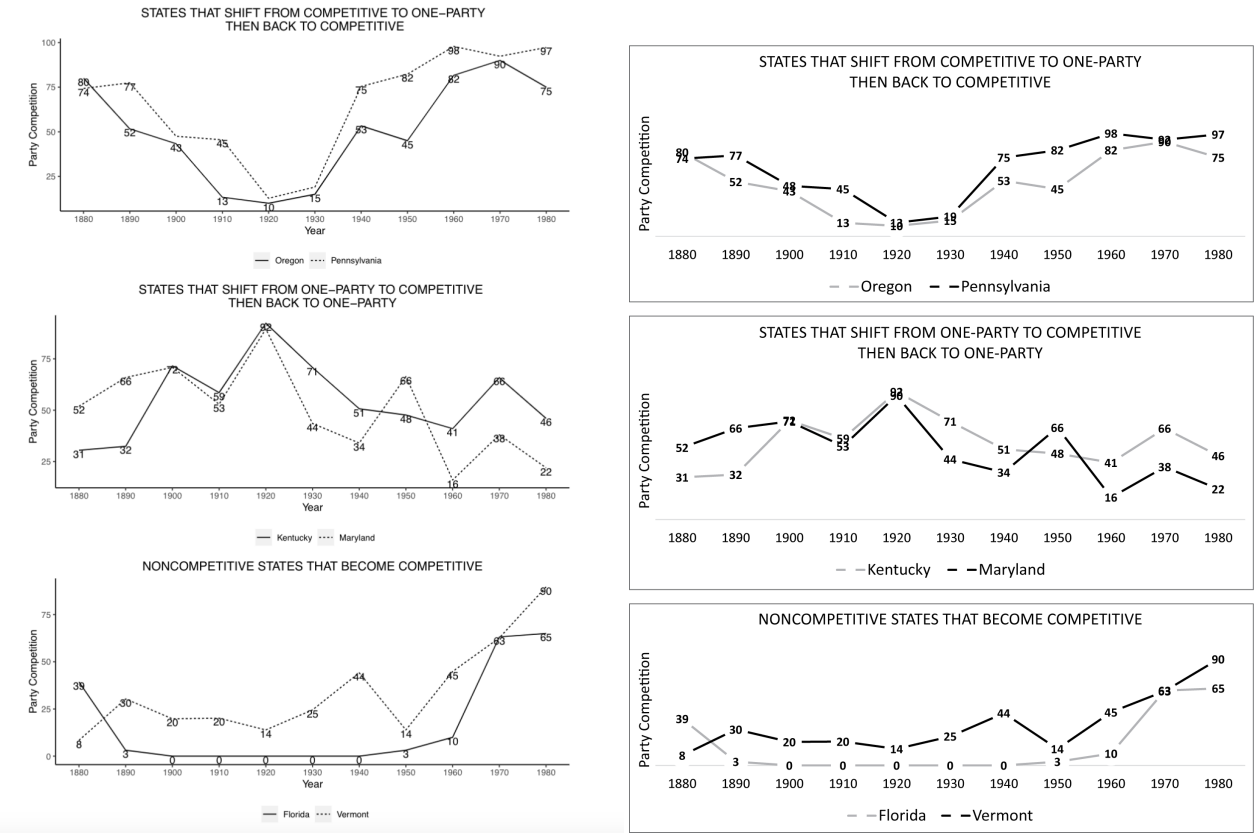
This plot showed three archetypes for the different types of states, ones that are competitive, ones that are one-party states, and ones that switch between the two over time. This figure replicated nicely. One slight change that we could look at would be if we could stop the lines from actually connecting at the numbered points, as was done in the original paper. (This would make the replicated figure a bit easier to read). If we also wanted to faithfully replicate the original, we could remove the x-axis label and the box around the plot.

Figure 2

```
df_fig2a <- df %>%
  filter(state %in% c('Oregon', 'Pennsylvania')) %>%
  filter(year > 1870 & year < 1990) %>%
  select(year, state, leg_party_competition)

df_fig2b <- df %>%
  filter(state %in% c('Kentucky', 'Maryland')) %>%
  filter(year > 1870 & year < 1990) %>%
  select(year, state, leg_party_competition)

df_fig2c <- df %>%
  filter(state %in% c('Florida', 'Vermont')) %>%
  filter(year > 1870 & year < 1990) %>%
  select(year, state, leg_party_competition)


P1 <- df_fig2a %>%
  ggplot(aes(x = year, y = leg_party_competition,
             label = round(leg_party_competition, digits = 0))) +
  geom_line(aes(linetype = state)) +
  theme_bw() +
  scale_x_continuous(limits = c(1880, 1980),
                     breaks = c(1880, 1890, 1900, 1910, 1920, 1930,
                                1940, 1950, 1960, 1970, 1980)) +
  geom_text(check_overlap = TRUE, vjust = 1) +
  xlab("Year") +
  ylab("Party Competition") +
  ggtitle("STATES THAT SHIFT FROM COMPETITIVE TO ONE-PARTY \nTHEN BACK TO COMPETITIVE") +
  theme(panel.grid.major = element_blank(),
        panel.grid.minor = element_blank(),
        panel.background = element_blank(),
        axis.line = element_line(colour = "black")) +
  theme(legend.position = 'bottom',
        plot.title = element_text(hjust = 0.5),
        legend.background = element_rect(fill = 'white')) +
  labs(linetype = '')

P2 <- df_fig2b %>%
  ggplot(aes(x = year, y = leg_party_competition,
             label = round(leg_party_competition, digits = 0))) +
  geom_line(aes(linetype = state)) +
  scale_x_continuous(limits = c(1880, 1980),
                     breaks = c(1880, 1890, 1900, 1910, 1920, 1930,
                                1940, 1950, 1960, 1970, 1980)) +
  geom_text(check_overlap = TRUE, vjust = 1) +
  xlab("Year") +
  ylab("Party Competition") +
  ggtitle("STATES THAT SHIFT FROM ONE-PARTY TO COMPETITIVE \nTHEN BACK TO ONE-PARTY") +
  theme(panel.grid.major = element_blank(),
        panel.grid.minor = element_blank(),
        panel.background = element_blank(),
        axis.line = element_line(colour = "black")) +
```

```
  theme(legend.position = 'bottom',
        plot.title = element_text(hjust = 0.5),
        legend.background = element_rect(fill = 'white')) +
  labs(linetype = '')

P3 <- df_fig2c %>%
  ggplot(aes(x = year, y = leg_party_competition,
             label = round(leg_party_competition, digits = 0))) +
  geom_line(aes(linetype = state)) +
  scale_x_continuous(limits = c(1880, 1980),
                     breaks = c(1880, 1890, 1900, 1910, 1920, 1930,
                                1940, 1950, 1960, 1970, 1980)) +
  geom_text(check_overlap = TRUE, vjust = 1) +
  xlab("Year") +
  ylab("Party Competition") +
  ggtitle("NONCOMPETITIVE STATES THAT BECOME COMPETITIVE") +
  theme(panel.grid.major = element_blank(),
        panel.grid.minor = element_blank(),
        panel.background = element_blank(),
        axis.line = element_line(colour = "black")) +
  theme(legend.position = 'bottom',
        plot.title = element_text(hjust = 0.5),
        legend.background = element_rect(fill = 'white')) +
  labs(linetype = '')

library(cowplot)

plot_grid(P1, P2, P3, nrow = 3)

ggsave(path = "figures", filename = "fig2.png")
```

```
knitr::include_graphics(path = c("figures/fig2.png", "orig_figures/fig2.png"))
```



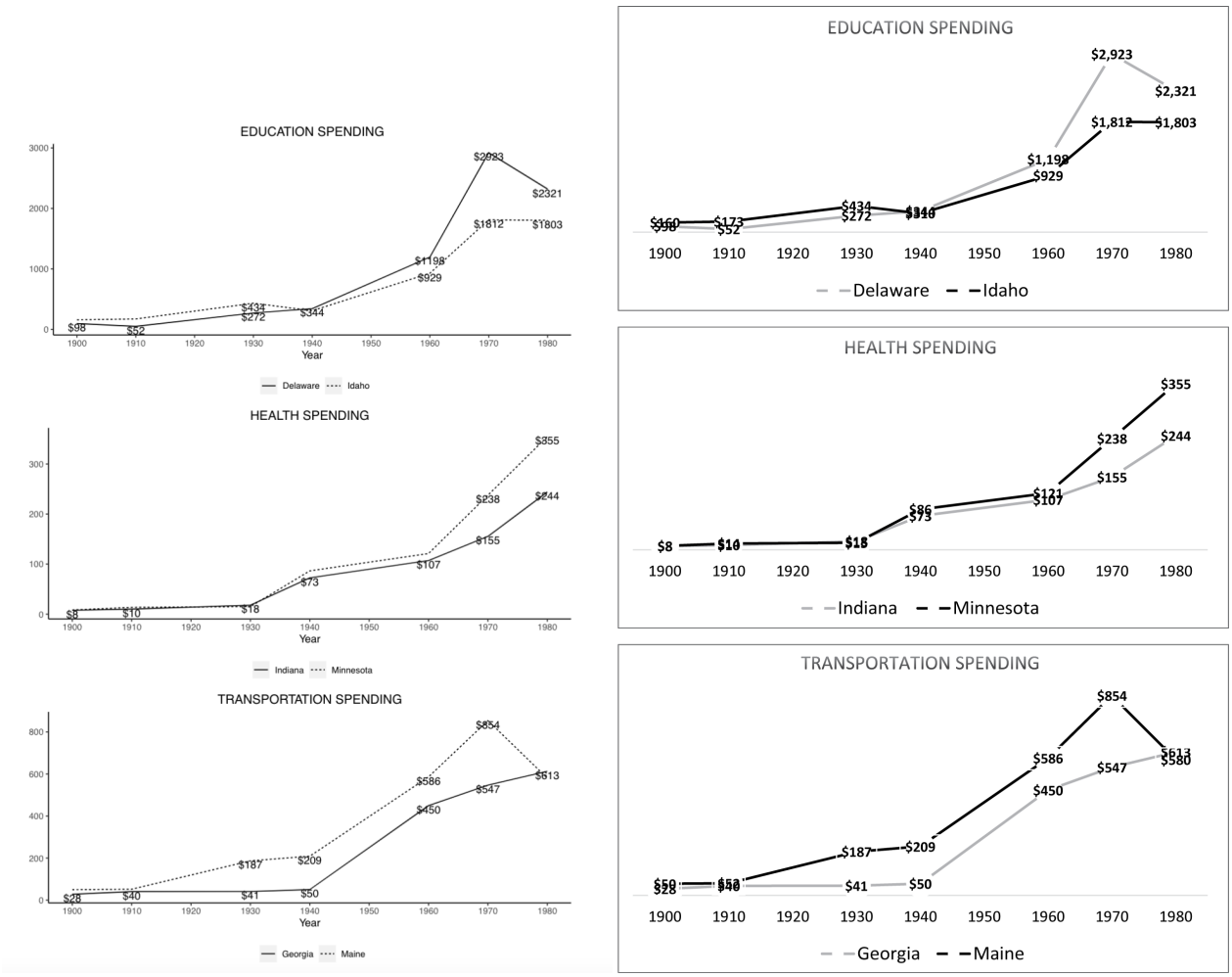
This plot served to show examples of the three different types of archetypal states (as described earlier). In each subplot, there is a typical Republican state and a typical Democratic state to show that there are states from both sides of the aisle that fall into these categories. This figure also replicated nicely. Again, to stay true to the original paper, we could try to have the lines disjoint at the numbered points, remove the x-axis labels, and remove the y-axis line. In addition, we could try to remove the boxes around the legend line-types.

Figure 3

```
df_fig3a <- df %>%
  filter(state %in% c('Delaware', 'Idaho')) %>%
  filter(year >= 1900 & year < 1990) %>%
  select(year, state, Education_pc) %>%
  drop_na()

df_fig3b <- df %>%
  filter(state %in% c('Indiana', 'Minnesota')) %>%
  filter(year >= 1900 & year < 1990) %>%
  select(year, state, HealthSewerSanitation_pc) %>%
  drop_na()

df_fig3c <- df %>%
  filter(state %in% c('Georgia', 'Maine')) %>%
  filter(year >= 1900 & year < 1990) %>%
  select(year, state, Transportation_pc) %>%
  drop_na()


P1 <- df_fig3a %>%
  ggplot(aes(x = year, y = Education_pc,
             label = paste0('$', round(Education_pc, digits = 0)))) +
  geom_line(aes(linetype = state)) +
  scale_x_continuous(limits = c(1900, 1980),
                     breaks = c(1900, 1910, 1920, 1930,
                                1940, 1950, 1960, 1970, 1980)) +
  geom_text(check_overlap = TRUE, vjust = 1) +
  xlab("Year") +
  ylab("") +
  ggtitle("EDUCATION SPENDING") +
  theme(panel.grid.major = element_blank(),
        panel.grid.minor = element_blank(),
        panel.background = element_blank(),
        axis.line = element_line(colour = "black")) +
  theme(legend.position = 'bottom',
        plot.title = element_text(hjust = 0.5),
        legend.background = element_rect(fill = 'white')) +
  labs(linetype = '')

P2 <- df_fig3b %>%
  ggplot(aes(x = year, y = HealthSewerSanitation_pc,
             label = paste0('$', round(HealthSewerSanitation_pc, digits = 0)))) +
  geom_line(aes(linetype = state)) +
  scale_x_continuous(limits = c(1900, 1980),
                     breaks = c(1900, 1910, 1920, 1930,
                                1940, 1950, 1960, 1970, 1980)) +
  geom_text(check_overlap = TRUE, vjust = 1) +
  xlab("Year") +
  ylab("") +
  ggtitle("HEALTH SPENDING") +
  theme(panel.grid.major = element_blank(),
        panel.grid.minor = element_blank(),
```

```
        panel.background = element_blank(),
        axis.line = element_line(colour = "black")) +
  theme(legend.position = 'bottom',
        plot.title = element_text(hjust = 0.5),
        legend.background = element_rect(fill = 'white')) +
  labs(linetype = '')

P3 <- df_fig3c %>%
  ggplot(aes(x = year, y = Transportation_pc,
             label = paste0('$', round(Transportation_pc, digits = 0)))) +
  geom_line(aes(linetype = state)) +
  scale_x_continuous(limits = c(1900, 1980),
                     breaks = c(1900, 1910, 1920, 1930,
                                1940, 1950, 1960, 1970, 1980)) +
  geom_text(check_overlap = TRUE, vjust = 1) +
  xlab("Year") +
  ylab("") +
  ggtitle("TRANSPORTATION SPENDING") +
  theme(panel.grid.major = element_blank(),
        panel.grid.minor = element_blank(),
        panel.background = element_blank(),
        axis.line = element_line(colour = "black")) +
  theme(legend.position = 'bottom',
        plot.title = element_text(hjust = 0.5),
        legend.background = element_rect(fill = 'white')) +
  labs(linetype = '')

library(cowplot)

plot_grid(P1, P2, P3, nrow = 3)

ggsave(path = "figures", filename = "fig3.png")
```

```
knitr::include_graphics(path = c("figures/fig3.png", "orig_figures/fig3.png"))
```



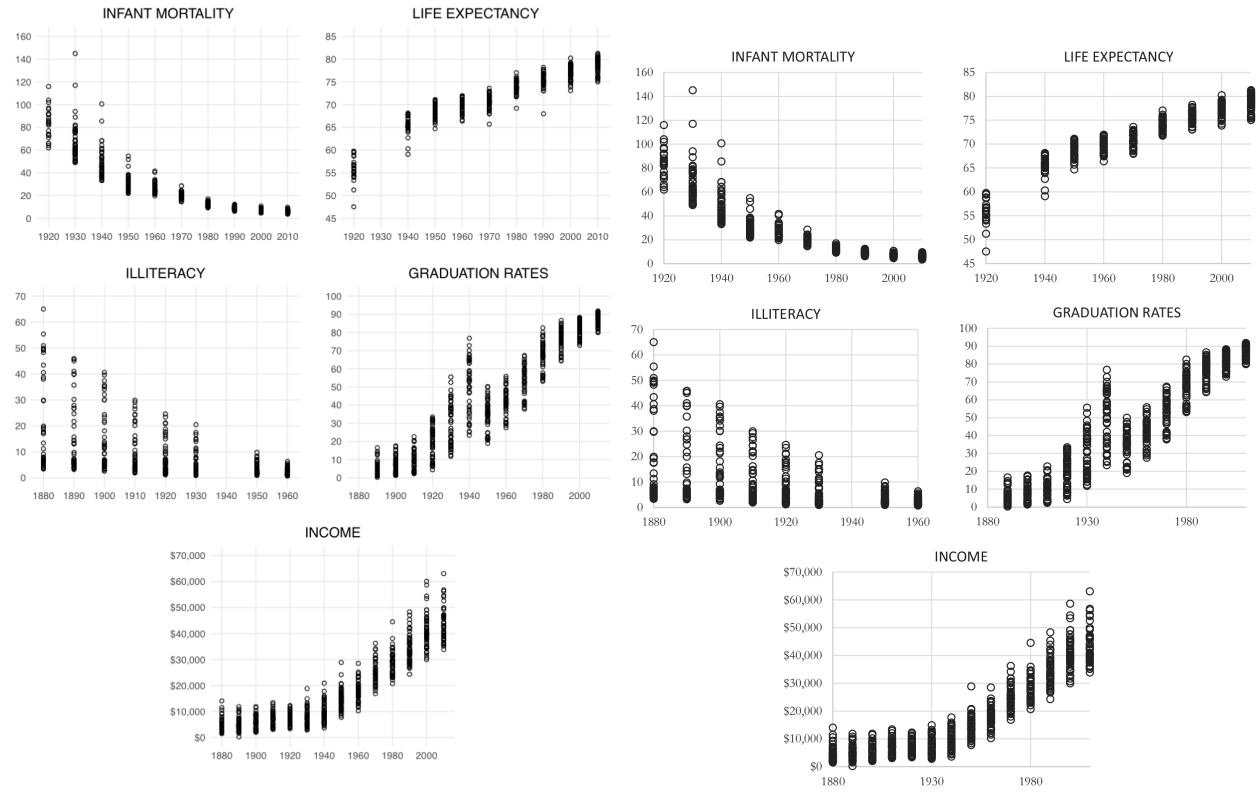
This plot shows the general trends over time in public-goods spending (along three different dimensions), between various pairs of states. This figure was almost the same as the previous (in terms of format). Again, to stay true to the original paper, we could try to have the lines disjoint at the numbered points, remove the x-axis labels, and remove the y-axis line. In addition, we could try to remove the boxes around the legend line-types.

Figure 4

```
df_fig4a <- df %>%
  filter(year >= 1920 & year <= 2010) %>%
  select(year, state, infantmortality)

df_fig4b <- df %>%
  filter(year >= 1920 & year <= 2010) %>%
  select(year, state, at_birth_life_expectancy)

df_fig4c <- df %>%
  filter(year >= 1880 & year <= 2010) %>%
  select(year, state, illiteracy)

df_fig4d <- df %>%
  filter(year >= 1880 & year <= 2010) %>%
  select(year, state, graduation_combined)

df_fig4e <- df %>%
  filter(year >= 1880 & year <= 2010) %>%
  select(year, state, CPI_per_capita_income)

p1 <- df_fig4a %>%
  ggplot() +
  geom_point(aes(x = year, y = infantmortality), shape = 1) +
  theme_minimal() +
  scale_x_continuous(limits = c(1920, 2010),
                     breaks = c(1920, 1930, 1940, 1950, 1960, 1970, 1980,
                                1990, 2000, 2010)) +
  scale_y_continuous(limits = c(0, 160),
                     breaks = seq(0, 160, 20)) +
  labs(x = '',
       y = '',
       title = 'INFANT MORTALITY') +
  theme(plot.title = element_text(hjust = 0.5)) +
  theme(panel.grid.minor = element_blank())

p2 <- df_fig4b %>%
  ggplot() +
  geom_point(aes(x = year, y = at_birth_life_expectancy), shape = 1) +
  theme_minimal() +
  scale_x_continuous(limits = c(1920, 2010),
                     breaks = c(1920, 1930, 1940, 1950, 1960, 1970, 1980,
                                1990, 2000, 2010)) +
  scale_y_continuous(limits = c(45, 85),
                     breaks = seq(45, 85, 5)) +
  labs(x = '',
       y = '',
       title = 'LIFE EXPECTANCY') +
  theme(plot.title = element_text(hjust = 0.5)) +
  theme(panel.grid.minor = element_blank())

p3 <- df_fig4c %>%
  ggplot() +
```

```
  geom_point(aes(x = year, y = illiteracy), shape = 1) +
  theme_minimal() +
  scale_x_continuous(limits = c(1880, 1960),
                     breaks = seq(1880, 1960, 10)) +
  scale_y_continuous(limits = c(0, 70),
                     breaks = seq(0, 70, 10)) +
  labs(x = '',
       y = '',
       title = 'ILLITERACY') +
  theme(plot.title = element_text(hjust = 0.5)) +
  theme(panel.grid.minor = element_blank())

p4 <- df_fig4d %>%
  ggplot() +
  geom_point(aes(x = year, y = graduation_combined), shape = 1) +
  theme_minimal() +
  scale_x_continuous(limits = c(1880, 2010),
                     breaks = seq(1880, 2010, 20)) +
  scale_y_continuous(limits = c(0, 100),
                     breaks = seq(0, 100, 10)) +
  labs(x = '',
       y = '',
       title = 'GRADUATION RATES') +
  theme(plot.title = element_text(hjust = 0.5)) +
  theme(panel.grid.minor = element_blank())

p5 <- df_fig4e %>%
  ggplot() +
  geom_point(aes(x = year, y = CPI_per_capita_income), shape = 1) +
  theme_minimal() +
  scale_x_continuous(limits = c(1880, 2010),
                     breaks = seq(1880, 2010, 20)) +
  scale_y_continuous(limits = c(0, 70000),
                     breaks = seq(0, 70000, 10000),
                     labels=scales::dollar_format()) +
  labs(x = '',
       y = '',
       title = 'INCOME') +
  theme(plot.title = element_text(hjust = 0.5)) +
  theme(panel.grid.minor = element_blank())

library(gridExtra)

grid.arrange(p1, p2, p3, p4, p5, nrow =3,
             widths = c(1, 1, 1, 1),
             layout_matrix = rbind(
               c(1, 1, 2, 2),
               c(3, 3, 4, 4),
               c(NA, 5, 5, NA)
             ))

ggsave(path = "figures", filename = "fig4.png")
```

```
knitr::include_graphics(path = c("figures/fig4.png", "orig_figures/fig4.png"))
```

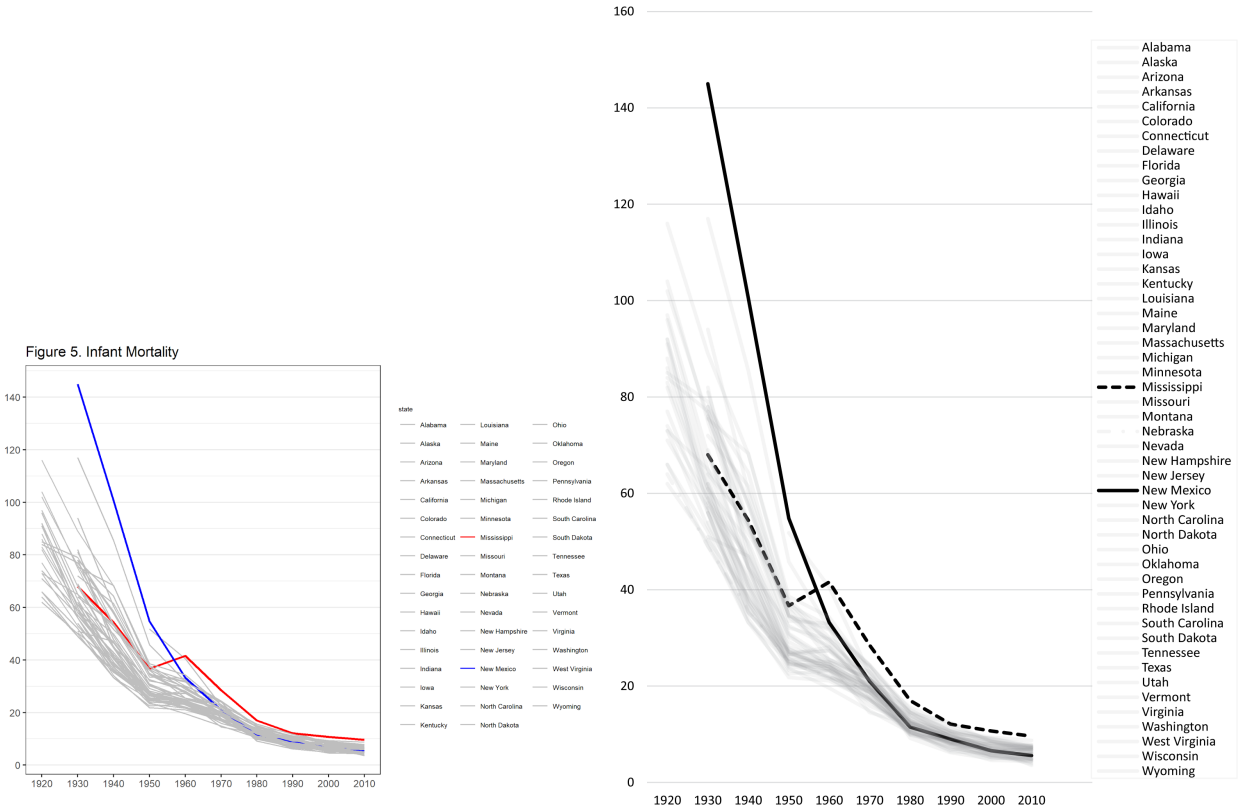


The trends in this figure serve as an overview for the following 5 figures. It shows how infant mortality and illiteracy rates fell across all 50 states over the last several decades, while income, life expectancy, and high school graduation rates increased. There were no problems replicating this figure.

Figure 5

```
df_fig5 <- df %>%
  filter(!is.na(infantmortality)) %>%
  filter(year >= 1920 & year <= 2010) %>%
  select(state, year, infantmortality) %>%
  mutate(stateid = ifelse(state == "New Mexico", 1, ifelse(state == "Mississippi", 2, 0))) %>%
  mutate(stateid = as.factor(stateid))

color_fn <- function(x){ifelse(x == "New Mexico", "blue", ifelse(x == "Mississippi", "red", "gray"))}
colors <- c(color_fn(unique(df_fig5$state)))
size_fn <- function(x){ifelse(x == "New Mexico" | x == "Mississippi", 1, 0.5)}
sizes <- c(size_fn(unique(df_fig5$state)))

f5 <- df_fig5 %>%
  ggplot(aes(x = year, y = infantmortality, group = state, color = state)) +
  geom_line(aes(size = state)) +
  scale_color_manual(values = colors) +
  scale_x_continuous(breaks = c(1920, 1930, 1940, 1950, 1960,
                                1970, 1980, 1990, 2000, 2010)) +
  scale_y_continuous(breaks = c(0, 20, 40, 60, 80, 100, 120, 140, 160)) +
  scale_size_manual(values = sizes) +
  ggtitle("Figure 5. Infant Mortality") + xlab("") + ylab("") +
  theme_bw() +
  theme(panel.grid.major.x = element_blank(),
        panel.grid.minor.x = element_blank())

ggsave(path = "figures", filename = "fig5.png", plot = f5, height = 6, width = 8)
```

```
knitr::include_graphics(path = c("figures/fig5.png", "orig_figures/fig5.png"))
```

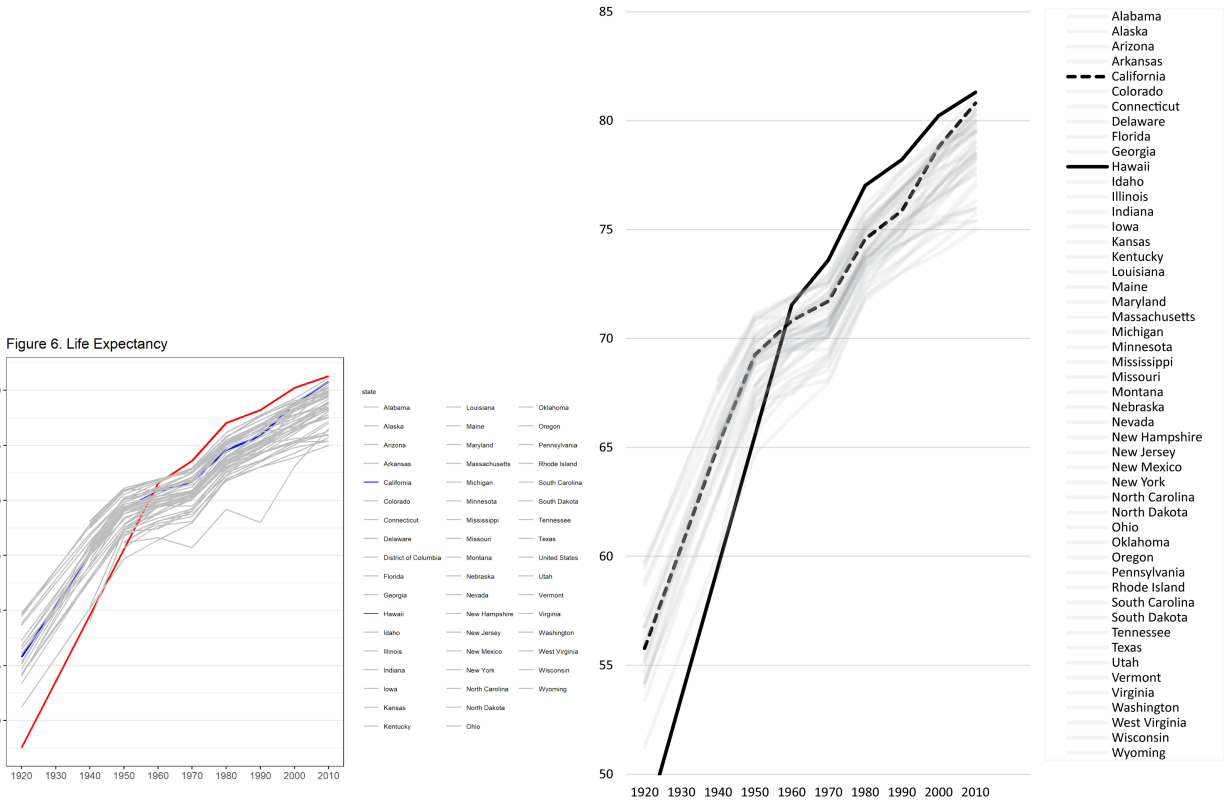


This plot shows how infant mortality steadily decreased across all 50 states over the last several decades, highlighting New Mexico and Mississippi specifically. This plot was simple to replicate, and we used color to highlight the two specific states rather than dotted lines, which makes the plot clearer. To improve this graph, we could try to make the blue and red lines drawn last so that they do not get covered by gray lines.

Figure 6

```
df_fig6 <- df %>%
  filter(!is.na(at_birth_life_expectancy)) %>%
  filter(year >= 1920) %>%
  select(state, year, at_birth_life_expectancy) %>%
  mutate(stateid = ifelse(state == "California", 1, ifelse(state == "Hawaii", 2, 0))) %>%
  mutate(stateid = as.factor(stateid))

color_fn <- function(x){ifelse(x == "California", "blue", ifelse(x == "Hawaii", "red", "gray"))}
colors <- c(color_fn(unique(df_fig6$state)))
size_fn <- function(x){ifelse(x == "California" | x == "Hawaii", 1, 0.5)}
sizes <- c(size_fn(unique(df_fig6$state)))

f6 <- df_fig6 %>%
  ggplot(aes(x = year, y = at_birth_life_expectancy, group = state, color = state)) +
  geom_line(aes(size = state)) +
  scale_color_manual(values = colors) +
  scale_x_continuous(breaks = c(1920, 1930, 1940, 1950, 1960, 1970, 1980, 1990, 2000, 2010)) +
  scale_y_continuous(breaks = c(50, 55, 60, 65, 70, 75, 80, 85)) +
  scale_size_manual(values = sizes) +
  ggtitle("Figure 6. Life Expectancy") + xlab("") + ylab("") +
  theme_bw() +
  theme(panel.grid.major.x = element_blank(),
        panel.grid.minor.x = element_blank())

ggsave(path = "figures", filename = "fig6.png", plot = f6, height = 6, width = 8)
```

```
knitr::include_graphics(path = c("figures/fig6.png", "orig_figures/fig6.png"))
```

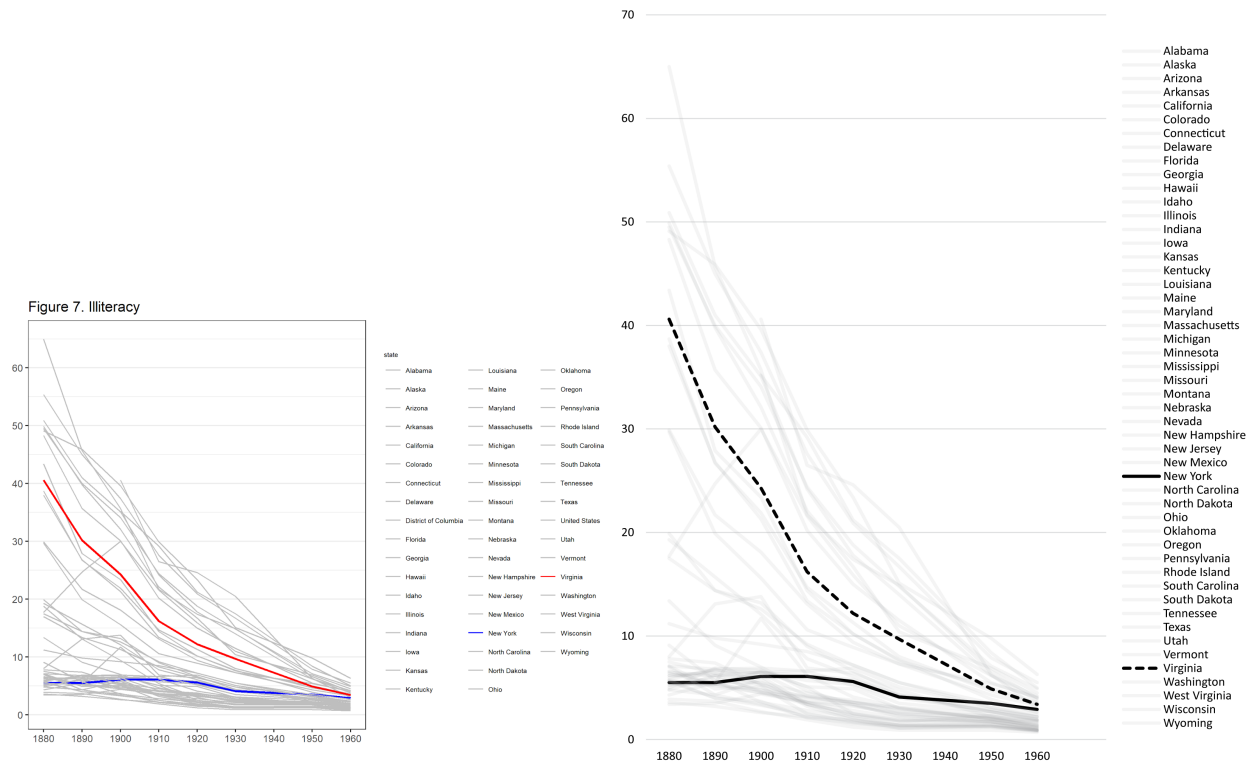


This plot shows how life expectancy steadily increased across all 50 states over the last several decades, highlighting California and Hawaii specifically. This plot was simple to replicate, and we once again added colors. Again, to improve this graph, we could try to make the blue and red lines drawn last so that they do not get covered by gray lines.

Figure 7

```
df_fig7 <- df %>%
  filter(!is.na(illiteracy)) %>%
  filter(year >= 1880 & year <= 1960) %>%
  select(state, year, illiteracy) %>%
  mutate(stateid = ifelse(state == "New York", 1, ifelse(state == "Virginia", 2, 0))) %>%
  mutate(stateid = as.factor(stateid))

color_fn <- function(x){ifelse(x == "New York", "blue", ifelse(x == "Virginia", "red", "gray"))}
colors <- c(color_fn(unique(df_fig7$state)))
size_fn <- function(x){ifelse(x == "New York" | x == "Virginia", 1, 0.5)}
sizes <- c(size_fn(unique(df_fig7$state)))

f7 <- df_fig7 %>%
  ggplot(aes(x = year, y = illiteracy, group = state, color = state)) +
  geom_line(aes(size = state)) +
  scale_color_manual(values = colors) +
  scale_x_continuous(breaks = c(1880, 1890, 1900, 1910, 1920, 1930, 1940, 1950, 1960)) +
  scale_y_continuous(breaks = c(0, 10, 20, 30, 40, 50, 60, 70)) +
  scale_size_manual(values = sizes) +
  ggtitle("Figure 7. Illiteracy") + xlab("") + ylab("") +
  theme_bw() +
  theme(panel.grid.major.x = element_blank(),
        panel.grid.minor.x = element_blank())

ggsave(path = "figures", filename = "fig7.png", plot = f7, height = 6, width = 8)
```

```
knitr::include_graphics(path = c("figures/fig7.png", "orig_figures/fig7.png"))
```

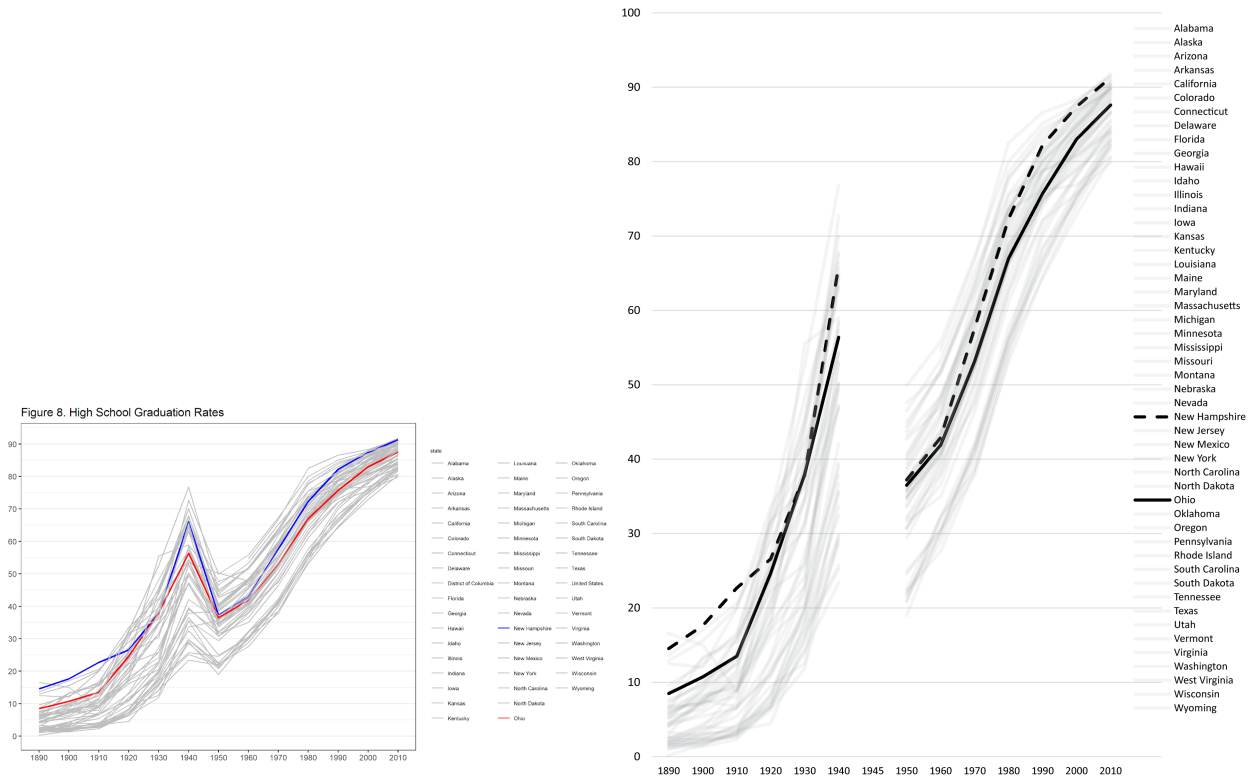


This plot shows how illiteracy steadily decreased across all 50 states over the last several decades, highlighting Virginia and New York specifically. This plot was simple to replicate, and we once again added colors. To improve this graph, we could try to make the blue and red lines drawn last so that they do not get covered by gray lines.

Figure 8

```
df_fig8 <- df %>%
  filter(!is.na(graduation_combined)) %>%
  filter(year >= 1890 & year <= 2010) %>%
  select(state, year, graduation_combined) %>%
  mutate(stateid = ifelse(state == "New Hampshire", 1, ifelse(state == "Ohio", 2, 0))) %>%
  mutate(stateid = as.factor(stateid))

color_fn <- function(x){ifelse(x == "New Hampshire", "blue", ifelse(x == "Ohio", "red", "gray"))}
colors <- c(color_fn(unique(df_fig8$state)))
size_fn <- function(x){ifelse(x == "New Hampshire" | x == "Ohio", 1, 0.5)}
sizes <- c(size_fn(unique(df_fig8$state)))

f8 <- df_fig8 %>%
  ggplot(aes(x = year, y = graduation_combined, group = state, color = state)) +
  geom_line(aes(size = state)) +
  scale_color_manual(values = colors) +
  scale_x_continuous(breaks = c(1890, 1900, 1910, 1920, 1930, 1940, 1950,
                                1960, 1970, 1980, 1990, 2000, 2010)) +
  scale_y_continuous(breaks = c(0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100)) +
  scale_size_manual(values = sizes) +
  ggtitle("Figure 8. High School Graduation Rates") + xlab("") + ylab("") +
  theme_bw() +
  theme(panel.grid.major.x = element_blank(),
        panel.grid.minor.x = element_blank())

ggsave(path = "figures", filename = "fig8.png", plot = f8, height = 6, width = 10)
```

```
knitr::include_graphics(path = c("figures/fig8.png", "orig_figures/fig8.png"))
```

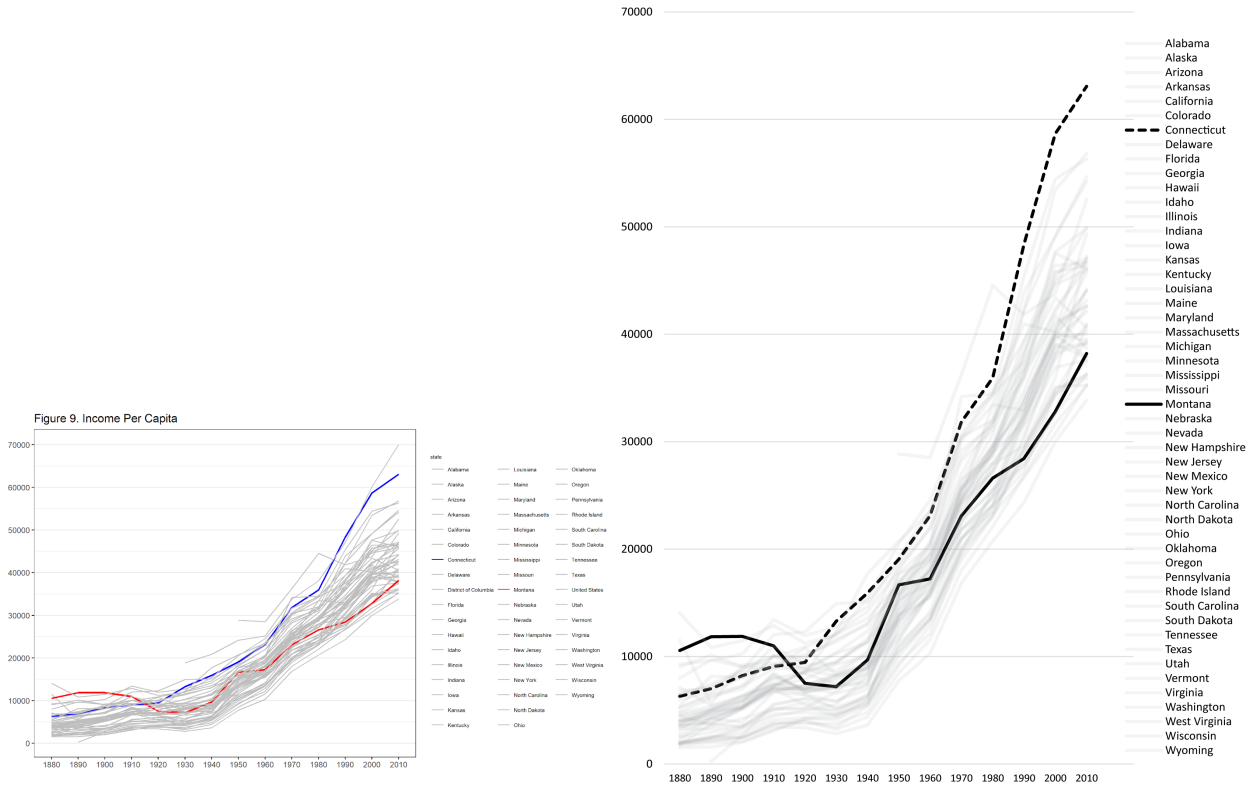


This plot shows how high school graduation rates steadily increased across all 50 states over the last several decades, highlighting New Hampshire and Ohio York specifically. This plot was simple to replicate, and we once again added colors. To stay true to the original paper, we could attempt to remove the middle section of the graph.

Figure 9

```
df_fig9 <- df %>%
  filter(!is.na(CPI_per_capita_income)) %>%
  filter(year >= 1880 & year <= 2010) %>%
  select(state, year, CPI_per_capita_income) %>%
  mutate(stateid = ifelse(state == "Connecticut", 1, ifelse(state == "Montana", 2, 0))) %>%
  mutate(stateid = as.factor(stateid))

color_fn <- function(x){ifelse(x == "Connecticut", "blue", ifelse(x == "Montana", "red", "gray"))}
colors <- c(color_fn(unique(df_fig9$state)))
size_fn <- function(x){ifelse(x == "Connecticut" | x == "Montana", 1, 0.5)}
sizes <- c(size_fn(unique(df_fig9$state)))

f9 <- df_fig9 %>%
  ggplot(aes(x = year, y = CPI_per_capita_income, group = state, color = state)) +
  geom_line(aes(size = state)) +
  scale_color_manual(values = colors) +
  scale_x_continuous(breaks = c(1880, 1890, 1900, 1910, 1920, 1930, 1940,
                                1950, 1960, 1970, 1980, 1990, 2000, 2010)) +
  scale_y_continuous(breaks = c(0, 10000, 20000, 30000, 40000, 50000, 60000, 70000)) +
  scale_size_manual(values = sizes) +
  ggtitle("Figure 9. Income Per Capita") + xlab("") + ylab("") +
  theme_bw() +
  theme(panel.grid.major.x = element_blank(),
        panel.grid.minor.x = element_blank())

ggsave(path = "figures", filename = "fig9.png", plot = f9, height = 6, width = 10)
```

```
knitr::include_graphics(path = c("figures/fig9.png", "orig_figures/fig9.png"))
```



This plot shows how income per capita steadily increased across all 50 states over the last several decades, highlighting Montana and Connecticut specifically. This plot was simple to replicate, and we once again added colors.

Table 1

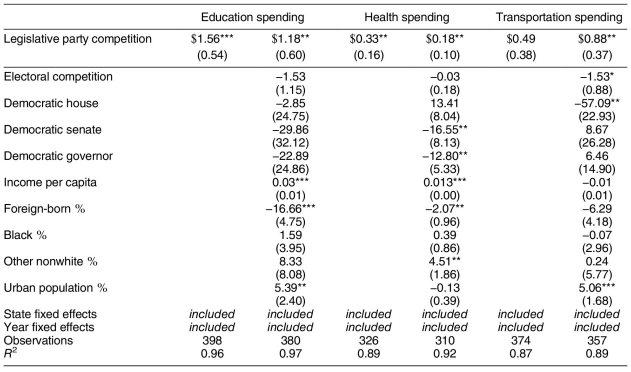
```
# subset data
df_tab1 <- df %>%
  filter(year > 1870 & year < 1990)

# regress model 1
tab1_mod1 <- plm(Education_pc ~ leg_party_competition + year_1890 + year_1900 +
                                year_1910 + year_1930 + year_1940 + year_1960 +
                                year_1970 + year_1980,
                 index = "state", data = df_tab1)
tab1_mod1_se <- coeftest(tab1_mod1, function(x) vcovHC(x, type = 'sss'))

# regress model 2
tab1_mod2 <- plm(Education_pc ~ leg_party_competition + Statewide_Competition +
                                house_dem + senate_dem + gov_dem +
                                CPI_per_capita_income + foreignborn_pct +
                                black_pct + othernonwhite_pct + urban_pct +
                                year_1890 + year_1900 + year_1910 + year_1930 +
                                year_1940 + year_1960 + year_1970 + year_1980,
                 index = "state", data = df_tab1)
tab1_mod2_se <- coeftest(tab1_mod2, function(x) vcovHC(x, type = 'sss'))

# regress model 3
tab1_mod3 <- plm(HealthSewerSanitation_pc ~ leg_party_competition +
                                year_1890 + year_1900 + year_1910 + year_1930 +
                                year_1940 + year_1960 + year_1970 + year_1980,
                 index = "state", data = df_tab1)
tab1_mod3_se <-coeftest(tab1_mod3, function(x) vcovHC(x, type = 'sss'))

# regress model 4
tab1_mod4 <- plm(HealthSewerSanitation_pc ~ leg_party_competition +
                                Statewide_Competition  + house_dem + senate_dem +
                                gov_dem + CPI_per_capita_income + foreignborn_pct +
                                black_pct + othernonwhite_pct + urban_pct +
                                year_1890 + year_1900 + year_1910 + year_1930 +
                                year_1940 + year_1960 + year_1970 + year_1980,
                 index = "state", data = df_tab1)
tab1_mod4_se <-coeftest(tab1_mod4, function(x) vcovHC(x, type = 'sss'))

# regress model 5
tab1_mod5 <- plm(Transportation_pc ~ leg_party_competition +
                                year_1890 + year_1900 + year_1910 + year_1930 +
                                year_1940 + year_1960 + year_1970 + year_1980,
                 index = "state", data = df_tab1)
tab1_mod5_se <-coeftest(tab1_mod5, function(x) vcovHC(x, type = 'sss'))

# regress model 6
tab1_mod6 <- plm(Transportation_pc ~ leg_party_competition +
                                Statewide_Competition  + house_dem + senate_dem +
                                gov_dem + CPI_per_capita_income + foreignborn_pct +
                                black_pct + othernonwhite_pct + urban_pct +
                                year_1890 + year_1900 + year_1910 + year_1930 +
                                year_1940 + year_1960 + year_1970 + year_1980,
```

```
                index = "state", data = df_tab1)
tab1_mod6_se <-coeftest(tab1_mod6, function(x) vcovHC(x, type = 'sss'))

# print Table 1
stargazer(tab1_mod1_se, tab1_mod2_se, tab1_mod3_se,
          tab1_mod4_se, tab1_mod5_se, tab1_mod6_se,
          header = F, type = 'latex', digits = 2,
          title = "Party Competition Predicts Higher Human Capital and Infrastructure Spending, 1880-198
          column.labels = c("Education spending", "Health spending",
                            "Transportation spending"),
          column.separate = c(2, 2, 2),
          covariate.labels = c("Legislative party competition",
                               "Electoral competition", "Democratic house",
                               "Democratic senate", "Democratic governor",
                               "Income per capita", "Foreign-born percentage",
                               "Black percentage", "Other nonwhite percentage",
                               "Urban population percentage"),
          omit = c("Constant", "year_1890", "year_1900", "year_1910", "year_1920",
                   "year_1930", "year_1940", "year_1960", "year_1970", "year_1980"),
          add.lines = list(c("State fixed effects", "included", "included",
                             "included", "included", "included", "included"),
                           c("Year fixed effects", "included", "included",
                             "included", "included", "included", "included"),
                           c("Observations", "398", "380", "326", "310", "374", "357"),
                           c("R-Squared", "0.96", "0.97", "0.89", "0.92", "0.87", "0.89")))

gtsave(path = 'figures', filename = 'tab1.png')
```

```
knitr::include_graphics(path = c("figures/tab1.png", "orig_figures/tab1.png"))
```

Table 1: Party Competition Predicts Higher Human Capital and Infrastructure Spending, 1880-1980

	Dependent variable:					
	Education spending		Health spending		Transportation spending	
	(1)	(2)	(3)	(4)	(5)	(6)
Legislative party competition	1.56***	1.18*	0.33**	0.17*	0.49	0.88**
	(0.54)	(0.60)	(0.16)	(0.10)	(0.38)	(0.37)
Electoral competition		−1.53		−0.03		−1.53*
		(1.15)		(0.18)		(0.88)
Democratic house		−2.85		13.41*		−57.09**
		(24.72)		(8.02)		(22.89)
Democratic senate		−29.86		−16.55**		8.67
		(32.08)		(8.11)		(26.24)
Democratic governor		−22.89		−12.80**		6.46
		(24.83)		(5.32)		(14.88)
Income per capita		0.03***		0.01***		−0.01
		(0.01)		(0.002)		(0.01)
Foreign-born percentage		−16.66***		−2.07**		−6.29
		(4.74)		(0.95)		(4.17)
Black percentage		1.59		0.39		−0.07
		(3.95)		(0.86)		(2.95)
Other nonwhite percentage		8.32		4.51**		0.24
		(8.07)		(1.86)		(5.76)
Urban population percentage		5.39**		−0.13		5.06***
		(2.40)		(0.39)		(1.68)
State fixed effects	included	included	included	included	included	included
Year fixed effects	included	included	included	included	included	included
Observations	398	380	326	310	374	357
R-Squared	0.96	0.97	0.89	0.92	0.87	0.89
Note:						*p<0.1; **p<0.05; ***p<0.01



	Education spending		Health spending		Transportation spending	
Legislative party competition	$1.56***	$1.18**	$0.33**	$0.18**	$0.49	$0.88**
	(0.54)	(0.60)	(0.16)	(0.10)	(0.38)	(0.37)
Electoral competition		−1.53		−0.03		−1.53*
		(1.15)		(0.18)		(0.88)
Democratic house		−2.85		13.41		−57.09**
		(24.75)		(8.04)		(22.93)
Democratic senate		−29.86		−16.55**		8.67
		(32.12)		(8.13)		(26.28)
Democratic governor		−22.89		−12.80**		6.46
		(24.86)		(5.33)		(14.90)
Income per capita		0.03***		0.013***		−0.01
		(0.01)		(0.00)		(0.01)
Foreign-born %		−16.66***		−2.07**		−6.29
		(4.75)		(0.96)		(4.18)
Black %		1.59		0.39		−0.07
		(3.95)		(0.86)		(2.96)
Other nonwhite %		8.33		4.51**		0.24
		(8.08)		(1.86)		(5.77)
Urban population %		5.39**		−0.13		5.06***
		(2.40)		(0.39)		(1.68)
State fixed effects	*included*	*included*	*included*	*included*	*included*	*included*
Year fixed effects	*included*	*included*	*included*	*included*	*included*	*included*
Observations	398	380	326	310	374	357
R^2	0.96	0.97	0.89	0.92	0.87	0.89

This regression table is intended to reflect the article's main argument that legislative party competition leads to higher spending on various forms of human capital and infrastructure, a relationship that holds with or without control variables. The results did not exactly replicate: while the coefficients are all identical,

there are very small differences in some of the standard errors (although these differences are all less than 0.05). We did translate the table's code from Stata to R, so we speculate that the differences were likely caused by slight variations in how Stata and R compute the errors.

Table 2

```
# subset for model 1
df_tab2_mod1 <- df %>%
  filter(year >= 1930 & year < 2020)

# regress model 1
tab2_mod1 <- plm(infantmortality ~ HealthSewerSanitation_pc + CPI_per_capita_income +
                                   foreignborn_pct + black_pct + othernonwhite_pct +
                                   urban_pct + as.factor(year),
                 index = "state", data = df_tab2_mod1)
tab2_mod1_se <- coeftest(tab2_mod1, function(x) vcovHC(x, type = 'sss'))

# subset for model 2
df_tab2_mod2 <- df %>%
  filter(year >= 1880 & year <= 2010) %>%
  mutate(f3_at_birth_life_expectancy = dplyr::lead(at_birth_life_expectancy, 3)) %>%
  filter(year <= 1980)

# regress model 2
tab2_mod2 <- plm(f3_at_birth_life_expectancy ~ HealthSewerSanitation_pc +
                                   CPI_per_capita_income + foreignborn_pct +
                                   black_pct + othernonwhite_pct + urban_pct +
                                   as.factor(year),
                 index = "state", data = df_tab2_mod2)
tab2_mod2_se <- coeftest(tab2_mod2, function(x) vcovHC(x, type = 'sss'))

# subset for model 3
df_tab2_mod3 <- df %>%
  filter(year >= 1880 & year <= 2010)

# regress model 3
tab2_mod3 <- plm(graduation_combined ~ Education_pc + CPI_per_capita_income +
                                   foreignborn_pct + black_pct+ othernonwhite_pct +
                                   urban_pct + south + as.factor(year),
                 index = "state", data = df_tab2_mod3)
tab2_mod3_se <- coeftest(tab2_mod3, function(x) vcovHC(x, type = 'sss'))

# use same data as model 3, regress model 4
tab2_mod4 <- plm(illiteracy_proportional_30 ~ Education_pc + CPI_per_capita_income +
                                   foreignborn_pct + black_pct + othernonwhite_pct +
                                   urban_pct + south + as.factor(year),
                 index = "state", data = df_tab2_mod3)
tab2_mod4_se <- coeftest(tab2_mod4, function(x) vcovHC(x, type = 'sss'))

# print Table 2
stargazer(tab2_mod1_se, tab2_mod2_se, tab2_mod3_se, tab2_mod4_se,
          header = F, type = 'latex',
          title = "Spending Levels Predict Development, 1880-2010",
          column.labels = c("Infant mortality",
                            "Life expectancy (30 years later)",
                            "High school completion",
                            "Illiteracy rate (30 years later)"),
          covariate.labels = c("Health, sewer, sanitation spending per capita",
```

```
                              "Education spending per capita",
                              "Income per capita",
                              "Foreign-born percentage", "Black percentage",
                              "Other nonwhite percentage", "Urban population percentage"),
          omit = c("Constant", "south", "year"),
          add.lines = list(c("State fixed effects",
                             "included", "included", "included", "included"),
                           c("Year fixed effects",
                             "included", "included", "included", "included"),
                           c("Observations", "240", "272", "374", "168"),
                           c("R-Squared", "0.92", "0.98", "0.96", "0.43")))

gtsave(path = 'figures', filename = 'tab2.png')
```

```
knitr::include_graphics(path = c("figures/tab2.png", "orig_figures/tab2.png"))
```

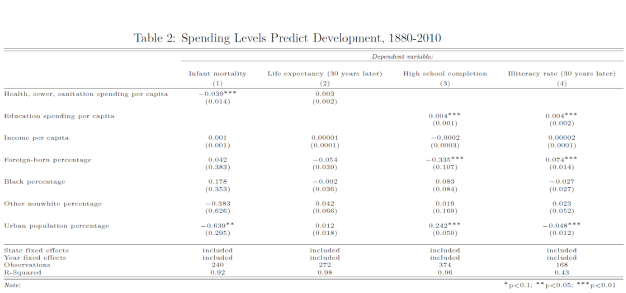


Table 2: Spending Levels Predict Development, 1880-2010

	Dependent variable:			
	Infant mortality (1)	Life expectancy (30 years later) (2)	High school completion (3)	Illiteracy rate (30 years later) (4)
Health, sewer, sanitation spending per capita	−0.039*** (0.014)	0.003 (0.002)		
Education spending per capita			0.004*** (0.001)	0.004*** (0.002)
Income per capita	0.001 (0.001)	0.00001 (0.0001)	−0.0002 (0.0003)	0.00002 (0.0001)
Foreign-born percentage	0.042 (0.383)	−0.054 (0.039)	−0.335*** (0.107)	0.074*** (0.014)
Black percentage	0.178 (0.353)	−0.002 (0.036)	0.083 (0.084)	−0.027 (0.027)
Other nonwhite percentage	−0.383 (0.626)	0.042 (0.066)	0.019 (0.169)	0.023 (0.052)
Urban population percentage	−0.639** (0.295)	0.012 (0.018)	0.242*** (0.050)	−0.048*** (0.012)
State fixed effects	included	included	included	included
Year fixed effects	included	included	included	included
Observations	240	272	374	168
R-Squared	0.92	0.98	0.96	0.43
Note:				*p<0.1; **p<0.05; ***p<0.01

	Infant mortality	Life expectancy (30 years later)	High school completion	Illiteracy rate (30 years later)
Health, sewer, and sanitation spending per capita	−0.039*** (0.01)	0.003* (0.002)		
Education spending per capita			0.004*** (0.001)	0.004** (0.002)
Income per capita	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Foreign-born %	0.04 (0.38)	−0.05 (0.04)	−0.34*** (0.11)	0.074*** (0.01)
Black %	0.18 (0.35)	0.00 (0.04)	0.08 (0.08)	−0.03 (0.03)
Other nonwhite %	−0.38 (0.63)	0.04 (0.07)	0.02 (0.17)	0.02 (0.05)
Urban population %	−0.64** (0.30)	0.01 (0.02)	0.24*** (0.05)	−0.048*** (0.01)
State fixed effects	*included*	*included*	*included*	*included*
Year fixed effects	*included*	*included*	*included*	*included*
Observations	240	272	374	168
R^2	0.92	0.98	0.96	0.43

This regression table is intended to reflect the second part of the article's main argument, that increased spending levels (from legislative party competition) leads to better outcomes for states' populations. Here, the results also did not exactly replicate: all the coefficients and standard errors are identical except one standard error (which is off by 0.01). Again, we did translate the table's code from Stata to R, so we speculate that the difference may be caused by slight variations in how Stata and R compute standard errors.

Table 3

```
# create new dataframe for "full sample"
df_tab3_full <- df

# subset for 1880-1940 sample
df_tab3_part <- df %>%
  filter(year >= 1880 & year <= 1940)

# regress model 1
tab3_mod1 <- plm(CPI_per_capita_income_next30 ~ HealthSewerSanitation_pc +
                   CPI_per_capita_income + foreignborn_pct + black_pct +
                   othernonwhite_pct + urban_pct + south + as.factor(year),
                 index = "state", data = df_tab3_full)
tab3_mod1_se <- coeftest(tab3_mod1, function(x) vcovHC(x, type = 'sss'))

# regress model 2
tab3_mod2 <- plm(CPI_per_capita_income_next30 ~ HealthSewerSanitation_pc +
                   CPI_per_capita_income + foreignborn_pct + black_pct +
                   othernonwhite_pct + urban_pct + south + as.factor(year),
                 index = "state", data = df_tab3_part)
tab3_mod2_se <- coeftest(tab3_mod2, function(x) vcovHC(x, type = 'sss'))

# regress model 3
tab3_mod3 <- plm(CPI_per_capita_income_next30 ~ Education_pc +
                   CPI_per_capita_income + foreignborn_pct + black_pct +
                   othernonwhite_pct + urban_pct + south + as.factor(year),
                 index = "state", data = df_tab3_full)
tab3_mod3_se <- coeftest(tab3_mod3, function(x) vcovHC(x, type = 'sss'))

# regress model 4
tab3_mod4 <- plm(CPI_per_capita_income_next30 ~ Education_pc +
                   CPI_per_capita_income + foreignborn_pct + black_pct +
                   othernonwhite_pct + urban_pct + south + as.factor(year),
                 index = "state", data = df_tab3_part)
tab3_mod4_se <- coeftest(tab3_mod4, function(x) vcovHC(x, type = 'sss'))

# regress model 5
tab3_mod5 <- plm(CPI_per_capita_income_next30 ~ Transportation_pc +
                   CPI_per_capita_income + foreignborn_pct + black_pct +
                   othernonwhite_pct + urban_pct + south + as.factor(year),
                 index = "state", data = df_tab3_full)
tab3_mod5_se <- coeftest(tab3_mod5, function(x) vcovHC(x, type = 'sss'))

# regress model 6
tab3_mod6 <- plm(CPI_per_capita_income_next30 ~ Transportation_pc +
                   CPI_per_capita_income + foreignborn_pct + black_pct +
                   othernonwhite_pct + urban_pct + south + as.factor(year),
                 index = "state", data = df_tab3_part)
tab3_mod6_se <- coeftest(tab3_mod6, function(x) vcovHC(x, type = 'sss'))

# print Table 3
stargazer(tab3_mod1_se, tab3_mod2_se, tab3_mod3_se,
          tab3_mod4_se, tab3_mod5_se, tab3_mod6_se,
```

```
        header = F, type = 'latex',
        title = "Health and Education Spending Levels Predict Income (Only in Pre-New Deal Period)",
        column.labels = c("Full sample", "1880-1940",
                          "Full sample", "1880-1940",
                          "Full sample", "1880-1940"),
        covariate.labels = c("Health, sewer, sanitation spending per capita",
                             "Education spending per capita",
                             "Transportation spending per capita",
                             "Income per capita",
                             "Foreign-born pct", "Black pct",
                             "Other nonwhite pct", "Urban population pct"),
        omit = c("Constant", "south", "year"),
        add.lines = list(c("State fixed effects",
                           "included", "included", "included",
                           "included", "included", "included"),
                         c("Year fixed effects",
                           "included", "included", "included",
                           "included", "included", "included"),
                         c("Observations", "336", "192", "408",
                           "264", "384", "240"),
                         c("R-Squared", "0.98", "0.99", "0.98",
                           "0.98", "0.98", "0.98")))

gtsave(path = 'figures', filename = 'tab3.png')
```

```
knitr::include_graphics(path = c("figures/tab3.png", "orig_figures/tab3.png"))
```

Table 3: Health and Education Spending Levels Predict Income (Only in Pre-New Deal Period)

	Income per capita (30 years later)					
	Full sample (1)	1880-1940 (2)	Full sample (3)	1880-1940 (4)	Full sample (5)	1880-1940 (6)
Health, sewer, sanitation spending per capita	−0.451 (4.252)	14.031*** (4.364)				
Education spending per capita			0.290 (0.913)	4.610** (1.812)		
Transportation spending per capita					−1.397 (1.517)	0.291 (2.679)
Income per capita	−0.226 (0.200)	−1.118*** (0.084)	−0.064 (0.189)	−0.806*** (0.089)	−0.104 (0.179)	−0.833*** (0.094)
Foreign-born percentage	−187.541*** (69.461)	−32.419 (32.928)	−111.734** (46.789)	15.878 (31.650)	−151.163*** (49.799)	−14.037 (29.393)
Black percentage	−159.075*** (51.383)	−45.252 (41.556)	−121.902*** (42.401)	−38.404 (46.454)	−126.182*** (47.841)	−9.529 (52.326)
Other nonwhite percentage	17.789 (76.291)	−73.770 (73.158)	38.685 (80.142)	−71.910 (107.354)	33.936 (81.652)	−73.773 (76.452)
Urban population percentage	−94.396*** (29.362)	41.979** (21.214)	−101.181*** (25.780)	20.396 (18.841)	−94.838*** (26.732)	28.933 (22.083)
State fixed effects	included	included	included	included	included	included
Year fixed effects	included	included	included	included	included	included
Observations	336	192	408	264	384	240
R-Squared	0.98	0.99	0.98	0.98	0.98	0.98
Note:						*p<0.1; **p<0.05; ***p<0.01

	Income per capita (30 years later)		Income per capita (30 years later)		Income per capita (30 years later)	
	Full sample	1880–1940	Full sample	1880–1940	Full sample	1880–1940
Health, sewer, and sanitation spending per capita	−0.45 (4.26)	14.03*** (4.38)				
Education spending per capita			0.29 (0.91)	4.61** (1.82)		
Transportation spending per capita					−1.40 (1.52)	0.29 (2.69)
Income per capita	0.77*** (0.20)	−0.12 (0.08)	0.94*** (0.19)	0.19** (0.09)	0.90*** (0.18)	0.17* (0.09)
Foreign-born %	−188*** (69.57)	−32.42 (33.02)	−111.7** (46.85)	15.88 (31.71)	−151*** (49.87)	−14.04 (29.46)
Black %	−159*** (51.46)	−45.25 (41.67)	−122*** (42.45)	−38.40 (46.55)	−126** (47.91)	−9.53 (52.44)
Other nonwhite %	17.79 (76.41)	−73.77 (73.36)	38.69 (80.24)	−71.91 (107.60)	33.94 (81.76)	−73.77 (76.62)
Urban population %	−94*** (29.41)	41.98* (21.27)	−101*** (25.81)	20.40 (18.88)	−95*** (26.77)	28.93 (22.13)
State fixed effects	*included*	*included*	*included*	*included*	*included*	*included*
Year fixed effects	*included*	*included*	*included*	*included*	*included*	*included*
Observations	336	192	408	264	384	240
R^2	0.98	0.99	0.98	0.98	0.98	0.98

This regression table is intended to reflect a nuance in the article's argument: while higher spending leads to better outcomes in areas like infant mortality, the "better" does not apply to income, at least not in the post-New Deal era. Once again, the results did not exactly replicate: while most of the coefficients are identical, all six coefficients for the income per capita variable were incorrect; in addition, the standard errors are slightly different, with the biggest difference being 0.20). While the difference in standard errors may have been caused by variations in how Stata and R compute standard errors, we are uncertain about why the income coefficients are entirely different, especially because the rest of the coefficients are right.