Single-Member Districting Mandate

Introduction

For much of early American history, the Federal government gave states wide latitude to decide how their own elections would be held, even if those elections were to the U.S. Congress. This led to a number of interesting differences in how states conducted their elections. Some worked by first past the post systems; others established runoffs if no candidate won a majority, occasionally generating as many as seven elections for one seat. States held their elections at vastly different times, with elections for one Congress sometimes taking place over the course of a full year.

Additionally, states had the exclusive power to determine constituency. Along with controlling voting eligiblity, they were allowed to district how they pleased. Some states chose to set up at-large elections. Georgia, for instance, in 1838 could elect 9 representatives to the U.S. House. It chose to set up a system whereby each party would nominate 9 candidates, each voter would cast 9 votes, and the top 9 candidates receiving the most votes would win.

In 1842, Congress inserted a controversial provision into the Apportionment Act. It mandated that all states with multi-member districts to elect U.S. Representatives must switch to a single-member system. Why, exactly, did Congress want to make such a change?

Analysis

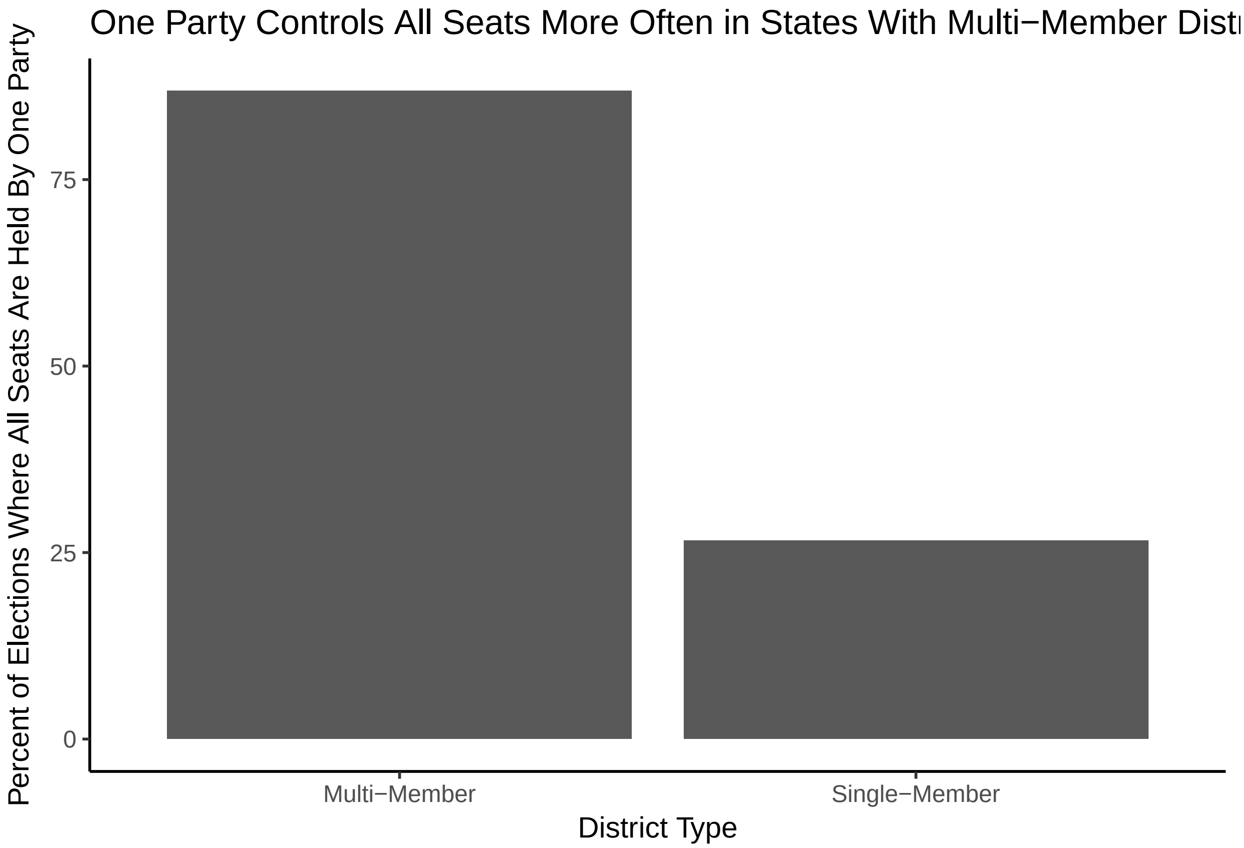
Stephen Calabrese hypothesized that because single-member districts tended to "sweep" a single party into every seat of the state's delegation, and because voters tend to "moderate" the majority party, the majority party (as represented in Congress) would be incentivized to mandate single-member districts to minimize losses.

```
second_largest = function(numbers) {
  largest = 0
  second_largest = 0
  for (number in numbers) {
    if (number > largest) {
      second_largest = largest
      largest = number
    } else if (number > second_largest) {
      second_largest = number
    }
  }
  return(second_largest)
}

filtered <-  filtered %>%
  filter(mean_elected_district >= 1, num_elected_state > 1, as.integer(congress) <= 27)

minority_power <- filtered %>%
  group_by(state, congress) %>%
  summarise(minority_stake=second_largest(actual), is_single_member=ifelse(first(mean_elected_district)
            num_elected_state=max(num_elected_state), mean_elected_district=mean(mean_elected_district))
  ungroup() %>% group_by(as.factor(num_elected_state)) %>% filter(n() > 1)


predicted_stake = predict(lm(minority_stake ~ num_elected_state, minority_power))
```

```
df <- table(as.factor(minority_power$minority_stake ==0), minority_power$is_single_member, as.factor(mi
mantelhaen.test(df)
```

```
##
##  Mantel-Haenszel chi-squared test with continuity correction
##
## data:  df
## Mantel-Haenszel X-squared = 55.929, df = 1, p-value = 7.514e-14
## alternative hypothesis: true common odds ratio is not equal to 1
## 95 percent confidence interval:
##  0.03618292 0.15964897
## sample estimates:
## common odds ratio
##        0.07600372
```

```
minority_power_summarised <- minority_power %>% group_by(is_single_member) %>%
  summarise(percentage_zero = sum(minority_stake == 0)/length(minority_stake))
ggplot(minority_power_summarised, aes(x=is_single_member, y=percentage_zero*100)) + geom_bar(stat="iden
  xlab("District Type") + ylab("Percent of Elections Where All Seats Are Held By One Party") +
  ggtitle("One Party Controls All Seats More Often in States With Multi-Member Districts") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
        panel.background = element_blank(), axis.line = element_line(colour = "black"))
```



A Mantel-Haenszel chi-squared test demonstrates that the difference is significant, even when accounting for the total number of representatives elected from each state.

So the data provides evidence of Quitt's "sweep effect". But does it provide evidence that this effect does, in fact, hurt the majority party in the next election?

```r
differences <- filtered %>% arrange(congress) %>%
  group_by(state, party) %>%
  mutate(is_majority = actual > .5) %>%
  mutate(difference = c(diff(actual), NA), lost_majority=shift(actual, type="lead") < .5) %>%
  filter(is_majority == T)

differences$district_type <- ifelse(differences$mean_elected_district > 1.5, "multi-member", "single-mem
wilcox.test(difference ~ district_type, differences)
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  difference by district_type
## W = 11906, p-value = 0.5179
## alternative hypothesis: true location shift is not equal to 0
```

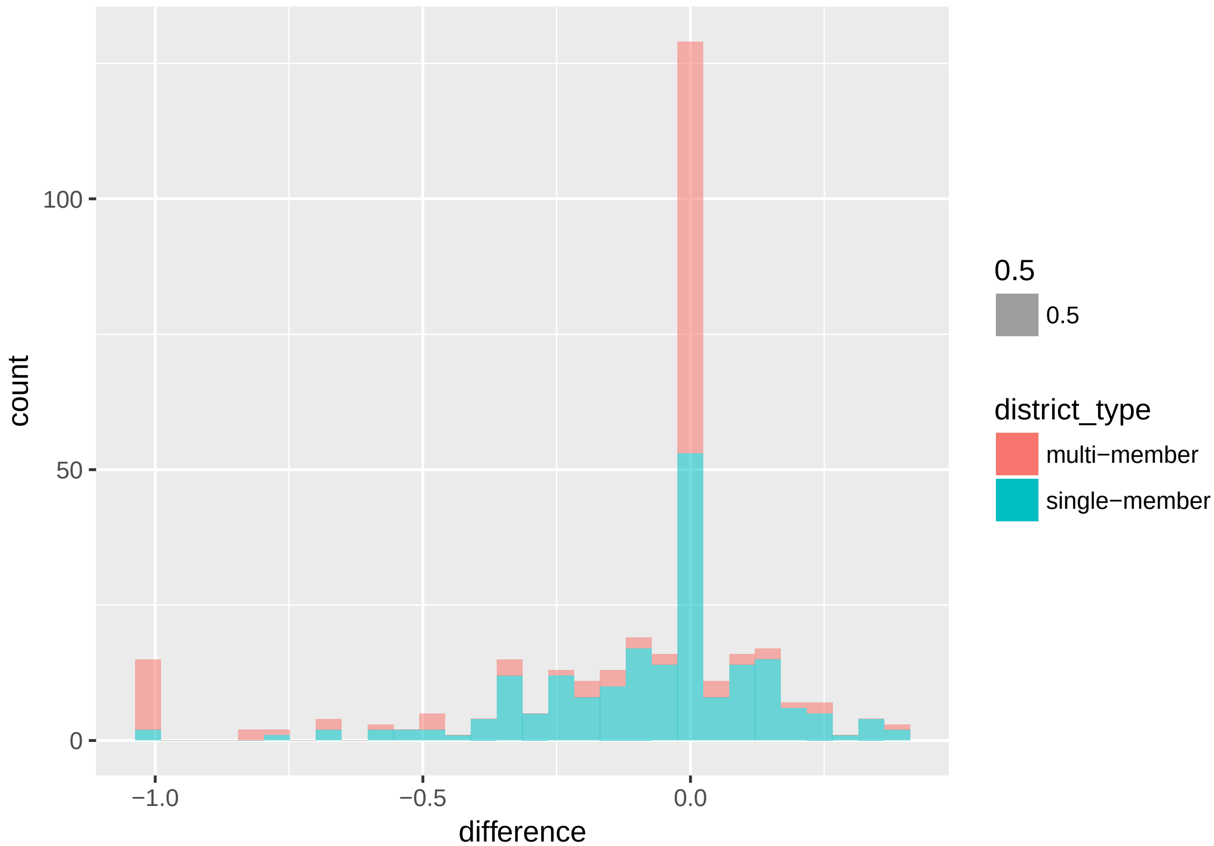
```r
wilcox.test(as.numeric(lost_majority) ~ district_type, differences)
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  as.numeric(lost_majority) by district_type
## W = 12352, p-value = 0.8941
## alternative hypothesis: true location shift is not equal to 0
```

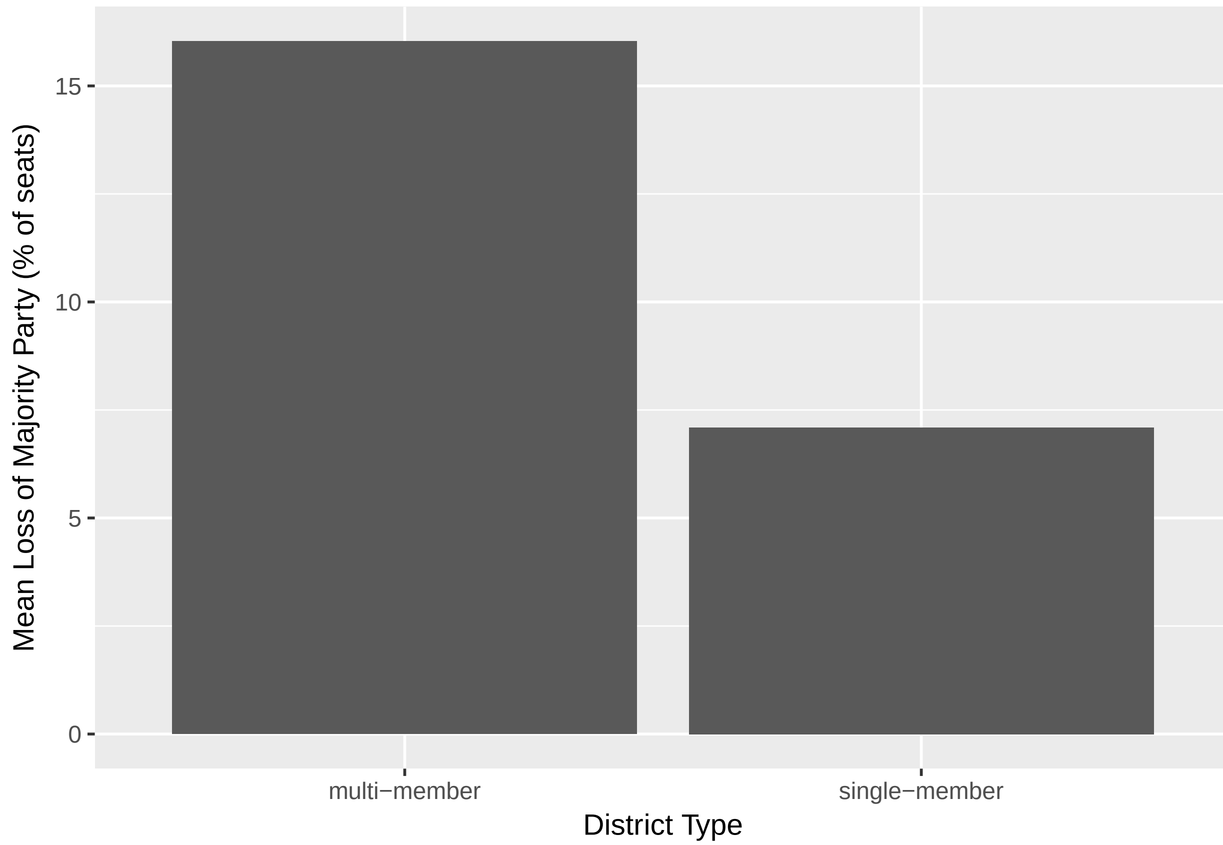
```r
ggplot(differences, aes(x=difference, alpha=0.5)) + geom_histogram(aes(fill=district_type))
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

```
## Warning: Removed 44 rows containing non-finite values (stat_bin).
```




```
differences_grouped <- differences %>% group_by(district_type) %>% summarise(diff=mean(difference, na.r
                                                                              lost_majority=mean(lost_ma
ggplot(differences_grouped, aes(x=district_type, y=-diff*100)) + geom_bar(stat="identity") + xlab("Dist
  ylab("Mean Loss of Majority Party (% of seats)")
```



From the data I have on hand, it does not look like it is the case that single-member districts tended to sweep out the majority party any more than multi-member districts.

Another hypothesis, raised by W. Mark Crain, posits that incumbents are disadvantaged in multi-member districts, and are reelected more often in single-member districts.

```
data <- data %>%
  filter(as.integer(congress) <= 27 & as.integer(congress) >= 2) %>%
  group_by(state, congress) %>%
  mutate(sum_votes = sum(votes * (type=="StandardElections"), na.rm=TRUE)) %>%
  group_by(state, congress) %>%
  mutate(num_districts = length(unique(district))) %>%
  group_by(state, congress, party) %>%
  mutate(theoretical = sum(votes * (type=="StandardElections"),na.rm=TRUE)/mean(sum_votes),
          actual=sum(result=="won")/max(num_elected_state),
          num_elected_state=max(num_elected_state), num_districts = max(num_districts)) %>%
  mutate(x= theoretical, theoretical = (theoretical-.5)*2+.5) %>%
  mutate(theoretical = ifelse(theoretical > 1, 1, ifelse(theoretical < 0, 0, theoretical)),
        mean_elected_district=num_elected_state/num_districts)

data$is_single_member <- ifelse(data$num_elected == 1, "Single-Member", "Multi-Member")

data_valid <- subset(data, incumbent == "True" & runoff=="False")
with(data_valid,
     summary(lm(percentage ~ num_elected + old_vote_share)))
```

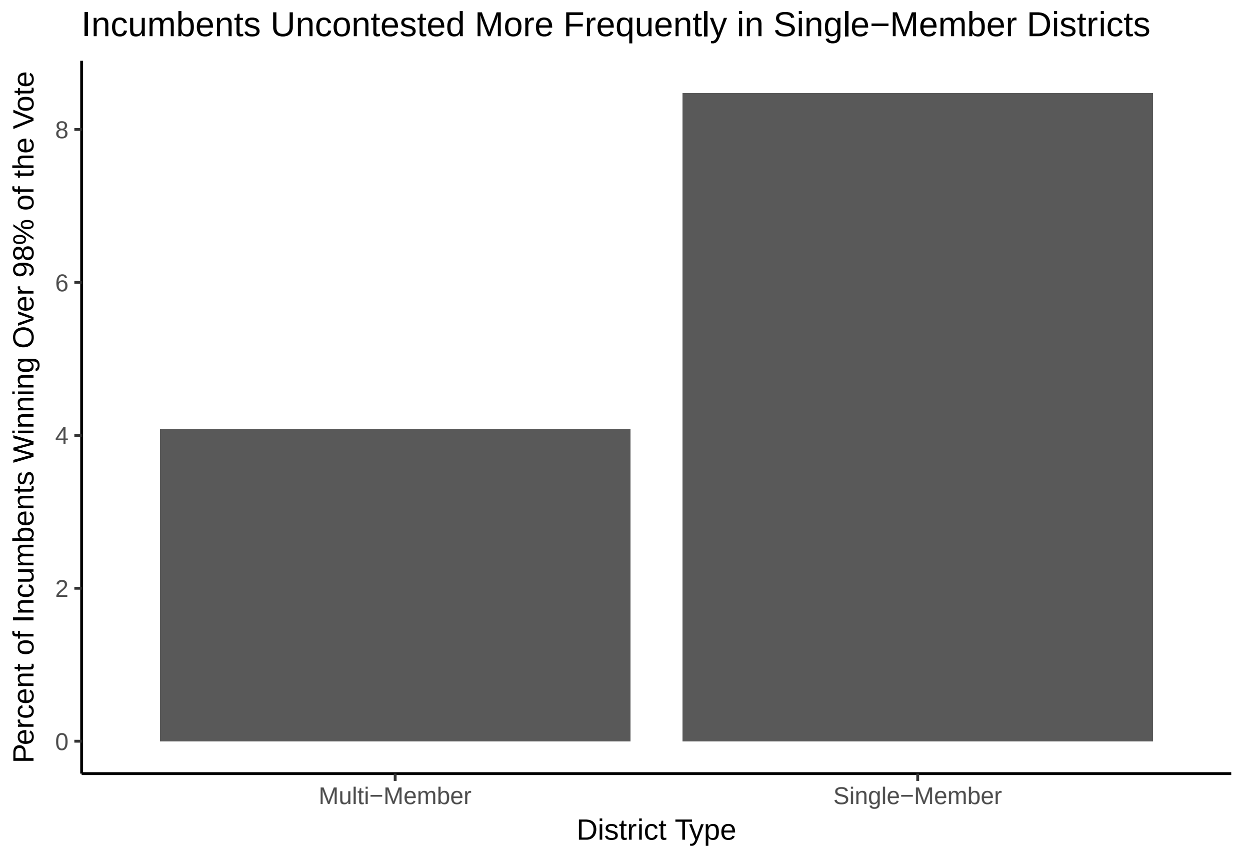
```
##
## Call:
## lm(formula = percentage ~ num_elected + old_vote_share)
```

```
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -70.207  -7.684  -2.215   7.419  48.161
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)    35.34545    2.07708  17.017   <2e-16 ***
## num_elected    -0.21896    0.25657  -0.853    0.394
## old_vote_share  0.39678    0.03205  12.382   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 17.77 on 1334 degrees of freedom
##   (783 observations deleted due to missingness)
## Multiple R-squared:  0.1034, Adjusted R-squared:  0.1021
## F-statistic: 76.92 on 2 and 1334 DF,  p-value: < 2.2e-16
```

```r
data_valid$uncontested <- data_valid$percentage > 98
chisq.test(as.factor(data_valid$uncontested), as.factor(data_valid$is_single_member))
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  as.factor(data_valid$uncontested) and as.factor(data_valid$is_single_member)
## X-squared = 7.678, df = 1, p-value = 0.00559
```

```r
uncontested_breakdown <- data_valid %>% group_by(is_single_member) %>% summarise(percent_uncontested = m

ggplot(uncontested_breakdown, aes(x=is_single_member, y=percent_uncontested*100)) + geom_bar(stat="iden
  xlab("District Type") + ylab("Percent of Incumbents Winning Over 98% of the Vote") +
  ggtitle("Incumbents Uncontested More Frequently in Single-Member Districts") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
        panel.background = element_blank(), axis.line = element_line(colour = "black"))
```



The number of representatives elected from an incumbent's district does not seem to have a significant effect on the candidate's eventual vote share; however, it makes a clear difference when looking at the percentage of incumbents running virtually uncontested (winning > 98% of the vote). Incumbents are far more likely to be uncontested from single-member districts, perhaps providing them with an incentive to establish single-member systems.

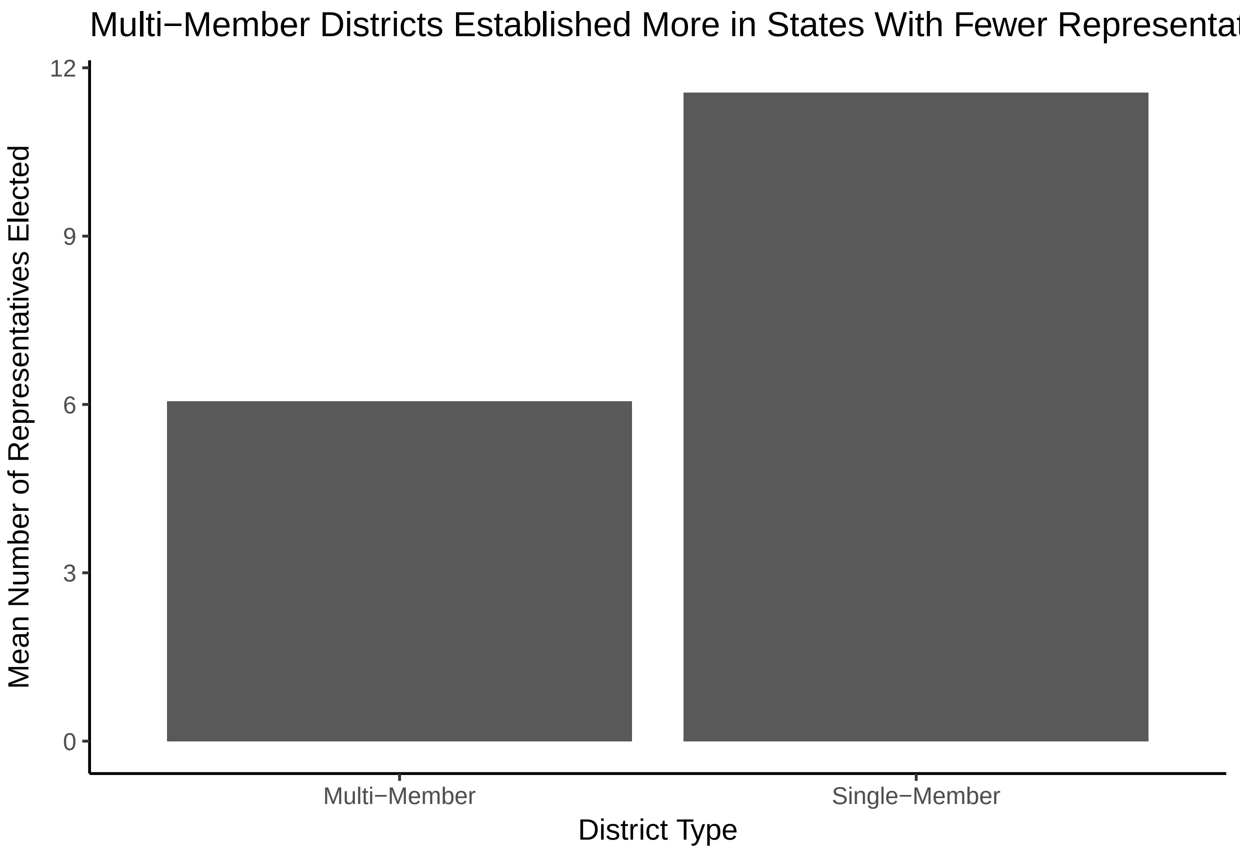
A common complaint of Whig lawmakers at the time of the Act was that small states using multi-member districts were able to send unified delegations consisting of one party to the house, while larger states with single-member districts were typically evenly split along party lines. This gave the smaller states disproportional power in the House. The data provides support for the idea that single-member districts usually were established from smaller states:

```
summary(lm(num_elected_state ~ as.factor(is_single_member), minority_power))
```

```
##
## Call:
## lm(formula = num_elected_state ~ as.factor(is_single_member),
##     data = minority_power)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -9.555 -4.055 -1.055  1.445 28.445
##
## Coefficients:
##                                        Estimate Std. Error t value
## (Intercept)                              6.0552     0.5689  10.644
## as.factor(is_single_member)Single-Member   5.5000     0.7086   7.762
##                                        Pr(>|t|)
```

```
## (Intercept)                              < 2e-16 ***
## as.factor(is_single_member)Single-Member 6.86e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.85 on 406 degrees of freedom
## Multiple R-squared:  0.1292, Adjusted R-squared:  0.1271
## F-statistic: 60.25 on 1 and 406 DF,  p-value: 6.855e-14
```

```
small_state <- minority_power %>% group_by(is_single_member) %>% summarise(mean_size=mean(num_elected_st
ggplot(small_state, aes(x=is_single_member, y=mean_size)) + geom_bar(stat="identity") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
        panel.background = element_blank(), axis.line = element_line(colour = "black")) +
  xlab("District Type") + ylab("Mean Number of Representatives Elected") +
  ggtitle("Multi-Member Districts Established More in States With Fewer Representatives")
```



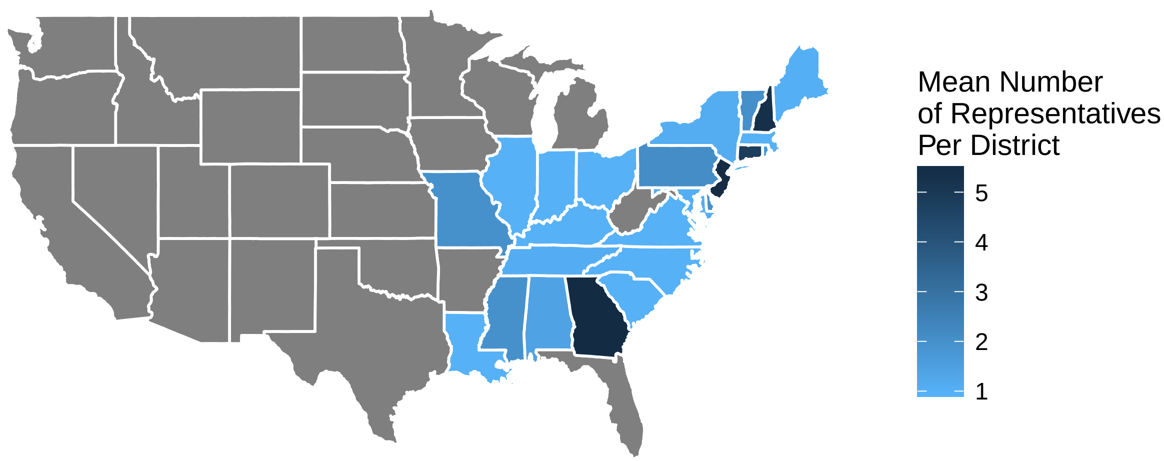
Single-member districts came from states with significantly fewer representatives.

However, it is possible that the relationship is confounded, and that the true influencer is not the number of representatives from a state but the geographic size of the state. Below is a map that may document this:

```
state_aggregate <- minority_power %>% group_by(state) %>% summarise(mean_num_elected_district = mean(me
state_aggregate$region <- tolower(state_aggregate$state)
states_with_data <- left_join(states, state_aggregate, by = "region")
ggplot(data = states_with_data) +
  geom_polygon(aes(x = long, y = lat, group = group, fill=mean_num_elected_district), color = "white") +
  coord_fixed(1.3) + scale_fill_gradient(low="#56B1F7", high="#132B43") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
        panel.background = element_blank(), axis.line = element_line(colour = "black"),
```

```
         axis.title=element_blank(),
         axis.text=element_blank(),
         axis.ticks=element_blank(),
         axis.line.x=element_blank(),
         axis.line.y=element_blank()) +
  labs(fill="Mean Number\nof Representatives\nPer District") +
  ggtitle("Distribution of Single- and Multi-Member Districts, 1788-1841")
```

Distribution of Single− and Multi−Member Districts, 1788−1841



To thoroughly test this hypothesis, I set up a logistic regression including both the number of representatives and the size of the state:

```
state_area <- read.csv("state_area.csv")
state_area$state <- state_area$State
state_area$Total.Area <- as.numeric(gsub(",", "", state_area$Total.Area))
states_with_area <- left_join(minority_power, state_area, by="state")
```

```
## Warning: Column `state` joining factors with different levels, coercing to
## character vector
```

```
wilcox.test(Total.Area ~ is_single_member, states_with_area)
```

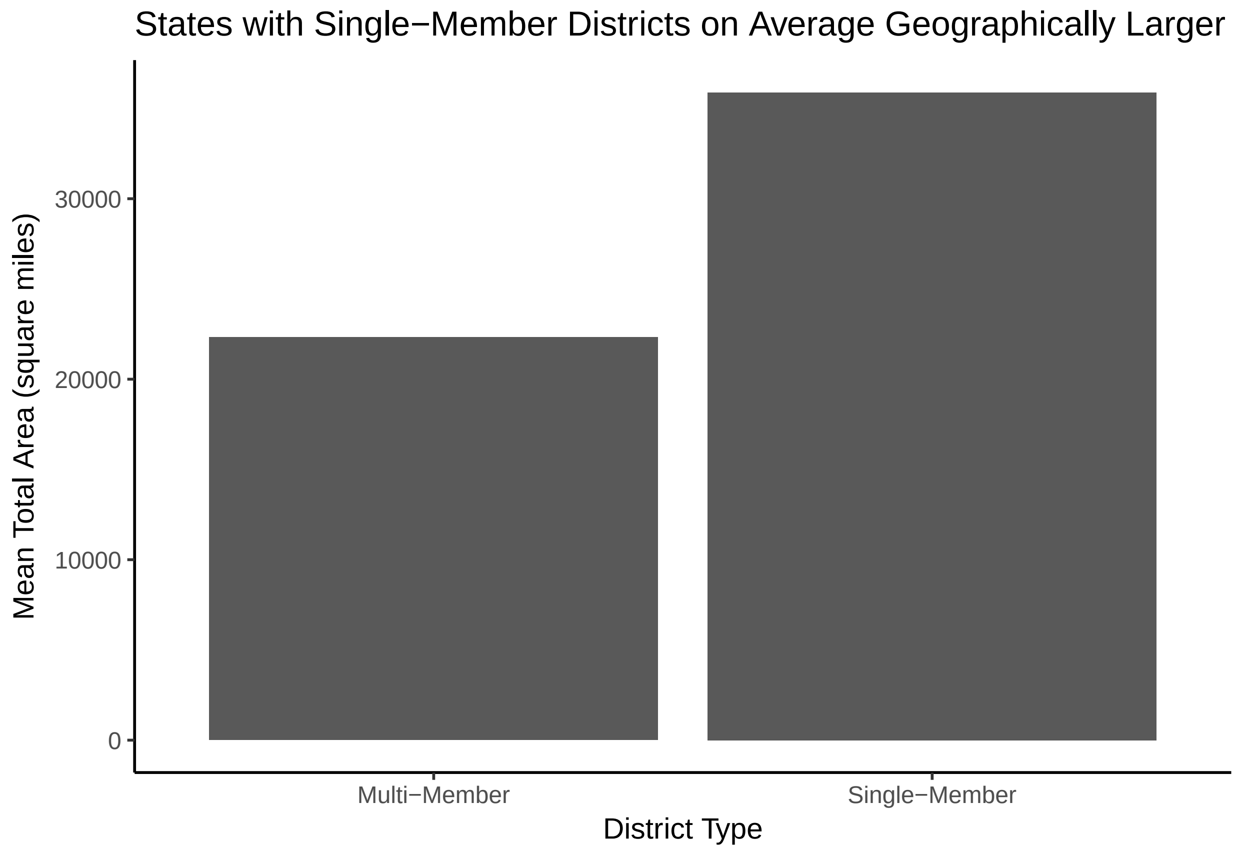
```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  Total.Area by is_single_member
## W = 11452, p-value = 2.253e-11
## alternative hypothesis: true location shift is not equal to 0
```

```
summary(glm(as.factor(is_single_member) ~ log(Total.Area) + log(num_elected_state),
            family = binomial(link = "logit"), data = states_with_area))
```

```
##
## Call:
## glm(formula = as.factor(is_single_member) ~ log(Total.Area) +
##     log(num_elected_state), family = binomial(link = "logit"),
##     data = states_with_area)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
```

```
## -2.3736  -1.0272   0.4787   0.7821   2.3746
##
## Coefficients:
##                          Estimate Std. Error z value Pr(>|z|)
## (Intercept)               -9.2966     1.3249  -7.017 2.27e-12 ***
## log(Total.Area)            0.7858     0.1307   6.012 1.83e-09 ***
## log(num_elected_state)     1.1095     0.1967   5.639 1.71e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 530.99  on 407  degrees of freedom
## Residual deviance: 414.66  on 405  degrees of freedom
## AIC: 420.66
##
## Number of Fisher Scoring iterations: 4
```

```
area_aggregate <- states_with_area %>% group_by(is_single_member) %>% summarise(mean_area = mean(Total.A

ggplot(area_aggregate, aes(x=is_single_member, y=mean_area)) + geom_bar(stat="identity") +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),
        panel.background = element_blank(), axis.line = element_line(colour = "black")) + xlab("District
  ylab("Mean Total Area (square miles)") +ggtitle("States with Single-Member Districts on Average Geogra
```



It appears that both variables have a significant effect on whether a state uses a multi-member district.

Lastly, I will go to the most likely source of information of representatives' true motivations: their vote record.

The final vote on the single-member districting mandate was 101-99, but unfortunately I can only account for 95 yes votes and 88 no votes in my electoral data. Still, it is possible to set up a model.

I attempt to take all investigated variables into account for this analysis:

num_elected_state is the number of representatives elected from the representative's state.

is_whigTRUE represents whether the representative was a Whig; Whigs generally supported the mandate.

inferred_percentage is the percentage of the vote the representative won in the last election.

num_elected represents the number of representatives elected from the given representative's district.

Since I am testing 4 variables, I will use a Bonferroni correction to set alpha to 0.0125.

```r
vote_record <- read.csv("vote_record.csv")

voters_raw <- data %>% filter(congress == 27)
voters <- left_join(voters_raw, vote_record, by = "clean_name")
```

```
## Warning: Column `clean_name` joining factors with different levels,
## coercing to character vector
```

```r
table(voters$vote)
```

```
##
##  no yes
##  88  95
```

```r
voters$is_whig <- as.factor(voters$party == "W")
summary(glm(as.factor(vote) ~ num_elected_state + is_whig + inferred_percentage + num_elected,
            family = binomial(link = "logit"),
            data = voters))
```

```
##
## Call:
## glm(formula = as.factor(vote) ~ num_elected_state + is_whig +
##     inferred_percentage + num_elected, family = binomial(link = "logit"),
##     data = voters)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.3598  -0.3071   0.3814   0.4749   2.4775
##
## Coefficients:
##                      Estimate Std. Error z value Pr(>|z|)
## (Intercept)          0.240171   1.413369   0.170  0.86507
## num_elected_state   -0.009056   0.023858  -0.380  0.70426
## is_whigTRUE          5.191429   0.720184   7.208 5.66e-13 ***
## inferred_percentage -0.046646   0.020518  -2.273  0.02300 *
## num_elected         -0.448016   0.137486  -3.259  0.00112 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 241.01  on 173  degrees of freedom
## Residual deviance: 102.15  on 169  degrees of freedom
##   (376 observations deleted due to missingness)
```

```
## AIC: 112.15
##
## Number of Fisher Scoring iterations: 6
```

The results suggest some interesting conclusions:

1. The total number of representatives elected from the representative's state does not have a significant effect on a representative's vote. Representatives from larger states were not significantly more likely to vote for the mandate.
2. Party is the strongest and most significant factor in the vote. Whigs were significantly more likely to vote for the amendment.
3. The percentage of the vote the representative won in the last election does not significantly effect the vote; vulnerable representatives do not appear to be significantly more likely to vote for the amendment.
4. The number of representatives elected from the representative's district does have a significant effect on the vote. Representatives from districts with multiple members were more likely to oppose the amendment.