Preliminary Primary Election Forensics: Governor and U.S. Senate, Orange County, June 5, 2018

R. Michael Alvarez and Yimeng Li

June 23, 2018

Executive Summary

In this preliminary report, we examine data from two statewide candidate races (the gubernatorial and U.S. Senate elections) in Orange County. Our analysis uses two different visualizations of preliminary precinct-level data from the Statement of Votes: histograms of candidate vote shares and scatter-plots of candidate vote shares by turnout. We find no evidence in these preliminary data of significant outliers or unusual distributions.

Introduction

As part of our work developing elections forensics methods for the recent June 5, 2018 primary in Orange County, we have begun an analysis of data from the County's preliminary Statement of Votes (SOV) — these are data on precinct-by-precinct votes cast for each candidate and ballot measure on the primary election ballot. In this initial study, we focus on two of the more salient statewide candidate races: the gubernatorial and U.S. Senate elections. Subsequent analyses will examine other statewide and down-ballot races on the Orange County primary election ballot.

Our initial analysis uses visualizations of the SOV data, taken from the June 15, 2018 report issued by the Orange County Registrar of Voters (OCRV). We downloaded the pdf of the SOV from the OCRV website, extracted from the pdf data on total ballots cast by precinct, as well as the ballots cast for all Gubernatorial and U.S. Senate candidates. We subset the data, removing all-mail precincts (they typically have very small numbers of registered voters, which tends to skew the visualizations). We then compute the vote shares (the percentage of votes currently tabulated for each candidate, of total ballots cast in that precinct). These data are preliminary, and may not reflect the final vote totals once the OCRV certifies the primary election.

There are two visualizations that we use here to look for potential anomalies in the precinct-level

data in these two races. The first approach is to examine histograms of candidate vote shares across precincts. We expect to see that the distributions of candidate vote shares (for the candidates receiving appreciable numbers of votes in the Orange County primary election) will have a relatively smooth and single-peaked distribution. Outliers, or multiple peaks, can be indications of precincts that might require further research.

The second visualization that we will present are candidate vote shares by precinct turnout. In general, we expect to see joint distributions between a candidate's vote share and precinct turnout that do not have any extreme outliers, and that the joint distributions do not appear to have multiple clusters of points in the scatter-plots we produce. In scatter-plots like these, anomalies that might call for further research would manifest themselves as extreme outliers, or situations where we see two clusters of points in the scatter-plot. For example, an unusual distribution might be a situation where a proportion of a candidate's vote shares come from moderate-turnout precincts, but the remaining proportion come from precincts with very high turnout and high vote shares for that candidate. In terms of our visualizations, this would manifest in highly irregular joint distributions: we would see clear evidence of multiple cluster of vote shares and turnout for a candidate in the scatter-plots.

All of these visualizations are attached to the this report. We will update this report as new data becomes available.

Preliminary Precinct Returns for Governor's Race

The first five histograms show the candidate vote shares, across the election precincts, for the candidates receiving most of the votes in Orange County: Cox (Figure 1), Allen (Figure 2), Newsom (Figure 3), Chiang (Figure 4), and Villaraigosa (Figure 5). Each histogram gives the distribution of candidate vote shares; the red dotted line shows the central tendency of the distribution.

The two Republican candidates receiving the most votes in the Orange County statewide primary are John Cox (Figure 1) and Travis Allen (Figure 2). For these two candidates, their vote share distributions are smoothly shaped, single-peaked, and they display no outliers. The three Democratic candidates receiving the larger vote shares in Orange County's primary, Newsome (Figure 3), Chiang (Figure 4), and Villaraigosa (Figure 5, each have histograms of vote share that are single peaked, though they each has a slight rightward skew, which is most noticeable for Villaraigosa. This indicates that these candidates likely have moderate levels of support in a certain set of precincts in Orange County, perhaps due to specific electoral segments supporting each candidate.

In Figure 6 we provide a single visualization of the joint distribution of each of these same gubernatorial candidate's vote share and precinct turnout. In this plot, we do not see any evidence for multiple clusters; rather we see that each of the joint distributions of vote share and turnout generally are grouped into single clusters, and there are relatively few outliers.

In Figure 6 we also plot a linear fit for each of these joint distributions (the shaded area around the

linear fit shows the uncertainty of the linear fit). We see evidence that there is a correlation between candidate vote share and turnout for Cox (positive) and Villaraigosa (negative), perhaps a reflection of each candidate's support in different precincts in Orange County. This will be the subject of further study; we see the same type of pattern in the U.S. Senate analysis below, indicating that this is a general feature of political behavior in Orange County.

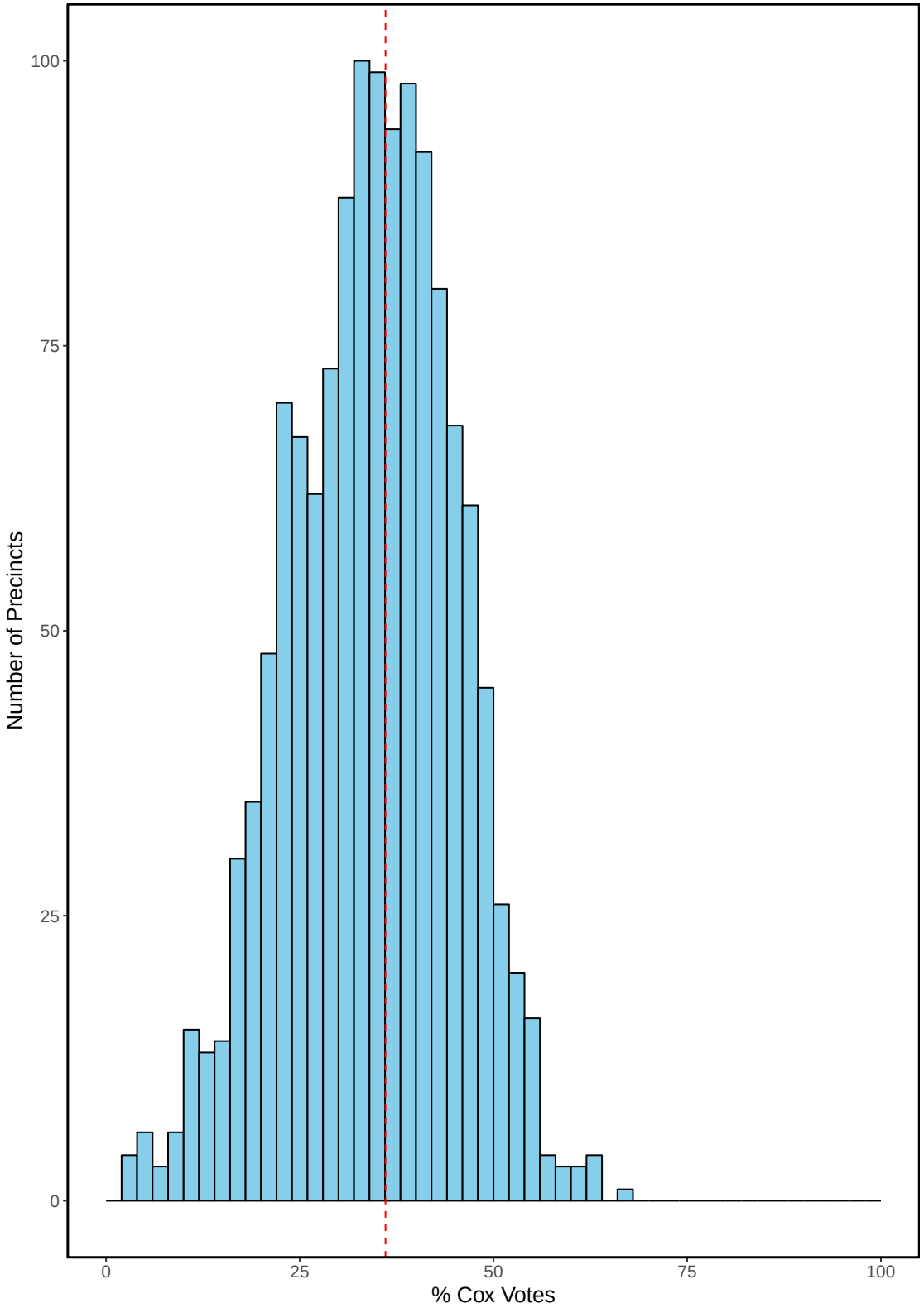
Preliminary Precinct Returns for U.S. Senate Race

Next, we present candidate vote share distributions for the three candidates receiving much of the vote in Orange County: Dianne Feinstein (Figure 7), Kevin de Leon (Figure 8), and James Bradley (Figure 9). The distribution of Feinstein's vote share is single-peaked, and appears very similar to a "normal" distribution, without any outliers. Bradley and de Leon both received much smaller vote shares across Orange County precincts, so their vote share distributions are much further to the left than Feinstein's. In the visualizations for Bradley and de Leon, their distribution of vote shares are single peaked, with a slight skew to the right; neither have any significant outliers.

In Figure 10 we provide the joint distributions of candidate vote shares and turnout. Again, here we are looking for visual evidence of odd joint distributions — which we do not see. There is no evidence of significant outliers in this scatter-plot, nor do we see evidence of multiple clusters of points for any candidate. And again, the linear fit shows a positive correlation between candidate vote share and turnout for Bradley, the Republican candidate, but a negative correlation for the two Democratic candidates in the election (Feinstein and de Leon). As we noted above, this general pattern requires further study, as it appears in both of these two salient statewide elections it is likely to be a reflection of political behavior in Orange County, and not an indication of any administrative issue requiring additional investigation.

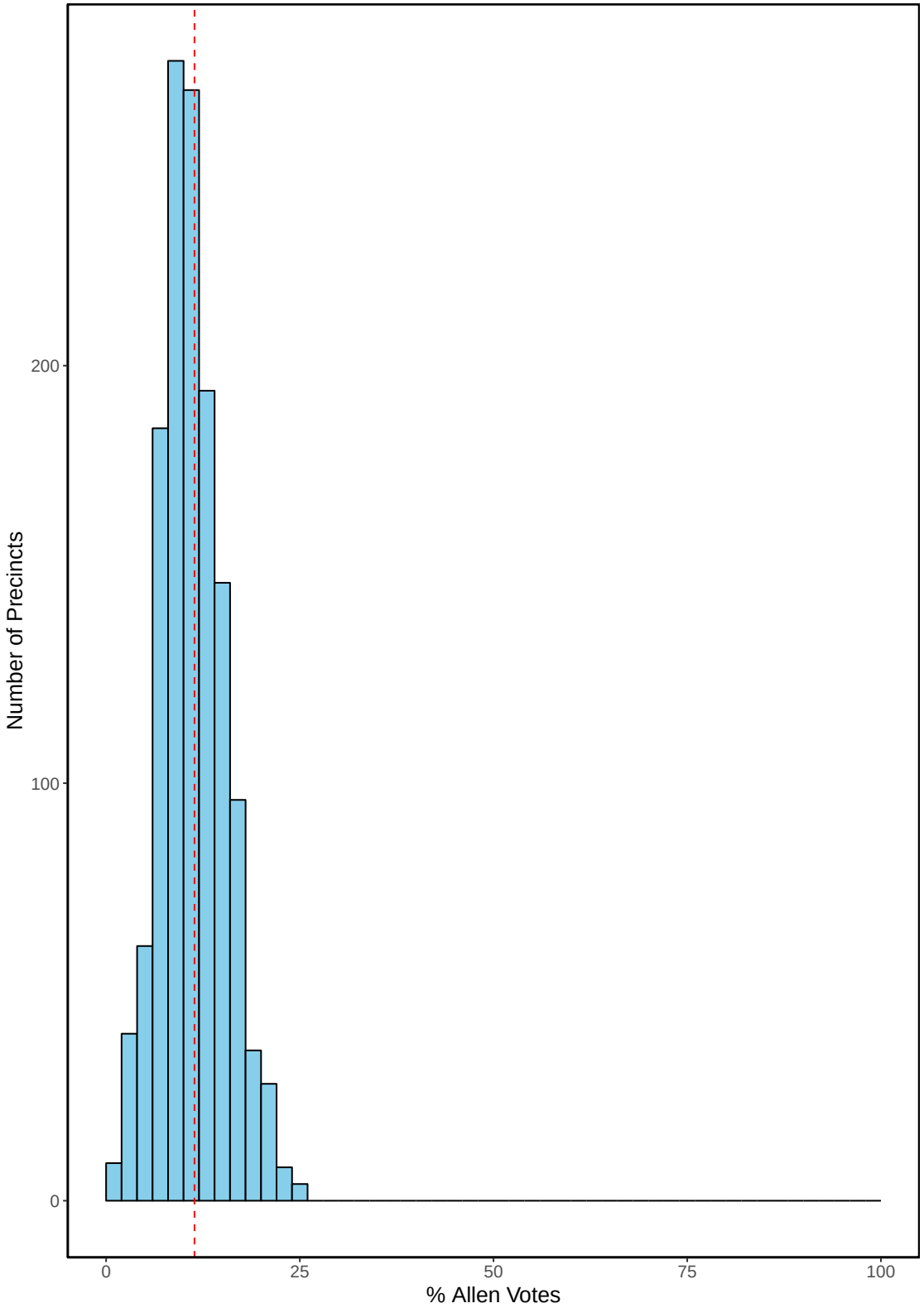
Conclusion

In this preliminary report we have provided a visual forensics analysis for candidates receiving larger vote shares in the gubernatorial and U.S. Senate elections in Orange County. In the candidate voter share histograms and in the scatter-plots of vote shares by turnout we see little evidence of anomalies needing further research. We do find that there is a correlation in both races between the vote shares and turnout for some candidates, which is consistent across the two statewide races we study in this preliminary report which we will continue to examine in future studies.



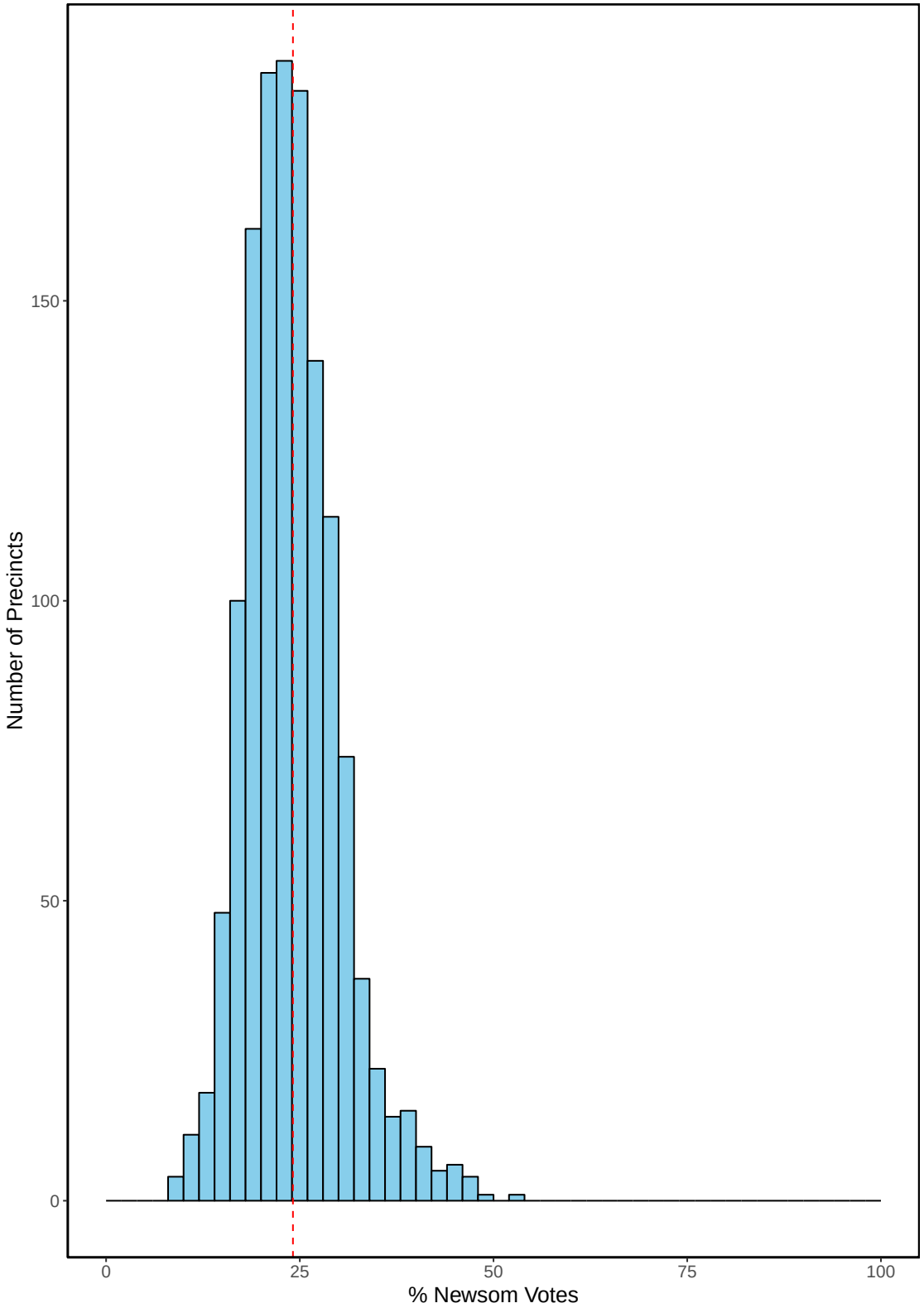
June 15, 2018 data.
Preliminary turnout data from the Orange County Registrar of Voters.
These figures are estimates, and may not reflect the final totals.

Figure 1: Cox Precinct Vote Shares



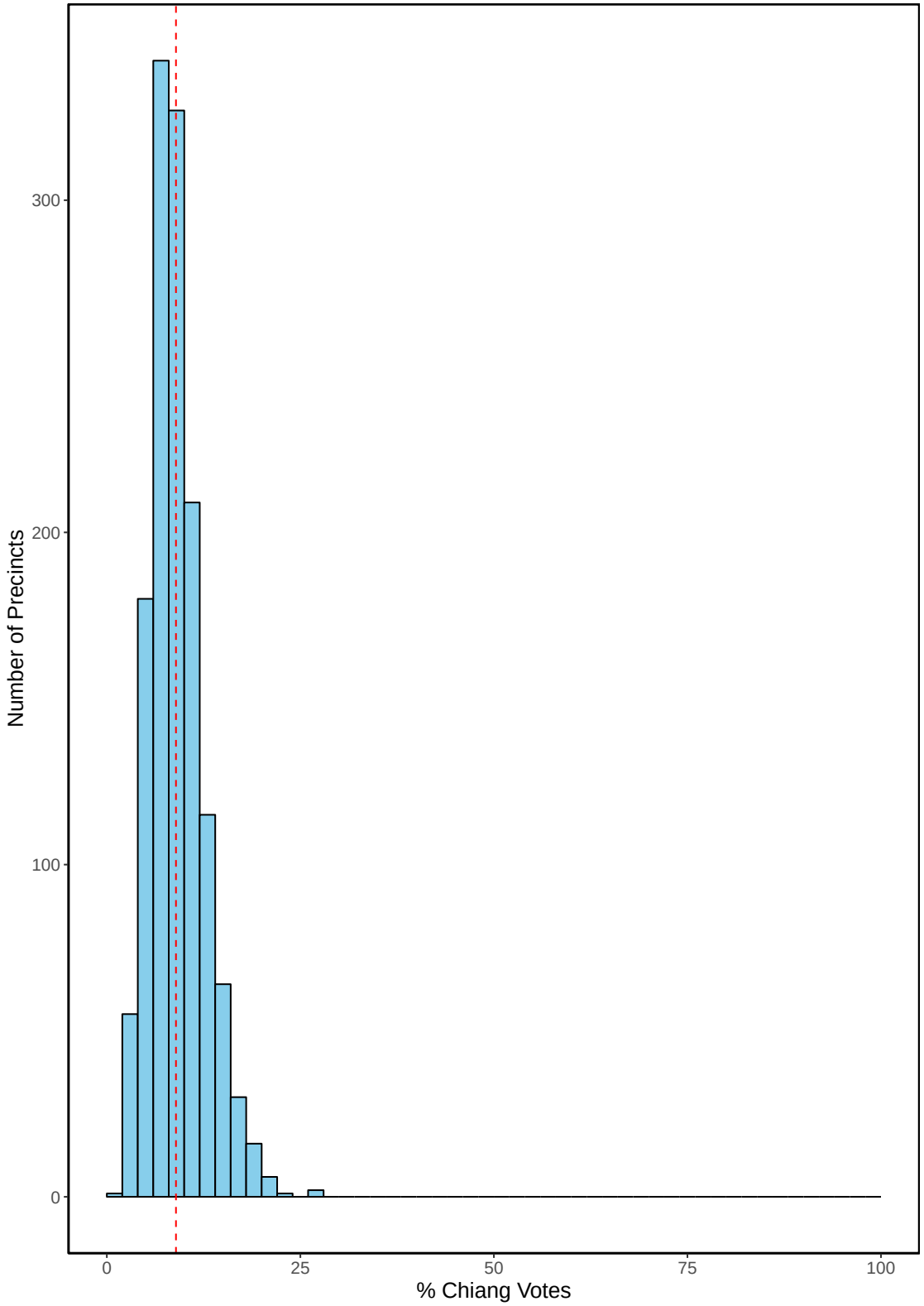
June 15, 2018 data.
Preliminary turnout data from the Orange County Registrar of Voters.
These figures are estimates, and may not reflect the final totals.

Figure 2: Allen Precinct Vote Shares



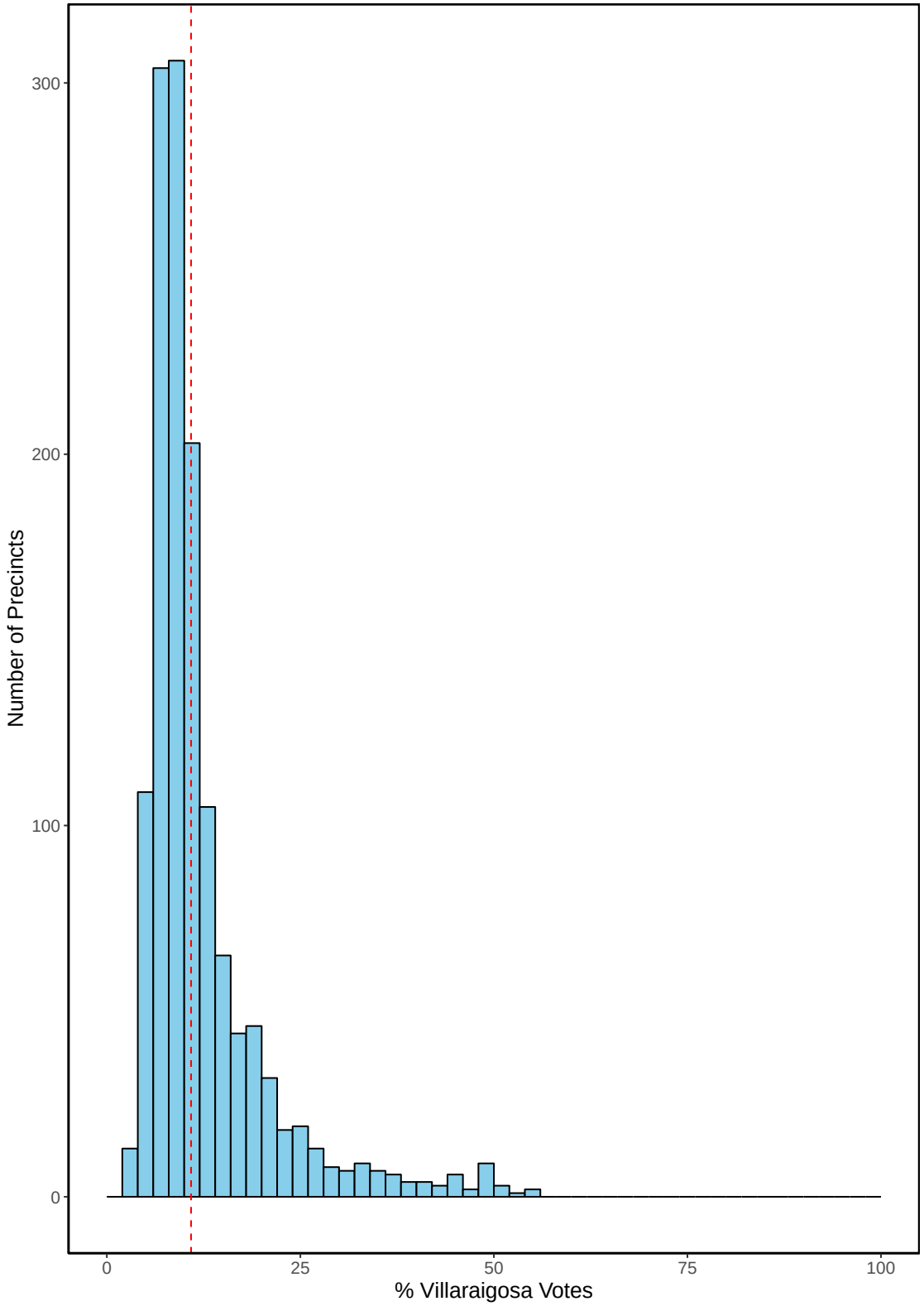
June 15, 2018 data.
Preliminary turnout data from the Orange County Registrar of Voters.
These figures are estimates, and may not reflect the final totals.

Figure 3: Newsom Precinct Vote Shares



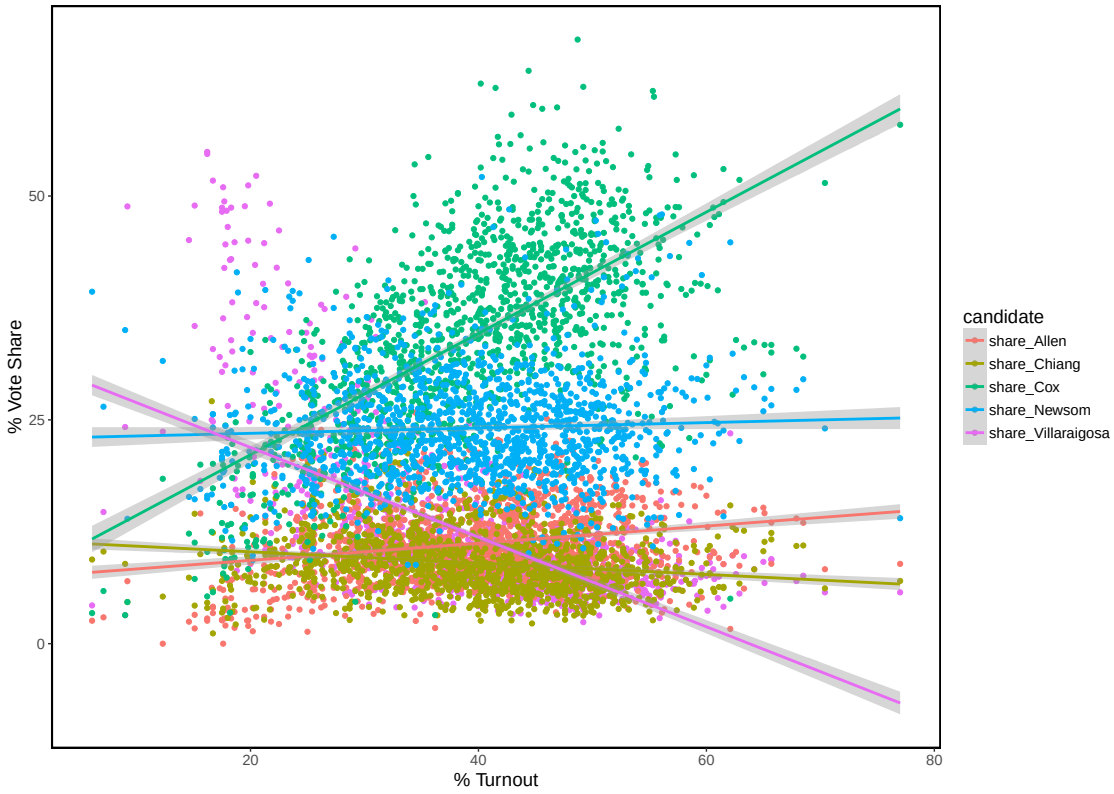
June 15, 2018 data.
Preliminary turnout data from the Orange County Registrar of Voters.
These figures are estimates, and may not reflect the final totals.

Figure 4: Chiang Precinct Vote Shares



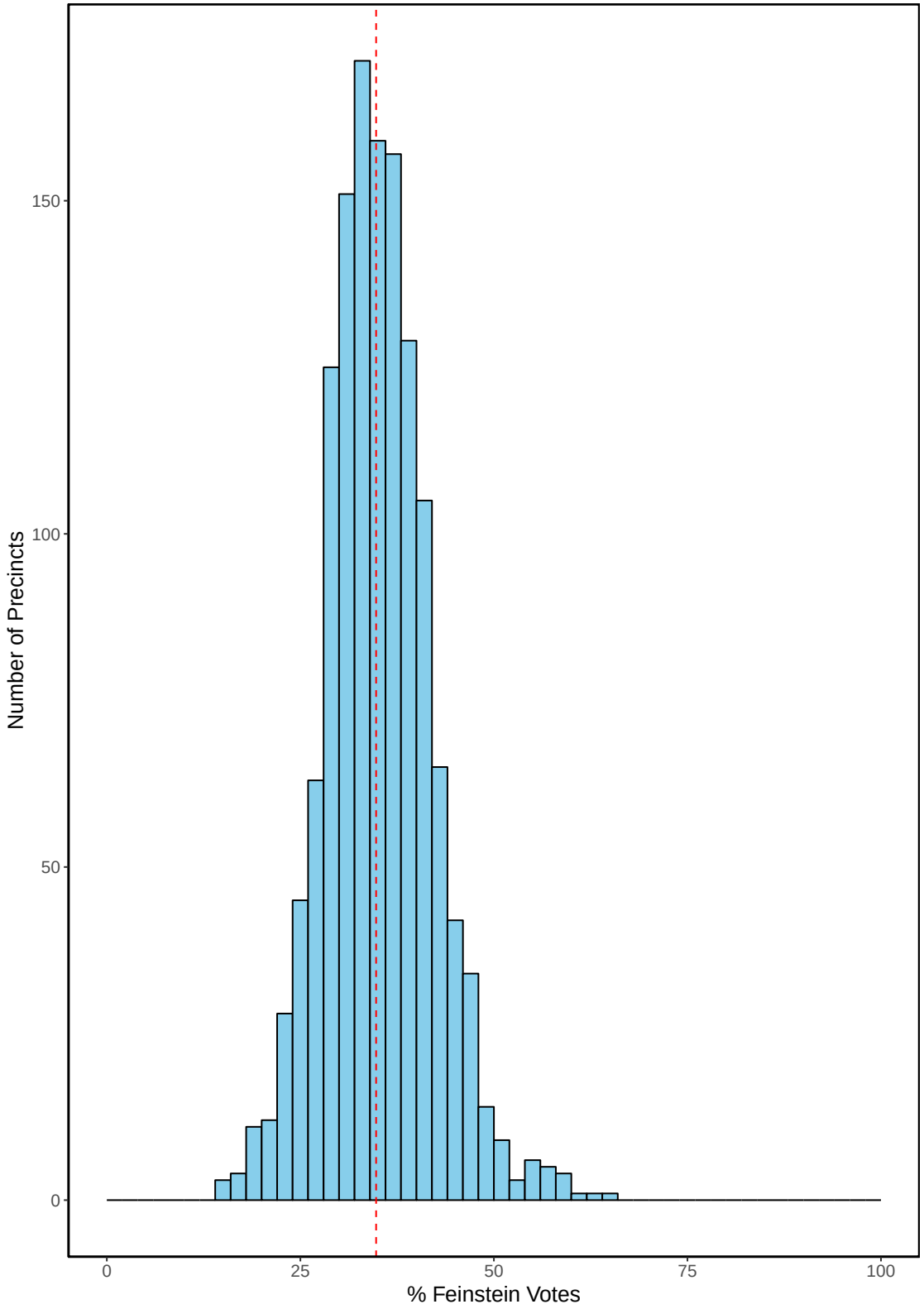
June 15, 2018 data.
Preliminary turnout data from the Orange County Registrar of Voters.
These figures are estimates, and may not reflect the final totals.

Figure 5: Villaraigosa Precinct Vote Shares



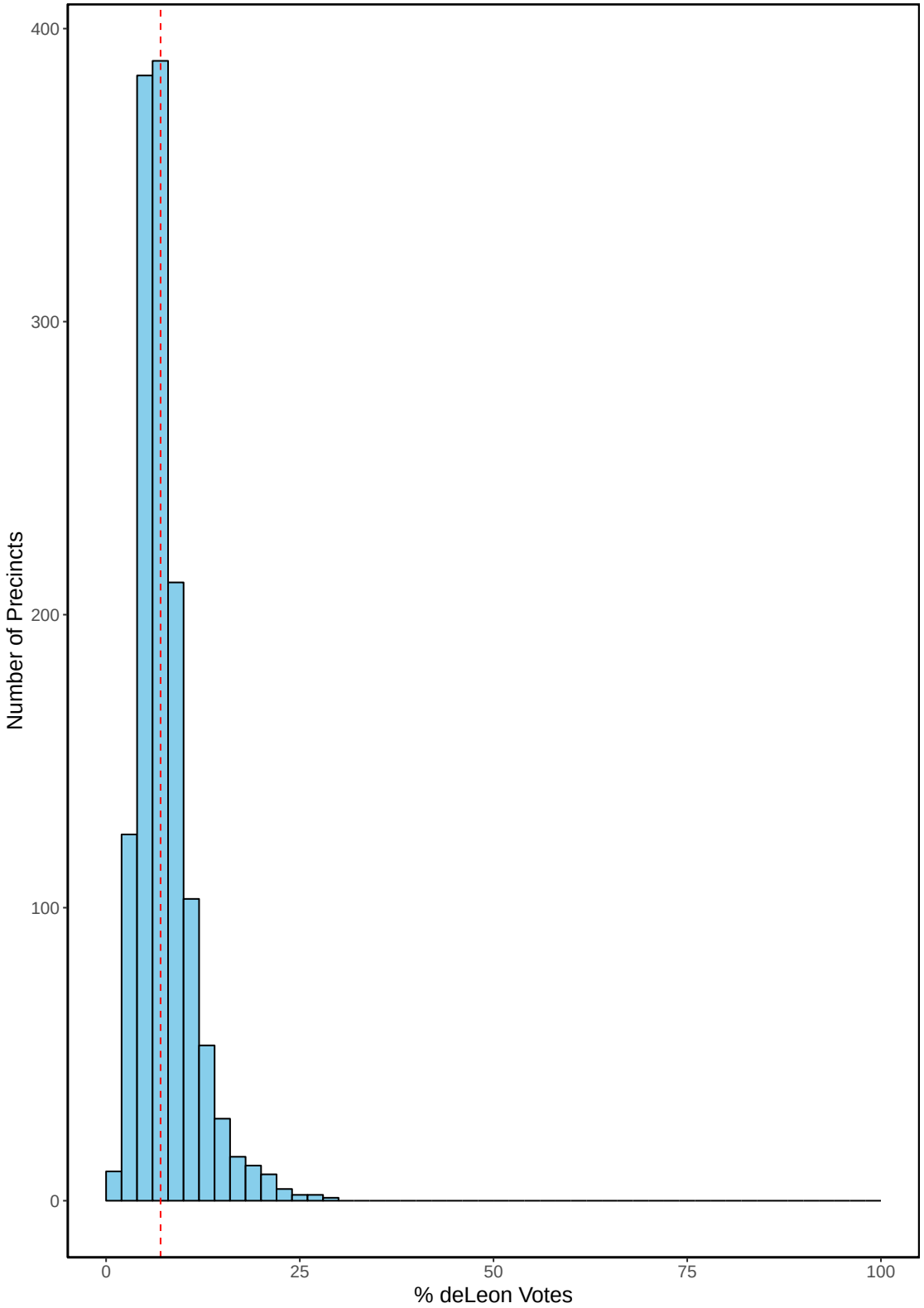
June 15, 2018 data.
Preliminary turnout data from the Orange County Registrar of Voters.
These figures are estimates, and may not reflect the final totals.

Figure 6: Gubernatorial Vote Share by Turnout



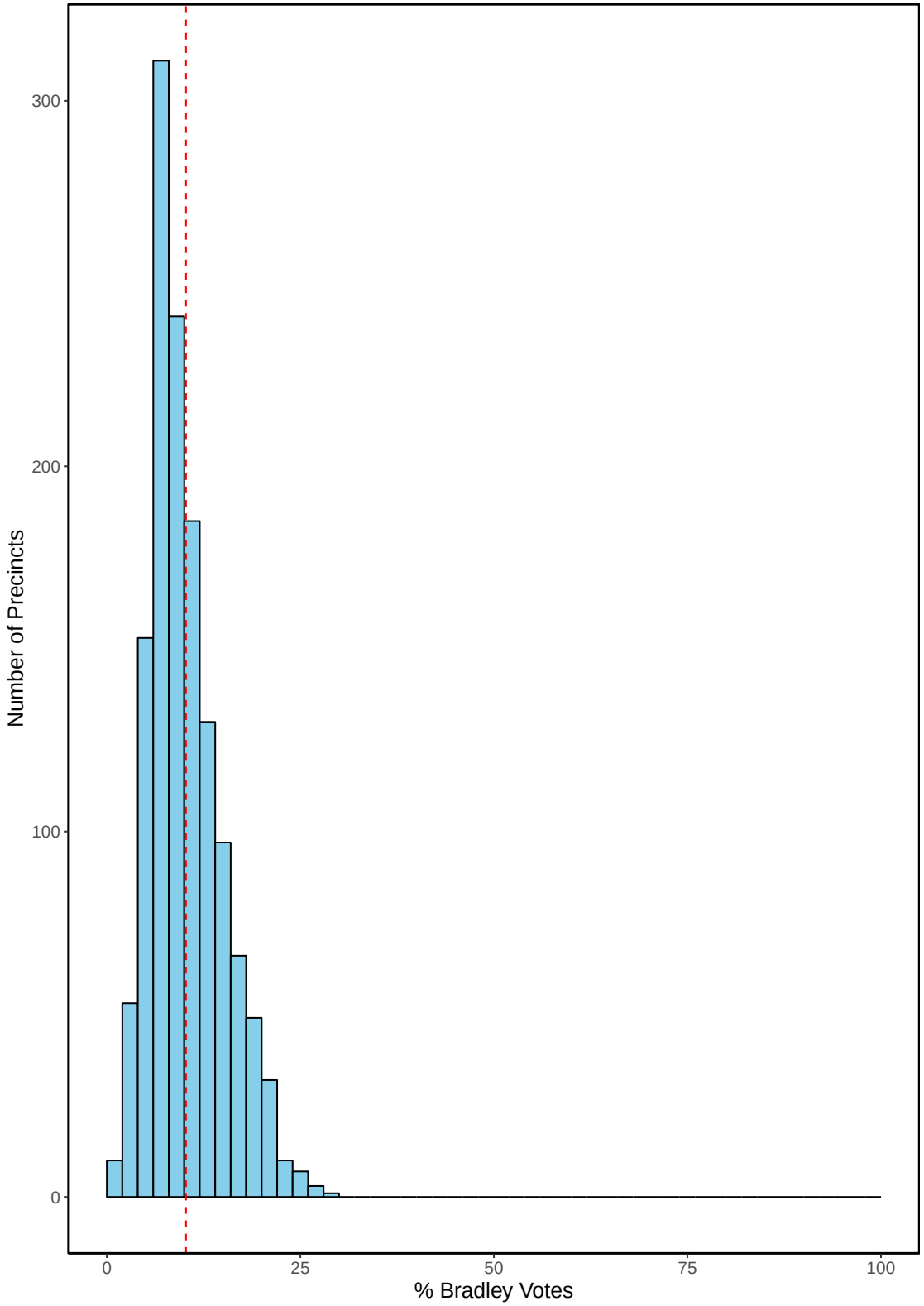
June 15, 2018 data.
Preliminary turnout data from the Orange County Registrar of Voters.
These figures are estimates, and may not reflect the final totals.

Figure 7: Feinstein Precinct Vote Shares



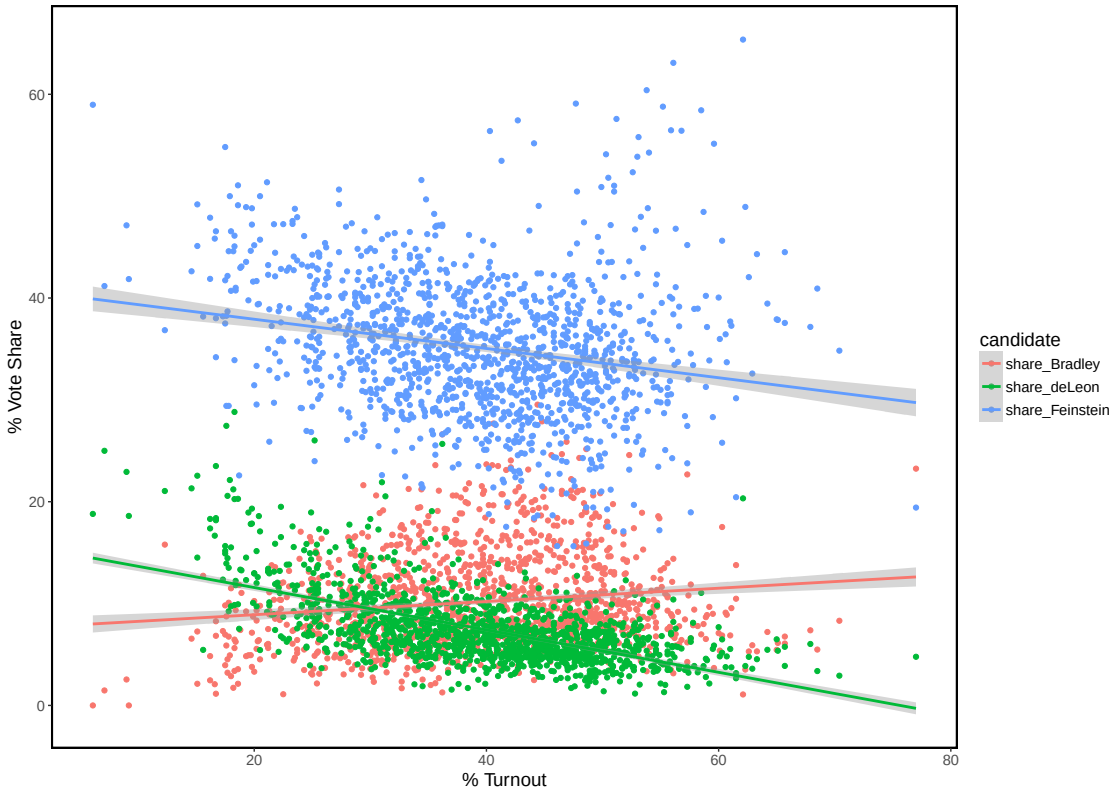
June 15, 2018 data.
Preliminary turnout data from the Orange County Registrar of Voters.
These figures are estimates, and may not reflect the final totals.

Figure 8: de Leon Precinct Vote Shares



June 15, 2018 data.
Preliminary turnout data from the Orange County Registrar of Voters.
These figures are estimates, and may not reflect the final totals.

Figure 9: Bradley Precinct Vote Shares



June 15, 2018 data.
Preliminary turnout data from the Orange County Registrar of Voters.
These figures are estimates, and may not reflect the final totals.

Figure 10: U.S. Senate Vote Share by Turnout