Precision and Bias in the US Electoral College: 1832-2016

Michael Geruso, Dean Spears, and Ishaana Talesara

UT Austin (Population Research Center and Economics Department)

Extended Abstract for PAA 2019

The 2016 United States presidential election reignited a public debate over the role of the Electoral College in American government. The debate is not new. For most years since Gallup began polling public opinion on the issue 50 years ago, more than half of Americans have preferred to abandon the Electoral College system as the method to award the presidency (Gallup 2017). Indeed, debate and contention over the seeming arbitrariness of Electoral College outcomes extend back to at least 1824, the first occasion on which a candidate (Andrew Jackson) won the popular vote and lost the presidency. Given the relatively small sample that history has provided, it remains unknown whether Electoral College-popular vote mismatches are flukes---that is, something possible but ex-ante unlikely. It also remains unknown whether the higher rate of mismatch in the recent period (2 of the last 5 Presidential races) is particular to the current political alignment of states, is fundamental to the Electoral College, or is merely a statistical anomaly.

In this paper, we systematically explore these issues through the lens of statistical theory. Given that attention is perennially focused on the extent to which the EC "gets wrong" the national popular vote, we frame our analysis in terms of the bias and precision of the EC as an *estimator* for the national popular vote. We approach these issues with an eye to contemporary problems, but also take an historical view, documenting long-lived patterns that have persisted for the last century and more. We examine the last stable party system prior to the Civil War (1832-1854), as well as the party system in the reconstruction era (1876 and 1892) and in the most recent period (1988-2016).[1] The particular parties, the political alignment of states, and indeed even the roster of states varied greatly across these periods. This is useful in disentangling whether the bias and precision properties of the EC are fundamental, or happenstantial--depending, for example, on the particular political alignment of states.

To set the stage, our paper and presentation will first illustrate several features of elector apportionment that drive the properties and outcomes we examine in our paper. Our results show, for the four most populous and four least populous states, the number of electors per state population, over two periods: 1832-1852 and 1970 to 2016. Three facts are apparent from this analysis. First, small states are favored by the system, which awards electors on the basis of

[1] Other time periods are less useful in providing identifying variation in electoral outcomes. For example, the period 1896 to 1932, the so-called "Fourth Party System" was comprised largely of Presidential landslide victories.

population (equal to the number of US representatives from the state) plus two (the number of US senators from the state). Whereas California has 1.5 elector votes per million residents today, Wyoming has 5.5 votes per million residents. Second, the coarseness of reassigning a small number of discrete electors (following each decennial census) leads to discrete jumps in elector counts along a continuous population distribution. This creates discontinuities in the voting power time series, such as the one in North Dakota following the 1970 Census, when electors per million persons fell from about 6.5 to 4.7 overnight. The coarseness in the number of assignable electors, which generates these jumps, also creates arbitrary discontinuities in voting power across states within any given election. Third, unequal EC voting power per population across states is correlated with state demographics. The smallest states are, for example, far whiter that the larger states.

In the next section of the paper, we begin by adopting an intuitive definition of *demographic group bias* linked to the EC votes per population. Following the 2016 election, news organizations and commentators have documented---often with significant analytical sophistication---a host of arguably undesirable properties of the Electoral College. In particular, commentators have noted that the number-of-congressmen-plus-two apportionment of electors favors voters who live in small population states, where the extra two electors are spread across fewer voters. Based on current demographics, this has the effect of disadvantaging blacks and Latinos, because these groups disproportionately live in large states. We briefly replicate and extend those facts, drawing out a number of surprising new findings. We show (i) that the bias against blacks that has been highlighted in the popular press is small and unstable through history, (ii) that in many cases the rank ordering of state population sizes does not match the rank ordering of voting population size, and in some cases, states with fewer voting-age adults are apportioned more EC votes than states with more such adults, (iii) that the small state favoritism of the EC implies that the highest voting power resides with rural whites of lower education. To a small extent, these bias also favor men over women, as women are very slightly more likely to reside in an urban area.

The new facts we lay out motivate the core empirical exercise of our paper: we perform Monte Carlo simulation to determine the **statistical properties of the EC as an estimator** for the national popular vote. The Monte Carlo procedure samples from actual state outcomes to generate counterfactual electoral outcomes from historical events.[2] That implied *ex ante* distribution of popular votes versus Electoral College victories reveals two striking patterns. First, largely regardless of the particular presidential election years included in the sampling scheme, there is significant probability of a EC-popular vote mismatch. **We predict the *ex ante* probability of a mismatch in the most recent period of about 15%.** Although it need not have

[2] The sampling frame consists of pairs of {electors awarded, votes received} in all the state × elections in a given period.

been true, the small sample of mismatches that history has provided (5 mismatches out of 48 presidential elections going back to 1824) appears fairly in line with the *ex ante* probability.

Second, we show that the mismatch is not likely to be symmetric across political parties due to the demographic differences across place. As long as geography predicts partisan affiliation in some way, this is a general property. In the most recent period, we find that a Republican candidate earning just over 50 percent of the popular vote has a 60-70% chance of claiming the electoral college, compared to just 30-40% for a Democrat with just over 50 percent of the vote. As an estimator for the national vote, the EC is thus significantly biased and inefficient. The Republican favor is a feature of the recent political alignment of states, though the mismatch is a general property. Similar Monte Carlo simulations drawing electoral outcomes from the period 1836-1852, a time when there were different dominant political parties (*i.e.* Whigs and Democrats) and when there were just 26 states, all in the east. We show a high ex-ante probability of mismatch which favored one 19th century political party over another.

We conclude by evaluating the extent to which several counterfactual policies address one or more of these underlying issues. In part, these counterfactuals highlight which underlying sources of bias and imprecision are the driving forces. For example, the "plus two Senators"[3] is significantly more impactful than the implicit coarseness of fixing the total number of electors at 538. This is because it serves as a statistical weight, in the sense of a weighted average.

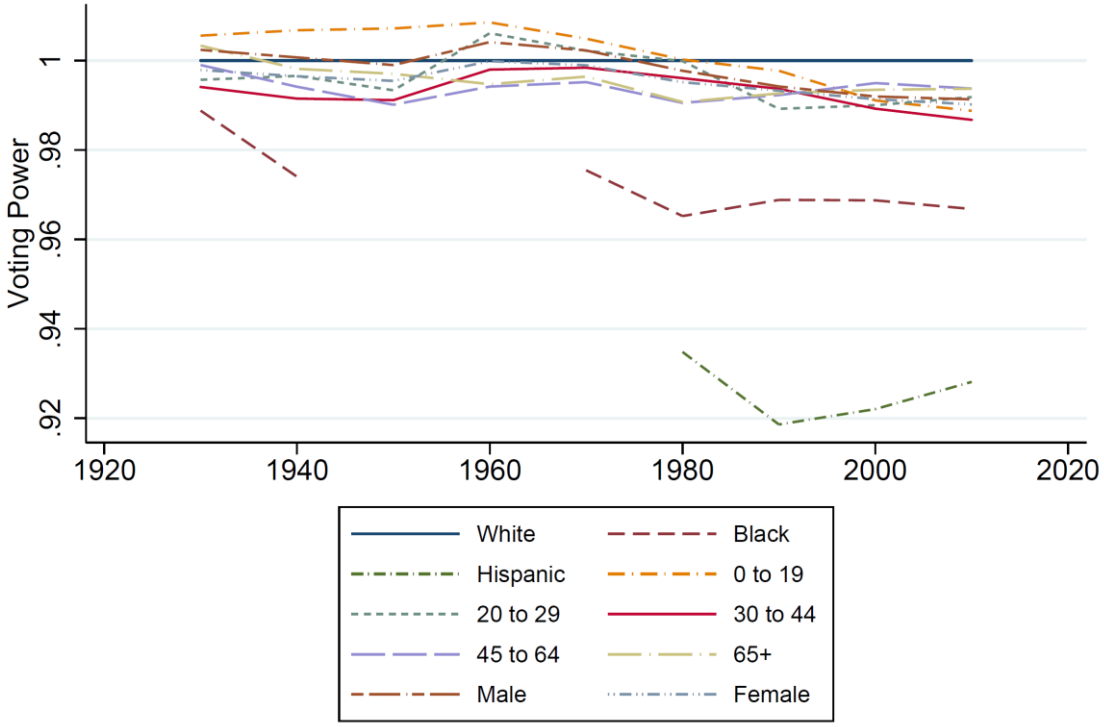
Our paper contributes to timely issue of public policy and popular attention. Although there are deep and difficult normative questions of representation wrapped up in the debate over the Electoral College (hereafter EC) that are purely subjective, there are also unresolved but answerable questions of empirical fact that could inform this debate. Paramount among these is how often and in whose favor should we expect a mismatch. Ours is the first paper to quantify the probability of an electoral mismatch using the estimator framework and Monte Carlo simulation. The findings here connect to a timely public debate, as well as several academic literatures in economics, political science, and demography.

The findings here also relate to a heated ongoing debate in public demography: how the design of the Census may effectively undercount non-citizens. Our findings demonstrate that such questions are important not only because the Census affects the geographic distribution of federal resources and the apportionment of US Representative to Congress, but also because the counting of non-citizens (as well as the counting of minors, legal resident aliens, etc.) plays in important role in the effective voting power of eligible voters in a Presidential race.

[3] A state's electoral college votes is its number of Representatives (which is proportional to the population) *plus two* representing the two Senators for each state.

Preliminary Figure 1: Voting Power by Demographic Groups

Although much of the prior literature has focused on the disadvantage of Blacks, we find that Hispanics have the largest disadvantage (and the females have less power than males).



Preliminary Figure 2: Monte Carlo simulations of counterfactual policies