Electoral College bias and the 2020 presidential election

Robert S. Erikson[a,1], Karl Sigman[b], and Linan Yao[a]

[a]Department of Political Science, Columbia University, New York, NY 10027; and [b]Department of Industrial Engineering and Operations Research, Columbia University, New York, NY 10027

Edited by Larry M. Bartels, Vanderbilt University, Nashville, TN, and approved September 7, 2020 (received for review July 7, 2020)

Donald Trump's 2016 win despite failing to carry the popular vote has raised concern that 2020 would also see a mismatch between the winner of the popular vote and the winner of the Electoral College. This paper shows how to forecast the electoral vote in 2020 taking into account the unknown popular vote and the configuration of state voting in 2016. We note that 2016 was a statistical outlier. The potential Electoral College bias was slimmer in the past and not always favoring the Republican candidate. We show that in past presidential elections, difference among states in their presidential voting is solely a function of the states' most recent presidential voting (plus new shocks); earlier history does not matter. Based on thousands of simulations, our research suggests that the bias in 2020 probably will favor Trump again but to a lesser degree than in 2016. The range of possible outcomes is sufficiently wide, however, to even include some possibility that Joseph Biden could win in the Electoral College while barely losing the popular vote.

elections | president | Electoral College | simulations

The Electoral College is often seen as an unfair institution that can deny the presidency to the popular vote winner, a circumstance sometimes called an electoral "inversion." Some argue that the Electoral College is biased in favor of small states on the grounds that their Electoral College allotments always include two extra votes representing the two senators that the state elects regardless of its population. Others claim that the bias actually favors more populous states because the winner takes all feature gives them excess pivotal power. Thoughtful evaluations of the Electoral College and the possibility of electoral inversions are plentiful [Gelman et al. (1), Miller (2), Cervas and Grofman (3), and Geruso et al. (4) plus the literature cited therein].

The possible distortion that gets the most attention is that the Electoral College, for some combination of reasons, has favored the Republican party historically. Trump's 2016 victory with a lesser share of the vote fuels that narrative. What are the chances of a repeat in 2020? If he loses the popular vote, could Trump still win an Electoral College majority?

This paper addresses the possibility of an electoral inversion in 2020. For this effort, we have examined the degree of Electoral College distortion in past elections and the degree to which it can be predicted in advance from prior state voting patterns. This knowledge is then applied to conditions in 2020. Here is what we found.

Electoral College Bias in Recent Elections

To start, we present a pattern that may surprise. Over the nine presidential elections leading up to 2016, the Electoral College presented little bias, even as it offers some threat of overturning the popular vote winner. Given the configuration of the relative vote divisions across the states, the popular vote winner could sometimes have been denied victory if the vote margin had turned out to be a very close election. However, despite common perceptions, there was no systematic distortion favoring one party over another. (Arguably, the notorious 2000 inversion would not have occurred without the miscounting of vote intentions in Florida.) Then came the 2016 election where Donald Trump was elected handily with less than 49% of the two-party vote.

To measure Electoral College bias in past elections, our tool is the uniform swing (1, 2). For any past election, we move every state's Democratic (or Republican) vote share of the two-party vote by a constant amount. This constant amount can vary, allowing us to calculate the Electoral College outcomes given different national popular votes.*

For instance, in 2016, Hillary Clinton lost in the Electoral College with 51.10% of the two-party vote. We apply the uniform swing rule to add 0.41-percentage points more of the vote to Clinton in every state, making the popular vote 51.51% Democratic. Clinton would have needed to exceed this vote margin to win all three famously pivotal states (Pennsylvania, Wisconsin, and Michigan) and gather enough Electoral College votes to win.†

Was the 2016 distortion typical for previous elections? Fig. 1 shows ranges of the national two-party popular vote (scaled as percentage Democratic) in which the popular vote winner would have lost the Electoral College vote, calculated according to the

Significance

Donald Trump's 2016 victory in the Electoral College without leading in the popular vote has led to speculation of a repeat in 2020. If the Democrats were to win the popular vote but by less than a landslide, could they be thwarted from stopping Trump once again? We examine how Electoral College outcomes are conditioned by how states had voted in previous elections. Our simulations suggest that in 2020, the Electoral College bias is likely to again favor the Republicans. However, the expectation is less distortion of the popular will than in 2016. There is even the possibility that if Trump was to win the 2020 popular vote by a slim margin, it could be he who loses the Electoral College.

Author contributions: R.S.E., K.S., and L.Y. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: rse14@columbia.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2013581117/-/DCSupplemental.

First published October 26, 2020.

*We assume that there are only two relevant parties, that all states are winner takes all (including Maine and Nebraska), and that there are no faithless electors. For this exercise, minor-party candidates can be ignored. In recent elections, only the Republicans and Democrats have come close to winning the plurality of a state's votes.

†For 2016, an Electoral Vote tie emerges in the narrow range of a hypothetical Clinton vote between 51.48 and 51.50%. This hypothetical tie results only when we ignore the separable electoral vote for Maine's second district. If we incorporate Maine 2's electoral vote, when the hypothetical vote is within this range, we would consistently take one electoral vote away from Clinton and give it to Trump, which breaks the tie and gives Trump the win.

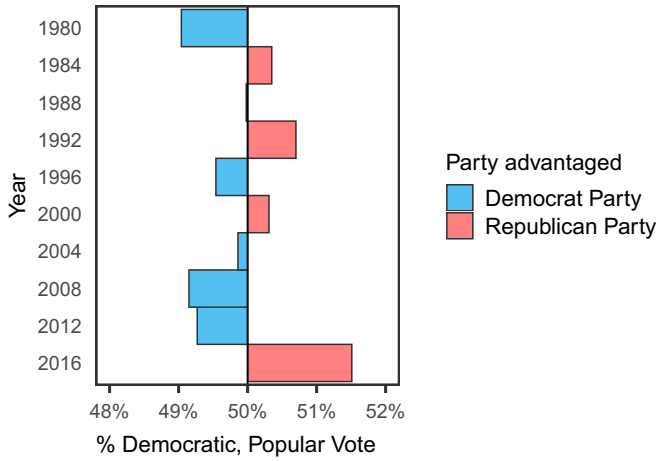

Fig. 1. Estimating the possibility of electoral inversions via uniform swing. The bars represent the range of the national popular vote for which one party would win the Electoral College without a popular vote plurality.

aforementioned procedure. The wider the bar in Fig. 1, the larger is the Electoral College distortion.

Over the 10 elections from 1980 to 2016, there was no obvious systematic bias tilting the Electoral College playing field in favor of one party or the other.[‡] In fact, the three presidential elections leading up to 2016 all showed the Electoral College working slightly in the Democrats' favor. Although it has not granted either party a persistent historical advantage, the Electoral College has offered a mild, seemingly random, perturbation to the outcome, which matters in close elections. The Electoral College's tilt toward Trump in 2016 stands out for its absolute magnitude, with the largest gap out of all elections. We return to the 2016 anomaly below.

Predicting the Next Election's Electoral College Bias from Available History

We would like to predict the likely Electoral College bias in 2020. For this task, we must predict 2020's relative partisan divisions of the states on the Republican vs. Democrat continuum. To do this, we need to take the states' past vote divisions into account. To find the right formula, we must look to the past to see how state vote divisions could be predicted from even earlier state vote divisions.

Fortunately, from the 1980 to 2016 era, a clear and simple answer emerges. To predict the relative positions of the states on the partisan continuum in terms of presidential voting in election year t, the one useful predictor is the set of state vote divisions in the previous election in year $t-4$. The state voting in earlier election years ($t-8$, $t-12$,...) does not seem to matter. This result greatly simplifies the analysis.

Statistically speaking, state presidential voting behaves as an autoregressive model of order 1. *SI Appendix* presents a full discussion.[§] As a basis for simulating possible 2020 Electoral College outcomes, we model state presidential voting over the nine presidential elections from 1984 to 2016.[¶] Pooling across years, we model each state's voting outcome (percentage of vote) in election year t as a function of state voting in election year $t-4$ (one election earlier) and obtain the following equation for each state:

$$Dem(t) = \alpha + \beta Dem(t-4) + u. \quad [1]$$

The α term is a separate constant for each election year from 1980 to 2016 ("year effects"). The parameter β is the regression coefficient for the lagged vote at time $t-4$ and is very close to 1 in our estimates: 0.98. The u term is random and has a normal probability distribution with mean zero; it is the prediction error: the shock from new sources of the vote in year t with an estimated SD (σ) of 3.5%. Armed with Eq. **1** as the data-generating function over all states (including DC), we can simulate from the distribution of possible Electoral College outcomes in 2020 as year t by simulating 51 independent copies of u, 1 for each state.

The Initial Test: Applying Our Model to Past Elections

Before we present the results of this exercise for 2020, we must ask: how useful is it for forecasting the Electoral College division in previous elections? For each election 1984 to 2016 as year t, we modify the data-generating function, modeling Eq. **1** separately for each election year and eliminating data from election year t, the election at hand. We thus have nine equations, each the basis of an out-of-sample retrospective "forecast" of the vote in the excluded election year. The predicted election year is excluded to make this exercise more comparable with that of predicting the 2020 election. Further technical details are presented in *SI Appendix*.

We proceed to simulate the Electoral College outcomes in the nine elections, each from an equation based on data from the other eight elections during the time window. With 10,000 draws for each contest, the result is a probabilistic distribution of Electoral College outcomes from knowing both the national popular vote and the states' immediately prior presidential voting. With this distribution of possible outcomes, we observe how close the actual Electoral College margin in election t was to the center of the distribution of the simulated ones.

Fig. 2 shows the simulations for the nine elections. The simulated Electoral College outcomes are conditional on the actual national popular vote as explained above. In other words, knowing the exact popular vote in election year t, the partisan division of states in election year $t-4$, plus Eq. $\mathbf{1}_t$, the distributions show the range of likely outcomes in the Electoral College. For elections leading up to 2016, the actual vote margins were well within the range of the after-the-fact simulated forecasts. For example, the 2012 analysis in Fig. 2*B* shows the retrospectively likely 2012 Electoral College division when using Eq. $\mathbf{1}_{2012}$, the state vote divisions in 2008, and the 2012 popular vote of 52.0% Barack Obama and 48.0% Mitt Romney. The actual Electoral College division (332 Obama, 206 Romney) was near the center of the simulated outcomes. We replicate similarly benign results for earlier elections. Still, even with knowing the seemingly best estimation equation, the actual popular vote, and the state divisions during the previous election year, getting the Electoral College winner and popular vote winner to match has elements of a lottery. Even with 52.0% of the popular vote, Obama loses the simulated 2012 Electoral College 3.36% of the time.[#]

[‡]As Miller (2) has shown, earlier elections, pre-1980, contained greater potential for vote inversion.

[§]A more elaborate model of state voting is in Gelman et al. (1). They weight the vote in each of the two previous elections equally and consider additional independent variables.

[¶]We avoid including elections prior to 1980/1984 in our historical base due to the greater volatility of states' election-to-election vote stability in the earlier era, particularly in the South. Also, earlier presidential elections tended to be less close.

[#]Of the nine distributions of simulated electoral vote outcomes, one is not like the others. The 1984 simulations show a much narrow range of outcomes, which is a reflection of this election being a Republican landslide. Few states had much probability of voting Democratic, given the 1980 state vote patterns and the fact that President Reagan won with 59% of the two-party presidential vote.

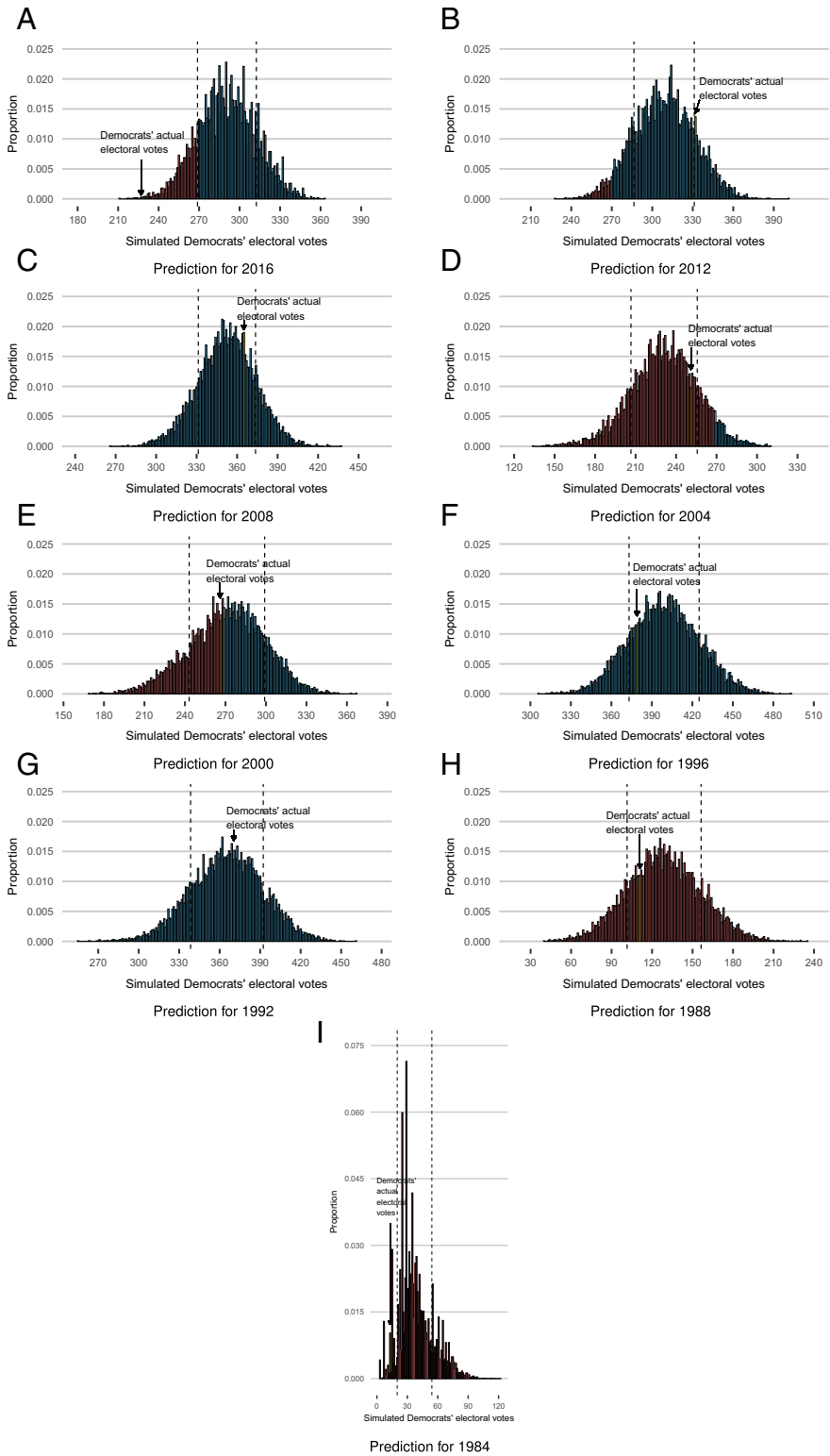

Fig. 2. Simulated and actual electoral votes, 1984 to 2016. The dashed lines represent one SD above or under the mean. The color of bars indicates whether the Democrat would win (blue), the Republican would win (red), or the two parties would tie (white), with the exception that the actual result is marked in yellow. The height of each bar represents the proportion of the 10,000 simulations that yield the corresponding value. Panels *A–I* each present one corresponding year's simulated and actual results, from the latest, 2016 election, to the earliest, 1984 election.

For 2016, the same prediction model goes haywire. Fig. 2*A* shows the range of predictions assuming Clinton's popular vote "victory" of 51.10% in the national popular vote, modified Eq. **1**, and state voting in 2012 as the baseline for prediction. Clinton would seem very likely to be the Electoral College victor, winning 83.04% of the simulations. If Clinton's popular vote margin had been known in advance, her anticipated chance of winning the Electoral College should have been slightly more than four

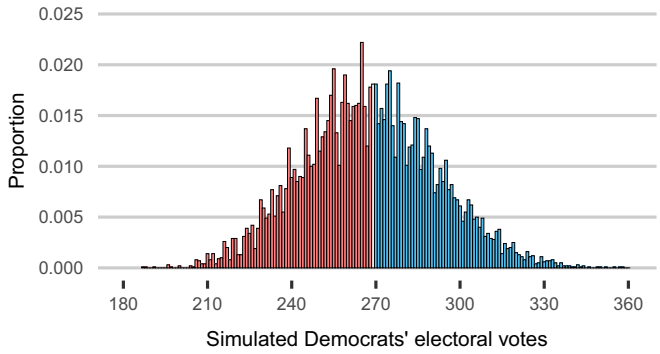

Fig. 3. Simulated electoral votes, 2020. In this graph, we assume that the popular vote margin is the same as in 2016: 51.10% Democrat and 48.90% Republican.

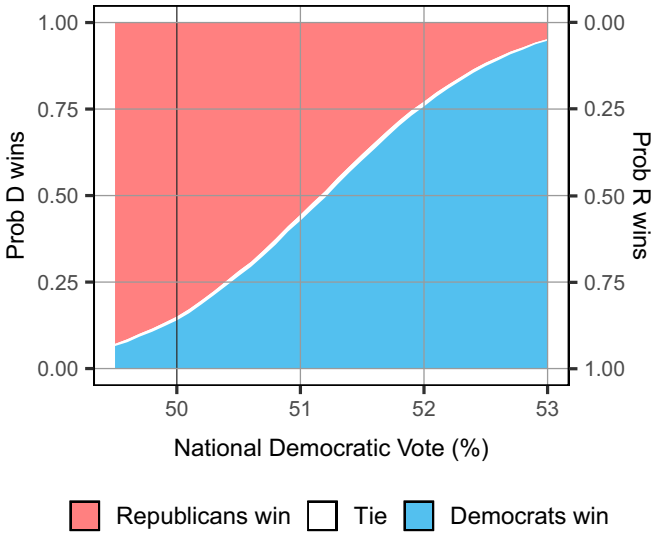

Fig. 4. Probable Electoral College winner, 2020, as function of the national (two-party) popular vote.

chances in five. The actual outcome of only 227 Electoral College votes was in the very tail of the distribution of simulated outcomes.

What happened? Donald Trump got lucky with the variation of the simulated shocks to the 2012 vote. The 2012 to 2016 vote shifts were more Democratic than average in the two largest states (California and Texas) without affecting the state winner. At the same time, these shifts were more Republican than average in three pivotal states (Pennsylvania, Michigan, and Wisconsin), allowing Trump to narrowly win there. If these five states had voted exactly as predicted by our Eq. **1**, the Electoral College would not have tilted for Trump.[‖]

This collection of nine simulations sets the stage for interpreting the predictions for 2020. Each outcome can be seen as a draw from the simulated distribution of likely outcomes. The previews from previous elections allow the predictions for 2020 to be gauged as follows. For the nine sets of simulations in Fig. 2, the mean SD is 23.89 electoral votes. For comparison, the mean of the nine deviations of the actual Electoral College vote from the simulation mean is a similar 21.12 Electoral College votes. We should consider these numbers when gauging the likely accuracy of the simulations for 2020. In effect, the distribution of the errors in prediction from the nine forecasts is similar to the average probability distribution of forecasts.

We are reassured to treat the simulations for 2020 similarly as for the previous elections, except now although we know state vote divisions in 2016, we do not know the 2020 popular vote [*NP*(2020)] in advance; so, we instead use hypothetical values of the popular vote, in which for each such value we perform separate simulated probabilistic distributions of the Electoral College verdict. We give our results in more detail next.

Predicting 2020

So what about 2020, which might be close enough for the Electoral College to matter? As mentioned earlier, for any value we assume for the 2020 popular vote [*NP*(2020)], we can apply Eq. **1** (with detailed analysis given in *SI Appendix*, Eqs. **S1**–**S3**) to simulate state vote divisions in 2020.** For illustration, Fig. 3 assumes the same popular vote margin as in 2016: 51.10% Democrat Biden and 49.10% Trump; that is, we use the value *NP*(2020) = 51.10%. We predict 2020 state outcomes as a function of their 2016 vote margins and independent simulated draws (of u shocks) from a normal distribution with an SD of 3.5%.[††]

The simulations of 2020 suggest once again a Republican bias, although less than in 2016. If Democrat Biden was to obtain 51.10% of the popular vote, he would have a 46.14% chance of winning the Electoral College. With 51.10% of the popular vote, Biden would have almost a 50% chance both of winning the electoral votes-rich states of Wisconsin, Michigan, and Pennsylvania and losing the less rich states of Minnesota, New Hampshire, and Nevada.[‡‡]

We conduct further simulations of the 2020 Electoral College for a range of popular vote outcomes and show the results in Fig. 4. If the popular vote would end in a virtual tie in 2020, our simulations would assign only a 12.00% chance of Trump losing. If the popular vote was 52 to 48 in favor of Biden, Biden would face a similar probability of losing. Note, too, the small probability of about 1.26% where, given any close popular vote, the Electoral College would itself be a tie and sent to the House of Representatives for decision.

Our findings are robust to alternative models of the data-generating process, the results of which we include in *SI Appendix*. *SI Appendix* considers alternative models taking into account the states' vote lagged two elections, year-to-year variations in estimated β, statistical error in the estimate of β, and the variability of the estimated SD of the errors, σ. Overall, we have very similar results: the answer always leads to the likely dividing line between Democrats and Republicans being favored in the Electoral College at about a 51 to 49 popular vote split, although to some degree, the degree of Republican bias in 2020 can be slightly influenced by the variance in u, the state-level shocks. (Greater than average shocks would increase the uncertainty while lowering the bias.) The popular vote ties and very

[‖]Rather than saying Trump got lucky, one could argue that Trump ran the stronger campaign in the battleground states where it mattered, while in California and Texas, each dominated by one party, both campaigns were largely silent. If the luck was by design, one could argue that states' shocks clustered rather than being 51 unrelated draws. However, if so, the exact nature of the clustering is not readily predicted in advance.

**For 2020, we must assume states' proportion of the two-party vote since like the actual vote, we do not know it in advance. Our solution is to ascribe each state's turnout in 2016 to 2020.

[††]The 10,000 simulations require 51 × 10,000 such independent normals. When we compare with using any new hypothetical value of *NP*(2020), we use common random numbers: that is, we use that same set of normals.

[‡‡]If the 2016 national outcome was to repeat in 2020, Biden's chances of winning each of 2016's six most pivotal states are 45.77% in Wisconsin, 47.98% in Michigan, 44.94% in Pennsylvania, 58.03% in Minnesota, 51.98% in New Hampshire, and 63.46% in Nevada.

close elections are likely to favor Trump, with a certain degree of built-in uncertainty.

Conclusion

Using past presidential elections as the testing grounds, we have verified a simulation procedure for forecasting the Electoral College for a given popular vote for a particular year. Our results consist of probabilistic distributions of discrete Electoral Vote divisions. When applied to 2020, we find that the Electoral College's pro-Republican bias that emerged in 2016 persists at about half as severe as in 2016. The inflection point between a probable Democratic or Republican win in the Electoral College is not at a 50 to 50 popular vote but rather, in the range of 51% Democrat and 49% Republican. Additionally, while 51 to 49 is the approximate division of the popular vote at which each candidate has an equal shot at winning, the Electoral College verdict remains probabilistic. Based on our modeling, Trump would have a remote chance of winning even if his support is as slim as 48% of the popular vote. Similarly, if the popular vote is a tie or Trump leads slightly in the popular vote, Biden would have a remote chance of winning, overturning the narrative that the Electoral College favors Republicans. The Electoral College distortion in 2020 will probably tilt in the Republicans' favor as it did in 2016 but to a lesser degree of magnitude, more in line with other recent elections.

The observed 2016 outcome and the probable 2020 pro-Republican bias in the Electoral College is a by-product of the distribution of the two-party vote division in the states. If this short-term tilt of the playing field leads Democrats to despair, they could be reassured by the possibility that the tilt might be transient. As we have seen, from 1980 to 2012 the Electoral College showed little bias and no long-term favoritism for either party. For all of the disruption in its wake, the Electoral College's Republican bias so evident in 2016 could recede in status to a historical anomaly.

Data Availability. All study data are included in the article and *SI Appendix*.

1. A. Gelman, J. Katz, G. King, "Empirically evaluating the electoral college" in *Rethinking the Vote: The Politics and Prospects of American Electoral Reform*, A. N. Crigler, M. R. Just, E. J. McCaffery, Eds. (Oxford University Press, New York, NY, 2004), pp. 75–88.
2. N. R. Miller, "Election inversions by the US electoral college" in *Electoral Systems: Paradoxes, Assumptions, and Procedures*, D. S. Felsenthal, M. Mosh'e, Eds. (Springer, Heidelberg, Germany, 2012), pp. 93–127.
3. J. R. Cervas, B. Grofman, Are presidential inversions inevitable? Comparing eight counterfactual rules for electing the president. *Soc. Sci. Q.* **100**, 1322–1342 (2019).
4. M. Geruso, D. Spears, I. Talesara, Inversions in US Presidential Elections 1826-2016. NBER Working paper 26247 (2020). https://www.nber.org/papers/w26247. Accessed 5 October 2020.