

The Effect of Caucuses on US Primary Elections

IDS 690: Unifying Data Science II

Anna Berman, Kate Coulter, Shota Takeshima

Background

Motivation

In the US presidential voting system, a preliminary round of voting takes place in which voters select a party nominee. The winner of this first election then goes on to run in the general election against the nominees of the other parties. This first round of voting can take place in two very different formats, a primary or a caucus. In primary elections, voters go to polls and cast ballots for their preferred candidate for their party in the general election. In caucuses, on the other hand, voters gather locally to openly debate and decide which candidate to support.

As Caitlin Jewitt put it during an NPR interview, “proponents of caucuses often say that it is deliberative democracy at its finest. It is the chance for people to show up and talk to their neighbors about politics, have informed conversation, talk to representatives from the campaigns and be persuaded” (Cornish 2020). Because caucuses are a more participatory form of elections, proponents of caucuses often believe they make voters feel that they have a voice and that leads to increased voter turnout. Conversely, proponents of primaries retain that caucuses are time consuming and thus less available to all voters. If caucus voters are required to spend hours talking about the candidates and standing in a gymnasium waiting for all people in the room to be counted, perhaps hiring a babysitter to stay home with their kids, that may very well be enough to prevent many from attending.

The following analysis examines the effect of election format on voter turnout. Specifically, do primaries result in higher voter turnout compared to caucus elections?

Research Design

National Voter Turnout Analysis

Without the capability to randomly assign election formats to each US state, we turn to voter records from presidential candidate primary and caucus elections as well as records of which format of election each state utilized. In 2020, several states switched from caucuses to primary presidential nominee elections including Colorado, Maine, Washington, Minnesota, Kansas, Idaho, Hawaii, and Nebraska. Our design will group states that made the switch from caucuses to primaries in the 2020 election and those that did not, but that had similar changes in voter turnout between 2012 and 2016. We will then do a difference in difference analysis between 2016 and 2020 democratic elections to assess the effect of election format on voter turnout.

Colorado Case Study

Building upon this initial analysis, we looked in-depth at one state that made the switch from caucus to primary format in 2020 - Colorado. Specifically, we examined the change in voter turnout in demographic groups, such as African Americans and Millennials, using a pre-post approach. Unfortunately, we were not able to locate individual-level voter records. However, by combining county-level demographic data with county-level turnout data, we can infer differences in effect on demographic groups based on the differences in demographics of each county.

Data

Data Sources

National Voter Turnout Analysis

Prior research around US primary elections suggests that the United States Elections Project is a gold-standard for voter turnout numbers (Rakich, 2020). Therefore, voter turnout and election format records were collected from the United States Elections Project at electprojec.org. This data is freely available and was retrieved March 22, 2020 (United States Election Project 2020). Elections included in this analysis include 2008, 2016 and 2020 US primaries. The 2012 primary election was excluded from this analysis as this was an Democratic incumbent year and many voter records were resultantly missing.

States included in the treatment group are Maine, Minnesota, Washington and Colorado as each of these states switched from caucus to primary elections in 2020. States included in the control group are states that had a consistent election format across all three election cycles. States with unique voting systems such as California were excluded from this analysis. States with presidential nominee elections that were affected by COVID-19 were also excluded from analysis. Finally, states without complete democratic voter turnout data for 2008, 2016, and 2020 elections were excluded from analysis. A complete list of states included in this analysis can be found in Appendix A.

We supplemented voter turnout data from the United States Elections Project with state-level demographic data collected by Kaiser Family Foundation. This supplemental dataset included population density and racial demographic features.

Colorado Case Study

In the second phase of analysis, county-level 2016 and 2020 presidential primary voter turnout was collected from the office of Colorado's Secretary of State (2020). County-level demographic data was collected from Colorado's State Demography Office (2020). Demographic data included population density, age, race, and ethnicity features.

Summary Statistics

National Voter Turnout Analysis

Population & Race

2020 population density is relatively similar across both treatment and control groups with mean populations of 4.92 and 6.75 million respectively (and median populations of 4.91 and 5.52 million for control and treatment groups respectively). These statistics are consistent across 2008-2020. Although population density is consistent across groups, racial

demographics diverge. Treatment groups have larger proportions of white citizens and a much lower proportions of black citizens compared to control groups. This divergence will be controlled for in the difference in difference analysis. The county pre-post analysis will further examine the effect of election format for specific demographic groups.

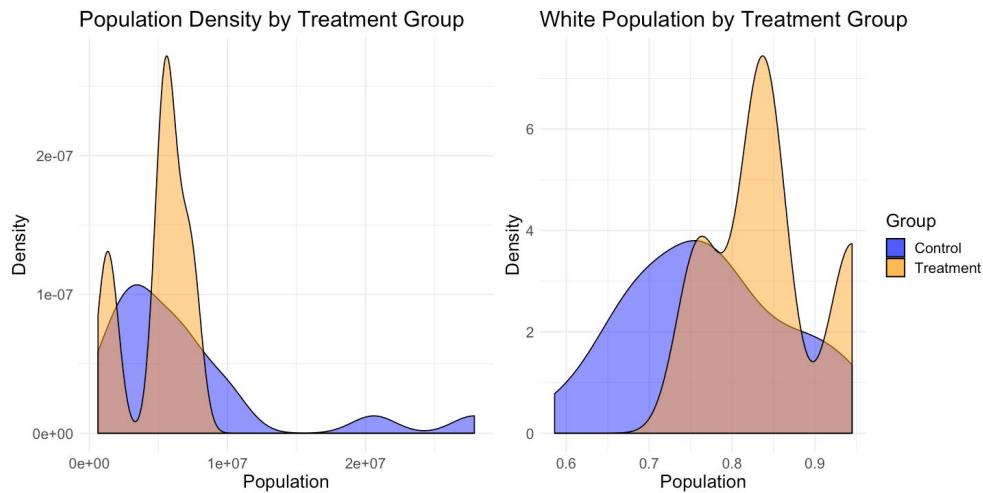


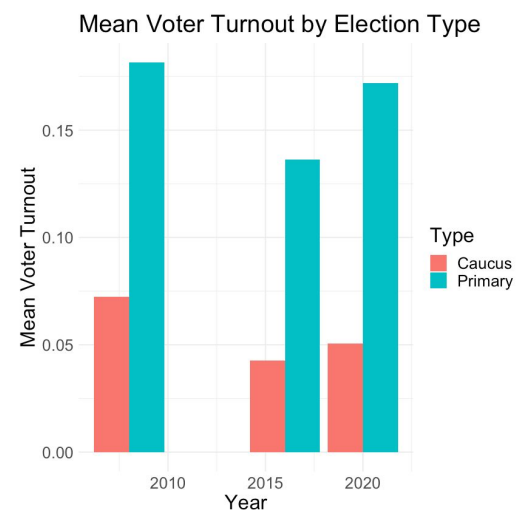
Fig 1. Kernel density plots for 2020 total population density and white population proportion in treatment and control groups. Population density is relatively similar across treatment and control groups while treatment groups show a higher proportion of white citizens compared to control groups.

Voter Turnout

On a National level, the mean voter turnout for the Democratic Primaries for 2008, 2016 and 2020 were 15.0%, 10.9% and 15.7% respectively. Each year, states with primaries observe greater voter turnouts with rates 2.5 - 3.4 times that of states with caucuses. This descriptive analysis does not imply causality, but does show evidence for a relationship between election format and voter turnout.

Year	Voter Turnout		Election Format	Voter Turnout	
	Mean	Median		Mean	Median
2008	15%	15.4%	Caucus	7.2%	5.7%
			Primary	18.2%	16.5%
2016	10.9%	10.3%	Caucus	4.3%	4.4%
			Primary	13.6%	11.8%
2020	15.7%	13.2%	Caucus	5%	4.9%
			Primary	17.2%	14.4%

Fig 2. Average voter turnout by election format. Primary election formats are associated with much higher voter turnout compared to caucus formats.



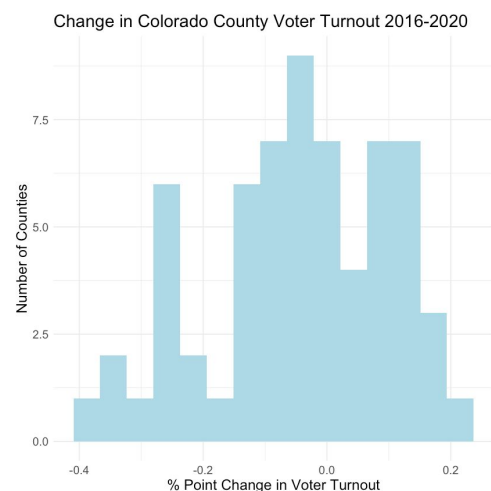
Colorado Case Study

Across Colorado's 64 counties, Colorado has a considerably high percentage of White citizens with the average county having 94% White citizens compared to the US average of 78%. Moreover, we see that counties in Colorado have a high percent of 18-25 year olds, with the average of 12% compared to the US average of 9% (United States Census 2010). We suspect that the change from caucuses to primaries would most affect minority groups such as non-Whites or Millennials. Thus, Colorado may not be the ideal state to study as our results may be smaller than anticipated for some demographic groups. However, given the limited data available, we maintain that Colorado is the best available option and kept these limitations in mind as we continued.

It should also be noted, that when not weighting statistics by county population density the average change in voter turnout in Colorado becomes -4%, compared to +7% when weighted by population. This is due to several small Colorado counties that saw large decreases in voter turnout between the 2016 and 2020 elections.

Statistic	Turnout - 2020	Turnout - 2016	% White	% 18-25	Change in Turnout
Mean	17.3%	21.7%	93.8%	11.9%	-4.3%
Median	16.4%	18.7%	95.3%	11.4%	-3.5%
Min	4.9%	7.2%	79.9%	6.5%	-32.7%
Max	37.4%	56.8%	98.6%	20.7%	23.2%

Fig 3. Summary statistics for counties in Colorado. Statistics not weighted by population. Colorado counties have a strong White majority and a high percentage of young residents. Average change in voter turnout is negative when population is not taken into account.



Results

National Voter Turnout Analysis

Our results show parallel trends between 2008 and 2016 Democratic primary elections. Between 2008 and 2016 control states saw a 4.3 percentage point drop in voter turnout while treatment states saw a 2.9 percentage point drop. On the other hand, between 2016 and 2020 we can observe stark differences in voter turnout rates. In control states there was a 2.2 percentage point increase in voter turnout while in treatment states there was a 17.9 percentage point increase. Using a difference in difference approach, the estimate for the effect of primary election format is 15.7 percentage points.

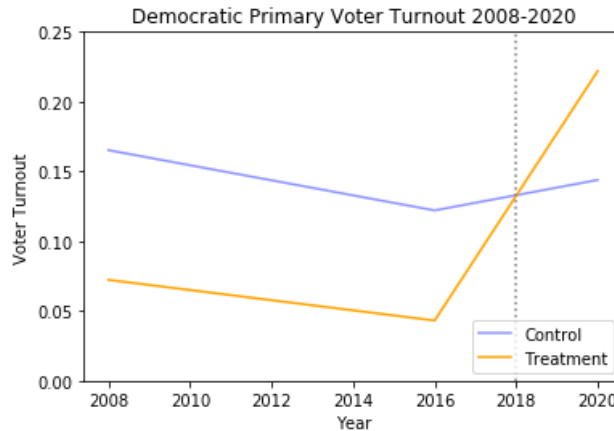


Fig 4. Democratic primary voter turnout between 2008 and 2020. Assumptions of parallel trends is met for elections between 2008 and 2020. Voter turnout greatly increased with the switch from caucus to primary election format. The causal effect of primary election format on voter turnout is 15.7 percentage points.

Colorado Case Study

Our results do not suggest any significant effect on the switch in format for any racial group, which is largely anticipated given our data and the small number of non-White population in any Colorado county.

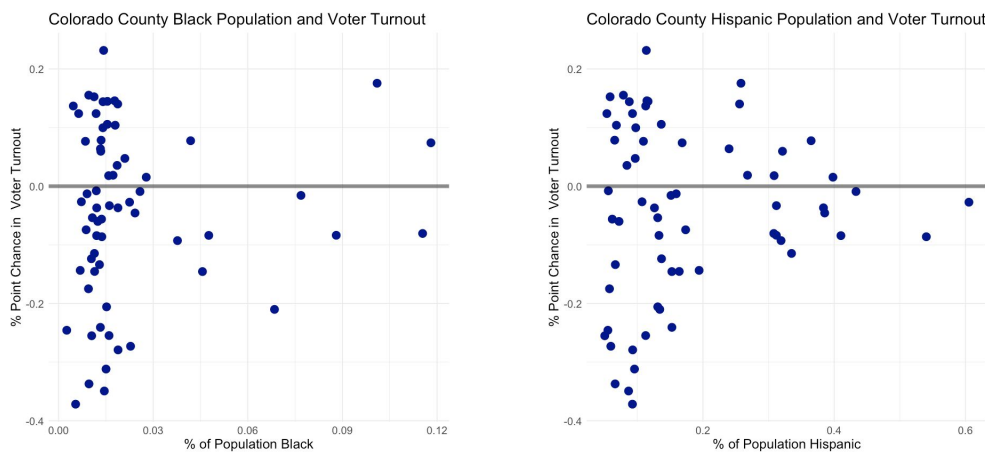


Fig 5. Change in voter turnout by demographic groups for Colorado counties. Little effect is observed between the proportion of racial minorities and changes in voter turnout.

However, results do suggest a relationship between increased voter turnout and proportion of county residents between the ages of 18 and 25. A simple linear regression suggests that each percentage point increase in millennial residents is associated with a 1.6% point increase in voter turnout (regression summary statistics can be found in Appendix B). This change may have additional impacts, as those that are voted for by younger groups are thought to be different than those who are older in certain cases (Milligan 2020). Moreover, the opposite effect is observed for proportion of senior residents with larger proportions being associated with decreased voter turnout.

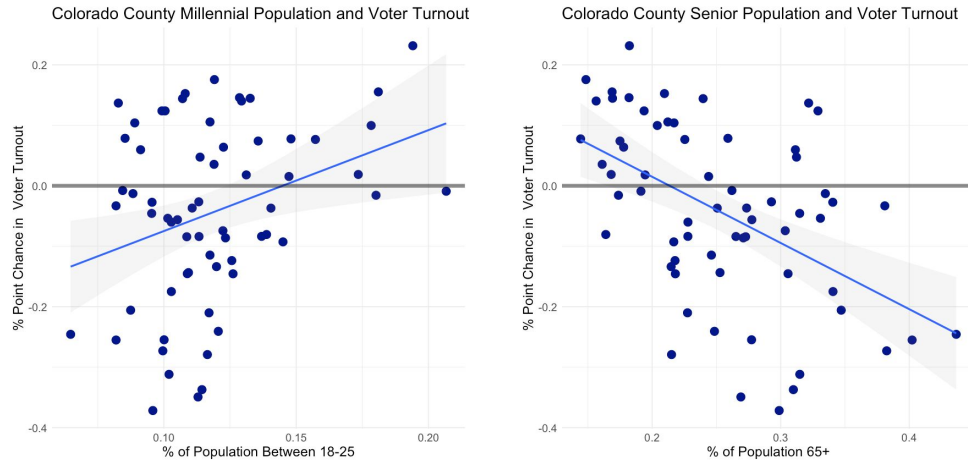


Fig 6. Change in voter turnout by age for Colorado counties. Increased young millennial populations are associated with increased voter turnout with opposite effects observed for senior citizens.

Discussion

Our findings suggest a considerable effect of election format on voter turnout. A 15.7 percentage point increase is enormous considering that the mean voter turnout for democratic caucus elections for 2008, 2016, and 2020 is only 5.6%. While voting regulations continue to be a state-regulated feature of presidential elections, the shift to having fewer caucuses and more typical primaries has continued throughout time and should continue to be monitored as voter turnout is an important feature of a democracy.

While our Colorado case study does not provide causal evidence for the effect of election format on specific demographic groups, it does support an association between election format and increased voter turnout for certain demographic groups. Specifically, increases in voter turnout were observed most often in counties with large millennial populations. As previously mentioned, an increased millennial voter turnout has the potential to impact the outcome of many elections as millennial voters tend to vote differently than older generations. Specifically, millennials tend to skew more Democratic and seem to care about social issues such as minority rights and climate change (Byler, 2019). If switching from a caucus to a primary system does, in fact, increase millennial voter turnout, election results may favor more Democratic candidates.

Future work should analyze the true impact that voting format has on minority groups to identify the true effect on turnout by demographic. This research could be done using individual voter roll and census demographic data which should become available in the coming year.

References

- Byler, M. (2019). Millennials could push American politics to the left — or totally upend them. Washington Post. Retrieved from <https://www.washingtonpost.com/opinions/2019/05/22/millennials-could-push-american-politics-left-or-totally-upend-them/>
- Colorado Secretary of State (2020). Election Results Archive. Retrieved from <https://www.sos.state.co.us/pubs/elections/Results/Archives.html>.
- Colorado State Demography Office (2020). Race and Ethnicity by Age Estimates - County. Retrieved from <https://demography.dola.colorado.gov/population/data/race-estimate/>.
- Cornish, A. (Host). (2020, February 5) Caucuses Or Primaries? Why States Might Pick One Or The Other. [Radio broadcast episode]. <https://www.npr.org/2020/02/05/803183343/caucuses-or-primaries-why-states-might-pick-one-or-the-other>
- Kaiser Family Foundation. Population Distribution by Race/Ethnicity. Retrieved from <https://www.kff.org/other/state-indicator/distribution-by-raceethnicity/?dataView=1¤tTimeframe=10&sortModel=%7B%22colId%22:%22Location%22,%22sort%22:%22asc%22%7D>
- Milligan, S. (2020, March 11). Young Voters Love Bernie, Just Not Enough. US News. Retrieved from <https://www.usnews.com/news/elections/articles/2020-03-11/young-voters-love-bernie-sanders-but-older-voters-gave-joe-biden-the-win>
- Rakich, N. (2020, March 12). Historic Turnout In 2020? Not so Far. Retrieved from <https://fivethirtyeight.com/features/historic-turnout-in-2020-not-so-far/>.
- United States Census (2010). Age and Sex Composition 2010. Retrieved from <https://www.census.gov/prod/cen2010/briefs/c2010br-03.pdf>.
- United States Elections Project. (Retrieved 2020, March 22). Voter Turnout Data. <http://www.electproject.org/home/voter-turnout/voter-turnout-data>

Appendix A

States included in the treatment group are states switched from caucus to primary elections in 2020. States included in the control group are states that had a consistent election format across all three elections. States with unique voting systems such as California were excluded from this analysis. States with presidential nominee elections that were affected by COVID-19 were also excluded from analysis. Finally, states without complete democratic voter turnout data for 2008, 2016, and 2020 elections were excluded from analysis.

Treatment group: Colorado, Washington, Minnesota, & Maine

Control Group: Alabama, Arizona, Arkansas, Florida, Iowa, Massachusetts, Michigan, Mississippi, Missouri, Nevada, New Hampshire, North Carolina, North Dakota, Oklahoma, South Carolina, Tennessee, Texas, Utah, Vermont, & Virginia

Appendix B

A simple linear regression predicting change in voter turnout based on the proportion of county millennial residents, Black residents, and Hispanic residents found that proportion of millennial residents was the most predictive feature.

Call:

```
lm(formula = delta ~ pop18.25 + Black + Hispanic, data = colorado)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.29296	-0.09232	0.01241	0.09833	0.24108

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-0.24493	0.07796	-3.142	0.00261	**
pop18.25	1.64859	0.66501	2.479	0.01600	*
Black	-0.07232	0.75296	-0.096	0.92380	
Hispanic	0.04047	0.14448	0.280	0.78037	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1432 on 60 degrees of freedom

Multiple R-squared: 0.1055, Adjusted R-squared: 0.06079

F-statistic: 2.359 on 3 and 60 DF, p-value: 0.08049