NBER WORKING PAPER SERIES

INVERSIONS IN US PRESIDENTIAL ELECTIONS: 1836-2016

Michael Geruso
Dean Spears
Ishaana Talesara

Working Paper 26247
http://www.nber.org/papers/w26247

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2019

This project is supported by the University of Texas Electoral College Study (utecs.org). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed a financial relationship of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w26247.ack

Inversions in US Presidential Elections: 1836-2016
Michael Geruso, Dean Spears, and Ishaana Talesara
NBER Working Paper No. 26247
September 2019, Revised in October 2020
JEL No. H0,J1,K16

ABSTRACT

Inversions—in which the popular vote winner loses the election—have occurred in four US presidential races. We show that rather than being statistical flukes, inversions have been ex ante likely since the early 1800s. In elections yielding a popular vote margin within one point (one-eighth of presidential elections), about 40% will be inversions in expectation. We show this conditional probability is remarkably stable across historical periods—despite differences in which groups voted, which states existed, and which parties participated. Our findings imply that the US has experienced so few inversions merely because there have been so few elections (and fewer close elections).

Michael Geruso
University of Texas at Austin
Department of Economics
1 University Station C3100
Austin, TX 78712
and NBER
mike.geruso@austin.utexas.edu

Dean Spears
University of Texas at Austin
Department of Economics
1 University Station C3100
Austin, TX 78712
and UTECS
dspears@utexas.edu

Ishaana Talesara
University of Texas at Austin
Department of Economics
2225 Speedway, Stop C3100
Austin, TX 78714
ishaana.talesara@gmail.com

A The University of Texas Electoral College Study is available at https://utecs.org

1 Introduction

Over the last two hundred years, the US presidential candidate with the most votes has lost the election about 8% of the time. The US Electoral College is a perennial focus of popular scrutiny and academic study because of these *electoral inversions*. Gallup has been consistently polling public opinion on the Electoral College since 1967.[1] In all but one survey year since then, between half and four-fifths of Americans have expressed a preference to discontinue the Electoral College system in favor of a national popular vote (Gallup, 2011, 2016, 2019).[2] The most common major concern among 2019 Gallup respondents about the Electoral College was that "the winner of the popular vote doesn't always win the election"—as opposed to, for example, "small-population states have a disproportionate influence on the outcome" or "candidates mostly focus their campaigns on voters in a small number of competitive swing states." In other words, most Americans who oppose the Electoral College do so because of the possibility of inversion.

Beliefs about the likelihood of inversions motivate practical efforts to change or eliminate the Electoral College, too. These efforts include more than one hundred failed Constitutional amendments proposed by members of Congress across many decades,[3] as well as the recent activity around the National Popular Vote Interstate Compact, which has been signed into law by 16 states as of 2020.[4] Meanwhile, many proponents of the current system explicitly motivate their support with claims about which party is likely to benefit from an inversion. Yet many basic facts at the root of these positions and policy actions remain unknown. How ex ante probable, in fact, are electoral inversions in US presidential races? Precisely how do inversion

[1] There was a 1948 Gallup poll with a similar question, but the survey did not ask the question of all respondents. It first screened on respondents correctly identifying what the Electoral College was. Of the respondents that did so, 53% favored discontinuing the Electoral College system; 31% favored retaining it.

[2] Late November 2016 was the low water mark for Americans supporting a constitutional amendment to establish a popular vote for President: 49% supported, with support divided along partisan lines. Except for following the 2000 and 2016 inversions, a majority of survey respondents opposed the Electoral College regardless of party identification. For example, in the 2012 Gallup survey, 71% of Democrat/Lean Democrat respondents and 53% of Republican/Lean Republican respondents favored amending the Constitution in favor of the popular vote to determine presidential elections. Support for abolishing the Electoral College system temporarily dipped among Republican respondents following the 2000 election and fell sharply to historical lows among Republican respondents following the 2016 election.

[3] See Edwards (2011) and Peirce and Longley (1968).

[4] Those 16 states cover 196 of the 270 electoral votes needed for activation.

probabilities vary with the closeness of the election? Is the over-representation of small state populations in the voting system likely to be pivotal in generating inversions (and today to favor Republican candidates), as is often claimed? The literature contains many studies of the Electoral College, but these key questions remain unanswered. The answers are not obvious because the sample of presidential elections is small—only 25 observations per century—making it difficult to distinguish statistical flukes from events that were ex ante probable.

Furthermore, the reform debate depends not only on how probable electoral inversions are today but also on how *stable* such statistical properties are in the face of changing demographics and politics. Two of the last five elections have awarded the Presidency to the loser of the popular vote. Is that indicative of the long-run characteristics of the Electoral College, or an artifact of a couple of unusual political moments? Reforms would affect not only the near future but could persist into the distant future when many features of the political system and electorate—even the set of states—could be different from today.

In this paper, we resolve these questions. We show, for the first time, that the probability of an inversion, conditional on a close election, is high and quantitatively stable—unchanging throughout the history of popular voting and unchanging across alternative future scenarios for presidential politics. We also decompose which features of the Electoral College's aggregation mechanism have contributed to partisan asymmetries, favoring one party or another at various points in history. To do so, we begin by defining a data-generating process for state vote totals in presidential races. Our statistical model is flexible enough to nest the various standard approaches to election modeling in the positive political science literature. Our data-generating process also nests the methods of statistically sophisticated professional forecasters and analysts (e.g., Silver, 2016). We estimate the model(s) using historical state-level voting data extending back to 1836.[5] Sampling from an estimated model yields a probable outcome for each state, and aggregating across states yields a probable national election. By sampling many thousands of these probable presidential elections, we characterize the joint distribution over the national popular vote and the Electoral College outcome.

Because US politics has changed dramatically over the last two centuries, we estimate all

[5]By 1836, citizens rather than state legislatures voted in presidential elections in all but one state.

outcomes of interest separately within historical periods. In particular, we generate results for the Antebellum, post-Reconstruction, and Modern periods, as well as for the early and mid twentieth century. This allows us to characterize whether and how the ex ante inversion probabilities have changed as the underlying data-generating process has evolved.

Because different structural restrictions on the data-generating process correspond to different substantive assumptions about the complex and interacting behaviors of voters, campaigns, and other social and economic forces, we generate results under many alternative estimation approaches within each period. Rather than preferring a particular set of assumptions, we show that the envelope of results described by *any plausible model* generates an informative lower bound on inversion probabilities in close elections. This is true even when we study partisan and geographic configurations of US politics that never have occurred—but could. For example, the hundreds of models we examine differ substantially in their implied probability distributions over the national popular vote and differ substantially in the covariance structure that links state-level voting shocks within an election year. The latter determines, for example, whether Florida and Ohio tend to move together in an election. Key results hold even when, in place of estimating model parameters, we iterate over a large grid of exogenously set variances and covariances representing uncertainty in state-level voting outcomes.

We find that in elections decided by a percentage point or less (equal to 1.3 million votes by 2016 turnout), the probability of inversion is at least about 40%. We show that significant probability of electoral inversion persists at much wider vote margins. These results hold across modeling and estimation approaches and are robust to excluding from our sampling frame the election-year observations in which an inversion actually occurred (1876, 1888, 2000, and 2016). The similarity of results across the diverse catalog of statistical models we examine implies that there need not be any consensus on the best model of election uncertainty to establish the high probability of inversion in close elections.

The results also shed new light on the extent of partisan asymmetry in the Electoral College. In the past 30 years, this asymmetry has favored Republicans. For example, conditional on an inversion occurring, the ex ante probability that it would have been won by a Republican ranges from 62% to 93% across models we analyze (in contrast to the ex post realization of 100%). But partisan advantage—unlike the chance of an inversion in a close race—has varied over time. In the post-Reconstruction period (1872-1888), for example, it was the Democrats who were advantaged.

In Section 5, we decompose how various features of the Electoral College's aggregation algorithm contribute to inversions and asymmetry over US history. These features include (i) the allocation of two electors to each state corresponding to the state's Senators and so not in proportion to population, (ii) the winner-takes-all awarding of state electoral votes, (iii) the rounding errors inherent in dividing the US population across just a few hundred indivisible electors (today, there are 435 House seats and so 538 electors), and (iv) the substantial demographic differences between residents-at-last-Census and voters-on-election-day.[6] Almost all popular accounts focus on (i) and (ii), but we develop several striking facts related to (iii) and (iv). In particular, during the post-Reconstruction era, Democrats' Electoral College advantage had little to do with malapportionment (the two Senator-linked electors) or the "wasted" votes in states won by large margins. Instead, it was closely tied to the wedge between residents-at-last Census used to determine Electoral College apportionment (which included blacks) and voters (which, due to suppression, did not). We show that the turnout-to-population heterogeneity across states remains an important driver of inversion possibilities today.[7] Our findings imply that even if all state governments changed their election laws to split the awarding of Electoral College votes across candidates in proportion to the state vote, and even if federal law changed to inflate the size of the US House to be arbitrarily large, and even if a constitutional amendment were passed that made Electoral College representation proportional to state population (removing the two Senator-linked electors per state), the possibility of a mismatch between the Electoral College outcome and the national popular vote would persist.

The Electoral College is a distinguishing feature of the US political system and so has

[6]Electoral College apportionment is based on the count of persons of any age and citizenship status.

[7]The turnout-to-representation mechanism highlighted here is important across the world as well. Notably, in India the apportionment of Parliamentary seats today is based on population counts in the 1971 Census. As population growth has trended differentially in the north and south of India over the last 50 years, the number of citizen votes that can elect a Minister of Parliament have diverged dramatically across regions of India, leading to skewed representation that favors populations in low-growth states. This creates a wedge between the popular vote and the electoral votes in Parliament, and has the potential to generate an inversion in the election of a Prime Minister. See Appendix A.2 for further discussion of the history and possibility of inversions around the world.

been widely studied across many fields (e.g., May, 1948; Peirce and Longley, 1968; Merrill, 1978; Ball and Leuthold, 1991; Garand and Parent, 1991; Katz, Gelman and King, 2004; see Miller, 2012 for a complete review). Nonetheless, the facts we document here are new. Much prior attention in the economics, law, and positive political science literatures has been focused on demographic inequalities and other facts about the Electoral College—such as effective voting power by geography or race (Banzhaf III, 1968; Sterling, 1978; Blair, 1979), the strategic deployment of campaign resources across states (Strömberg, 2008), the probability of a single voter being individually pivotal (Gelman, Silver and Edlin, 2012; Gelman and Kremp, 2016), or the voter response to perceptions around that probability (Gerber et al., 2019). However, very few empirical papers have quantified any aspect of the probability of an inversion.[8] Estimating the conditional probability of inversion as a function of the national popular vote is a main contribution of this paper and provides a new result to inform the ongoing debate.[9]

Another main contribution of our paper is to establish that the high probability of inversion at narrow vote margins is not a modern phenomenon but has been true for as long as citizens have cast votes for US Presidents. The prior literature has not established this fact, and forecasters whose methods are closest to the methods of this paper (e.g., Silver, 2016; The Economist, Gelman and Heidemanns, 2020) have focused almost exclusively on making predictions of inversion probabilities in the weeks or months preceding some future election—and only for the last few elections. By applying our estimation and simulation backwards through US history, we establish a surprising result. The core estimates of the likelihood of inversions in close races are essentially *unchanging* across radically different facts about politics, parties, and state and voter populations.

The US has grown from 24 states in 1836 to 50 states and the District of Columbia today. Over this time, larger shares of the population (non-whites, the poor, women) have been granted

[8]Although a recent theoretical literature has specifically examined the probability of Electoral College inversions (Lepelley et al., 2014; Kikuchi, 2017; de Mouzon et al., 2018; Kaniovski and Zaigraev, 2018), the stylized mathematical models underlying these studies do not take as inputs actual election-related data, or in some cases, even that there are 50 heterogeneous US states.

[9]Besides providing important new facts to inform the present debate and active legislation around a national popular vote, our study connects to recent work applying econometric techniques to issues at the intersection of economic demography and US politics (Vogl, 2014; Allcott and Gentzkow, 2017; Boxell, Gentzkow and Shapiro, 2017) and particularly to studies with a historic focus (Cascio and Washington, 2014; Gentzkow et al., 2015; Kuziemko and Washington, 2018; Cascio and Na'ama, 2020).

and have exercised the right to vote for President, and different sets of parties have participated in national politics. Our work shows that even as the Union has changed in these ways over the past two centuries, the high probability of an inversion has remained a constant feature of US presidential elections. Thus, the statistical randomness inherent in the Electoral College's tiered system of voting dominates the role of historical variation in demographic, institutional, cultural, or political factors in accounting for inversions. An implication is that the demographic and political changes likely to be experienced in the coming decades, which pale in comparison to the changes experienced over the study period (1836–2016), are unlikely to substantially alter the probability of inversion in a close election.

2 Background and Data

2.1 The Electoral College

The general provisions for the Electoral College (EC) system are established in Article One, Section 2 of the Constitution, though the particular method for determining the number of electors and allocating these across states has varied over time. EC electors are linked to Congressional apportionment, and so their number and geographic distribution have been affected by the various Apportionment Acts of Congress that have set the rules for allocating congressional seats across states over the past two centuries. In particular, the EC electors allocated to each state are equal in number to the state's voting members in the US House of Representatives plus two (for the two Senators of each state).[10] Today there are 538 electors in total: 435 corresponding to US Representatives, 100 corresponding to US Senators, and 3 electors for Washington DC. Washington DC's allocation was established by the Twenty-third Amendment. The present cap at 435 US Representatives and the method for apportionment of congressional seats across states was established by the Reapportionment Act of 1929. As is the case for US House seats, reapportionment of EC electors across states follows each decennial

[10]Below, we use *electors* to denote a state's number of representatives to the Electoral College. We also use *electors* or sometimes *elector ballots* to denote a quantity of apportioned seats in the Electoral College. We use *EC vote* to denote a ballot cast by an elector in the Electoral College as well as to denote a state's action in sending a pledged elector to cast such a ballot. So, a state has apportioned *electors* before election day, but its *EC votes* are only realized once the election occurs.

Census. States individually determine how to award their EC votes in an election. Currently, in all states except Maine and Nebraska, the statewide popular vote winner is awarded all of the state's EC votes, though there is no constitutional requirement to involve citizens in presidential elections at all.

Inversions have been possible—and, we show below, likely—over US history because of this tiered system of voting, in which citizens cast votes for electors, who in turn elect the President. Even absent the possibility of faithless electors,[11] the national popular vote and the EC outcome can diverge for a host of reasons that we detail below in Section 5, where we examine the aggregation mechanics of the EC. At a high level, inversions can occur when EC votes at the second tier can be captured by different numbers of citizen votes at the first tier.

2.2 Party Systems and Sample Periods

Figure 1 describes the periods in US history that we study. Political scientists have identified several stable party systems, characterized by competition between a fixed pair of parties with stable political properties. To define estimation samples, we further restrict attention within party systems to spans of years with stable partisan geographies.[12] This avoids, for example, grouping together election outcomes for the Democratic party before and after the 1960s partisan realignment of the North and South.

We begin our study in 1836, after the Twelfth Amendment changed the rules of the Electoral College and after various state-level reforms rendered the presidential election somewhat similar to our system today. Most importantly, we start only after all states (other than South Carolina) began allowing their citizens to vote in presidential elections. We do not study the Civil War era. Nor do we include in our main sample the first half of the twentieth century, which generated decades of consecutive landslide victories—first for Republicans, then for Democrats.

[11] We abstract away from faithless electors in the analysis, the existence of which could further impact mismatch between the popular and Electoral College outcomes. In addition to the fact that faithless electors are rare and have never flipped a presidential election, the Supreme Court of the United States unanimously ruled in Chiafalo v. Washington (2020) that states may require electors to vote for the state popular vote winner and may punish electors who fail to do so.

[12] Although our periods largely align with external accounts of the start and end dates of various party systems, an overriding principle in constructing the endpoints for our sample periods is the need to circumscribe a period of similar underlying geographic partisan alignment.

Landslides are less informative of the probability distribution of votes around the 50% threshold of interest. Given these restrictions, we study the Antebellum (1836–1852), post-Reconstruction (1872–1888), and Modern (1964–2016/1988–2016) periods.[13] For completeness (though with the caveats noted above) we also generate results for the party systems spanning the early and mid-twentieth century (1916–1956).[14]

The bottom panel of Figure 1 displays the popular vote margin of victory in each US presidential election in our study period.[15] There have been four electoral inversions over this time: in 1876, 1888, 2000, and 2016. There are also reasonable arguments that Kennedy's 1960 victory (outside of our sample) was an inversion too. (See Gaines, 2001 and our Appendix B.) The figure makes clear that, to date, electoral inversions have been limited to fairly close elections. One goal of this paper is to establish the conditional probability of inversion at any level of popular vote margin, including races that are not close. The figure also highlights the key inferential challenge in studying presidential elections: There have been just a few dozen elections in total. A credible empirical analysis has to contend with the model and parameter uncertainty arising from that fact.

2.3 Data

The key inputs to our analysis are the historical election returns by state for each presidential election. Data on state-level vote tallies for each candidate and the size of the state's EC delegation in each election come from the Leip (2018) compilation of state election returns. Where possible, we check these data against Federal Election Commission records. In the few state $\times$ election year instances of disagreement, we rely on state government election records where available. The discrepancies across data sources are generally small, often within single or double digit differences in the overall vote count statewide. Appendix C lists every case for which we update Leip's tallies. For data on state demographics including race and

[13] Because, by design, the study periods in our main analysis are characterized by tighter elections than the overall historical mean, the simulated NPV distributions for these periods are less dispersed than empirical frequency of close elections over the full history of the last 200 years.

[14] In Appendix A.1 we provide further historical context as it relates to sample definition.

[15] This is the absolute value of the percentage point difference in the vote share of the two major party candidates, with shares defined over the two party total. See Section 2.3. We denote this margin of victory measure with Δ.

education, we use IPUMS extracts from decennial Censuses (Ruggles et al., 2018) and the American Community Survey. Further data details are documented in Appendix C.

Following the literature (e.g., Vogl, 2014 and Cullen, Turner and Washington, 2018), we normalize vote shares as a fraction of the total won by the two major candidates/parties. The major parties were the Democrats and Whigs from 1836 to 1852 and Democrats and Republicans for the later periods we examine. The 50% share of the two-party vote is the relevant threshold for our analysis. For example, in 2000 the Republican candidate (Bush) won 48.847% of Florida citizen votes. This equaled 50.005% of votes cast for either of the two major parties. By crossing the 50%, two-party threshold, Bush took all of Florida's EC votes. This two-party normalization simplifies the graphical presentation but does not substantively impact our analysis, as third-party candidates won no EC votes over our primary study periods.[16]

A related but distinct issue is that a third-party candidate could be important in affecting the shares of votes won by the two major-party candidates. In some instances, third-party candidates won a large share of votes nationally. For example, Perot won 19% of the popular vote in the 1992 presidential election, despite receiving no EC votes. We examine sensitivity to various ways of handling third-party votes below.

3 Methods

We construct probability distributions over national election outcomes. We proceed in two stages: First we estimate the statistical model (i.e., the data-generating process) for presidential elections at the level of the states, and then we sample from the estimated model to build distributions of likely outcomes.[17] We do this many times, for many models.

[16] In a robustness check, we extend the Modern sample period back to 1964. In the 1968 election, Wallace won 46 electoral votes across five states. For those five states in 1968, we apply the two-party normalization to determine Democratic and Republican vote shares to estimate our model. For example, Arkansas went 31.0%/30.3%/38.7% for Nixon/Humphrey/Wallace. We calculate the Republican share of the two-party vote as 50.6%.

[17] It is not possible to estimate and simulate our data-generating process at the level of US congressional districts because the frequency of redistricting (after every decennial Census) means that we observe only two or three presidential election data points before the congressional electoral map is redrawn.

3.1 Data-Generating Process / Statistical Model

We flexibly model the data-generating process for a state-by-election-year (st) outcome as consisting of a state expectation, $\overline{\alpha_s}$, and a mean-zero shock (ϵ_{st}), which may be correlated across states in an election:

$$V_{st} = \overline{\alpha_s} + \epsilon_{st} \tag{1}$$

$$\epsilon_{st} = \gamma_t + \phi_{st} + \mathbf{X}_s \boldsymbol{\delta}_t$$

The outcome variable of interest is V_{st}, the two-party vote share for the indexed party (normalized to be Whigs before the Civil War and Republicans afterward) in the state-year, or the log-odds transformation of this vote share.

The compound shock ϵ_{st} includes an election year shock γ_t that is common to all states and independent across years. It also includes a state-specific shock ϕ_{st} that varies independently across states within each election year. The last component of ϵ_{st} is a vector $\boldsymbol{\delta}_t$ that accommodates correlation in the shocks experienced by different states in the same election year on the basis of common state characteristics—for example because some issue or candidate appeals to Western states (in which $\mathbf{X}_s$ is a vector of region indicators) or states with large non-white populations (in which $\mathbf{X}_s$ is the fraction of each state's population that is non-white). We defer parameterizing the distributions of γ_t, ϕ_{st}, and $\boldsymbol{\delta}_t$ until we discuss estimation below.

The statistical model in Equation 1 is a generalization of the consensus approach to modeling uncertainty in US election outcomes in political science. It nests the "unified method of evaluating electoral systems" (Gelman and King, 1994) and its more recent applications (e.g., Katz, Gelman and King, 2004). The unified method, as it is typically applied to legislative elections such as US House seats, estimates the variances of legislative district shocks (σ^2_ϕ) and a common shock (σ^2_γ). By varying the assumptions on the structure of ϵ_{st}, our statistical model can accommodate any typical approach in the positive political science literature.[18] It also nests contemporary forecasting approaches (e.g., Silver, 2016).

[18] See Appendix B for a complete discussion of how our statistical model nests other models in the literature and how it relates to deterministic methods like uniform partisan swing analysis.

Equation 1 serves as both a model to be estimated and—post-estimation—the process from which we sample Monte Carlo draws to generate distributions of probable elections. In the context of estimation, t corresponds to a particular election, like Hayes v Tilden 1876. In the context of simulation, t is a probable election that could have occurred during the period from which the parameters were estimated. In other words, t is a single simulation run, which contains S state realizations. (S=51 in the Modern sampling frame, which includes DC.) Aggregating V_{st} across states yields a national popular vote for each t. Aggregating EC votes, which are implied by each state's voting outcome, yields an EC winner for each t. For each model we estimate, we generate 100,000 election simulations to yield smooth joint probability distributions of popular vote and EC outcomes.

The reason for simulating election outcomes from these models is that the Electoral College is a complex statistical object. There is no analytical mapping from model estimates (i.e., the parameters defining γ_t, ϕ_{st}, and δ_t) to the outcomes of interest. Focusing our discussion on simulation outcomes rather than model parameters—in particular, focusing on the conditional expectation of an inversion for each level of the national popular vote—also facilitates comparisons across models with different assumed shock structures and so different sets of estimated parameters.

This flexibility in Equation 1 is central to our approach. Specification uncertainty is an important challenge in this context: The sample of elections is too small to be confident of any single model for the distribution of potential presidential election outcomes. Therefore, we estimate results under alternative sets of assumptions on ϵ_{st}, including restrictions on the correlation structure that links outcomes across states in an election year. We also vary whether the distribution is parametrically estimated following the literature or built up from bootstrap draws that avoid parametric assumptions.

Below, we report results produced by hundreds of parametric and bootstrap models. For tractability, we sometimes focus attention on 25 named models that span much of the relevant space of model uncertainty.[19] Rather than preferring any single model or estimation approach, we show that the envelope of results described by *any plausible model* generates an informative

[19] Figure A1 in the online appendix lists details for each named model compactly.

lower bound on inversion probabilities. Demonstrating robustness to model uncertainty is a key contribution of the paper.

3.2 Parametric Estimation

Because we estimate many variants of Equation 1, we give names to some focal models. In the "$x1$" set of models—A1, R1, M1, for application to the Antebellum, post-Reconstruction and Modern periods, respectively—shocks to the log-odds vote shares are assumed to be distributed as independent normals, with $\gamma_t \sim N(0, \sigma_\gamma)$ and $\phi_{st} \sim N(0, \sigma_\phi)$. $\mathbf{X}\delta$ is restricted to zero. Thus, each state draws an idiosyncratic shock from the same distribution, and all states receive a common national shock in each election t. This error structure, in which common national shocks are the only source of correlated shocks across states, aligns with the stylized fact in the elections literature that common, national shocks are an important component of the across-election-year variance. These baseline $x1$ models are similar to the model in the Katz, Gelman and King (2004) analysis of the Electoral College, though applied to different study periods and to answer a somewhat different set of questions. We estimate the parameter vector $\theta = \{\overline{\alpha}_{s=1}$, ... $\overline{\alpha}_{s=51}$, σ_γ, $\sigma_\phi\}$ via maximum likelihood.

In other models, we allow subnational correlation in state outcomes, though there are important constraints on our ability to estimate a variance-covariance matrix for state vote shares. For example, the modern study period includes observations for 51 states over the 8 elections that fall between 1988 and 2016. The unconstrained covariance matrix would be 51×51 triangular. Therefore, when effectively constraining this matrix by choosing the $\mathbf{X}$ vector—i.e., making substantive assumptions about which state characteristics could link the shocks between states—we follow the elections literature and recent practice in election forecasting.

In models M2, R2, and A2, we follow FiveThirtyEight's published methodology (Silver, 2016) in using fatter-tailed distributions and an alternative process for correlated shocks. In particular, we use t distributions with one degree of freedom fewer than the number of election years in the sample period. And, in addition to independent state and national shocks described by σ_γ and σ_ϕ, we specify an $\mathbf{X}$ vector that includes region indicators, fraction non-white in the state, and fraction with a college degree in the state. Other parametric models ($x5$, $x7$, $x8$,

$x9$, $x10$) vary the set of characteristics $\mathbf{X}$ permitted to link the state shocks, as indicated below. Parameters $\theta = \{\overline{\alpha}_{s=1}, ... \overline{\alpha}_{s=51}, \sigma_\gamma, \sigma_\phi, \sigma_{\delta_{\text{Region}}}, \sigma_{\delta_{\text{Ed}}}, \sigma_{\delta_{\text{Race}}}\}$ are estimated via maximum likelihood.

It is important to understand that the unknowns of interest here are parameters describing the uncertainty in election outcomes—i.e., the shock process described by $\epsilon_{st} = \gamma_t + \phi_{st} + \mathbf{X}_s\delta_t$—rather than parameters describing the expectations of state election outcomes in past elections. The best unbiased predictor of, for example, the expected Republican vote share in Ohio over elections in the last thirty years is arguably the observed mean of the Republican vote share in Ohio over that period. The challenge lies in statistically describing uncertainty around how these historical elections could have unfolded differently. Our focus on estimating spread is in contrast to studies investigating, for example, how ongoing demographic changes could shift *expectations* of states' future partisan alignment. Nonetheless, we examine below the potential impacts of shifting partisan alignment in key states.[20]

3.3 Hyperparameter grid

A challenge for any study of the EC is that with only a few elections per party system, it is impossible to be confident that estimates precisely reflect the true variance-covariance matrix of random shocks across states. We therefore investigate the sensitivity of our results to *assuming* model parameters that cover a large grid of national shock variances (σ_γ), state shock variances (σ_ϕ), and Census-region shock variances (σ_r). This exercise allows us to assess the importance of parameter uncertainty. Although arbitrarily specifying model parameters would in most settings and for most questions generate uninformative bounds, we show that for the probability of inversions in close presidential elections these bounds are informative (far from zero).

3.4 Non-Parametric, Bootstrap-Based Monte Carlo

Beyond assessing parameter uncertainty, we further address model uncertainty. In place of the parametric assumptions on the error process described above, we perform a bootstrap

[20]Our primary approach examines periods of stable partisan alignment of states in order to estimate unobserved parameters for the electoral data-generating process: variances, means, and covariances of election outcomes. By definition, it would not be possible to estimate these for the moment of a structural break such as a geographic partisan realignment, when model parameters would be changing. When we perform counterfactuals that evaluate possible realignments, it necessarily involves specifying rather than estimating parameters.

Monte Carlo in several forms. The bootstrap procedures conform to the data-generating process described in Equation 1, but, rather than making parametric assumptions on the shocks and estimating these parameters, we draw ϵ_{st} directly from the discrete distributions of historical events.

To generate a single counterfactual election (t), an actual election year outcome is drawn for each state from among the election years in the sampling frame. In the baseline, these draws of election years are independent across states and are made with equal probability among the election years included in the sampling frame. Thus, a simulated election during the Antebellum era might include the Whig vote share in Alabama in 1836, in Arkansas in 1852, in Connecticut in 1840, and so on. Combining a draw from each state yields a counterfactual election. Generating many such elections yields a probability distribution over election outcomes.

We also perform variants on the bootstrap procedure that preserve within-year, across-state correlation in outcomes to various degrees. In one set of simulations, we include a tunable parameter that places excess probability weight on drawing state outcomes from the same realized election year. Within each simulation t, we first randomly (with uniform probability) draw a focal year on which to apply the excess probability mass. In models M4, R4, and A4, we set this excess mass parameter to 0.50, so that for each state there is a 50% chance that the draw comes from the randomly selected focal year for that simulation. The remaining 50% probability is divided uniformly across all years in the sample frame to generate one simulated election.[21]

In models M5, R5, and A5, we use wild bootstrap draws (Cameron, Gelbach and Miller, 2008) from a common pool of discrete shocks experienced by all states over the sample period. Other bootstrap variants are reported below. Among these are cases in which we allow for swing-state bootstrap draws to be correlated.

4 Results

Using the parameter estimates $\widehat{\theta}$, or taking bootstrap draws in the case of non-parametric simulations, we draw Monte Carlo samples to find the joint distributions of national popular

[21] See Appendix D for complete details.

votes and Electoral College outcomes. The summary statistics of interest from these distributions are the conditional probability of an inversion at each popular vote level, Inv(NPV), and the conditional probability that the index party wins the Presidency at each popular vote level, Win(NPV).

4.1 Inversion Rates

Figure 2 reports baseline results from the $x1$ models over the Modern, post-Reconstruction, and Antebellum periods. We generate similar figures for a wider set of models below. In each panel of the figure, NPV along the horizontal axis is the share of the two-party vote won by the Republican candidate (or the Whig in the earliest period). The left panels show the probability distribution over the national popular vote as a histogram. These panels also plot Win(NPV) non-parametrically, as a series of non-smoothed means in narrow bins of the NPV vote share. Table A1 reports the parameter estimates behind these simulations and others described in this section.

If the EC and the national popular vote outcome always agreed, then Win(NPV) would follow a step function that increased from 0 to 1 as the national popular vote share crossed 0.50. For each of the historical periods, Figure 2 shows that Win(NPV) evolves smoothly across the 0.50 vote share threshold.

The mere fact that the Win(NPV) function is smooth rather than discontinuous at NPV=0.50 is not surprising given the history and known mechanics of the Electoral College. The value of these plots is in providing an estimate of the magnitude of inversion probabilities at any vote share. In the Modern period, Win(NPV) equals about 65% at 0.50 Republican vote share, implying that Republicans should be expected to win 65% of presidential contests in which they narrowly lose the popular vote.[22] In the Antebellum and post-Reconstruction eras, the estimated Win(NPV) function shifts but retains a similar overall shape. The slope $\frac{\partial \text{Win}}{\partial \text{NPV}}$ is roughly constant over a wide range—one or two percentage points depending on the historical period. Thus, the electoral system is similarly "responsive" in the Gelman and King (1990) sense

[22] For Republican/Whig vote shares less than 0.50, any Republican/Whig victory is an inversion, so $\text{Inv(NPV)} = \text{Win(NPV)}$ for NPV<0.50. For Republican/Whig vote shares greater than 0.50, $\text{Inv(NPV)} = 1 - \text{Win(NPV)}$.

to citizen votes at various margins of the national distribution.

In Panels B, D, and F of Figure 2, we restrict the axes to focus on closer elections and plot Inv(NPV). In the post-Reconstruction period spanning 1872–1888, Inv(49.99) is about 0.4, and so Inv(50.01) is about 0.6. Thus, a Democratic candidate from this period would be expected to win 60% of elections in which they narrowly lost the national popular vote.

A useful summary statistic is the probability of an inversion, conditional on the election being decided by within some popular vote margin. Denote the popular vote margin as $\Delta = |V^R - V^D|$, where V^P indicates party P's share of the two-party vote. We define $\pi(\Delta)$ as

$$\pi(\Delta) = \int_{0.5-0.5\Delta}^{0.5+0.5\Delta} \text{Inv}(\text{NPV})\, dF(\text{NPV}). \tag{2}$$

In relation to Figure 2, $\pi(\Delta)$ is the conditional inversion function (from the right panels) integrated over the predicted probability distribution of the popular vote (from the left panels) in the range $(0.50 - \frac{\Delta}{2}, 0.50 + \frac{\Delta}{2})$. Calculating $\pi(.01)$ from the estimates in Figure 2, we find that, for a race decided by a one percentage point margin or less, the probability that the result is an inversion is 42%, 44%, and 39% in the Modern, post-Reconstruction, and Antebellum periods, respectively.

Both twenty-first century inversions occurred in elections with small popular vote margins. Clinton in 2016 won the popular vote by a 2.1 point margin. Gore in 2000 won it by a 0.5 point margin. But Figure 2 warns against concluding that only in such close elections could inversions occur. The figure shows that probable EC–NPV disagreement persists even at wide popular vote losses for Whigs, Democrats, and Republicans. Our results indicate that a 3.0 point margin favoring a generic modern Democrat—i.e., 48.5% Republican vote share, or a gap of about 4 million votes by 2016 turnout—is associated with a 15% inversion probability.

In the Appendix, we compare results across 25 alternative models. Figure A1 reports key summary statistics from several parametric and bootstrap models, each with different assumptions and constraints on the data-generating process. The models covered in the figure track the main approaches to modeling election uncertainty from the political science literature and election forecasting professionals, though such models have generally not been applied backwards to the periods of US history we study. The catalog of results shows that the probability of an inversion in a close election is not very sensitive to the particular modeling assumptions, such as the structure that allows for correlated shocks across states in an election year or whether the shock distribution is assumed to have fat tails.

4.2 Stability, over US History and into Alternative Future Scenarios

An important question is whether the the high probability of inversions is a fundamental property of the EC. In other words, is the fact that two out of the last five elections (through 2016) have been inversions a reflection of some deep, stable property of the EC? Or is it the unlikely product of extraordinary circumstances—hanging chads in Florida in 2000 and the unique political moment and candidates in the 2016 election? This matters because the desirability of reform, for many commentators and observers, is tied to the probability that mismatches between the EC and national popular vote will continue.

In this section, we address this question of stability. The probability of an inversion in a close race is strikingly similar for all periods we study. In elections within a one percentage point margin—about 1.3 million votes, based on 2016 turnout—the probability of an inversion is around 40%. In historical fact, six presidential elections of the 46 since 1836 have yielded a popular vote margin within one percentage point. Two of these six have been inversions (three if one counts Kennedy/Nixon 1960).

The fact that $\pi(.01)$ is about 40% across history is true as well for presidential elections in the first half of the twentieth century, in which no inversions occurred. This period was excluded from the main analysis because it contained many landslide victories and so lacked variation around the 50% NPV threshold of interest.[23] With that caveat, we estimate statistics for this period in Appendix E.5. Figure A2 reports results. These show that the probability of an inversion in a race decided by less than one and two percentage points, respectively, is 40% and 36% over the period 1916–1932 and 43% and 36% over the period 1936–1956. The corresponding M1 estimates from the Modern period (1988–2016) are 42% and 35%. The corresponding

[23] This generates an additional dimension of uncertainty in the estimates of close-election inversions over this period.

estimates from the extended Modern period (1964–2016), which estimates a substantially higher unconditional probability of a Republican popular vote majority, are 41% and 33%.

$\pi(.01)$ and $\pi(.02)$ are thus remarkably stable over the entire twentieth-century US, which began in 1900 as a union of 45 states in which Republicans dominated on the West Coast and Northeast and Democratic power was concentrated in the South.[24] Extending the comparison further back in time, we note that even as the set of states has expanded from 24 in 1836; even as non-whites and women have been granted the vote; even as reforms like the 24th Amendment eliminated poll taxes and other obstacles to exercise that right; and even as different sets of political parties have dominated national politics, the conditional probability of an inversion in a close election has been stable. For as long as there has been a popular vote to compare to the EC outcome, the high probability of inversion in a close race has been a constant property of the Electoral College.[25]

But that is the past. One question frequently asked in the public sphere is how the changing demographics across states and in the nation overall will affect Electoral College outcomes in the near and distant future. Could inversion probabilities in close races meaningfully change? Our results suggest not: Though changing demographics may cause changes in party politics, the party alignment of states, and the presidential candidates chosen in the primaries and general election, our study suggests that one thing unlikely to change is the conditional probability of an EC inversion in a close election. Why? Because the type of demographic changes invoked in these hypotheticals—e.g., will Texas' growing non-white population make it a potential swing state in the coming decade?—pale in comparison to the historical demographic changes we study here. For example, our analysis spans periods in which Texas was and was not a state.

To further illustrate the irrelevance of such future possibilities to our main finding, we simulate a range of potential changes to the partisan alignment of voters across states. We do not model the underlying behaviors that might generate these outcomes, but simply ask

[24]An important difference between the early twentieth century results and results from the Antebellum, post-Reconstruction, and Modern periods is that the probability of a close election was much lower in the first half of the twentieth century, making the unconditional probability of an inversion lower.

[25]In the earliest US presidential elections, most states determined how to award EC votes by means other than a statewide vote. Even by the 1812 election, less than half of states determined how to cast their EC votes by a popular vote. We begin our first study period in 1836 both because it aligns with the start of a stable party system and because by this time almost all states determined EC votes by a statewide popular vote.

whether there could be any change to voters' party alignment in large states or swing states that would importantly change our conclusions. To do so, we counterfactually shift the state-level distribution of possible voting outcomes to be more Democratic- or Republican-leaning for individual states or groups of states.[26] As reported in detail in Tables A2 through A4, we variously shift swing states, non-swing states, other large states, states won by Clinton in 2016, states won by Trump in 2016, or groups of such states. The range of the hypothetical shifts we consider includes a 10 point margin shift toward the Republican and a 10 point margin shift toward the Democrat (all relative to the actual M1 estimate).

These permutations are plotted in Figure A3. First and unsurprisingly, the probability of a Democratic or Republican candidate winning the Presidency is highly sensitive to such assumed shifts in the electorate: The range of estimates for the unconditional expected probability of a Republican victory is 10%–83% across counterfactuals. Likewise, conditional on an inversion occurring, which party is likely to win it varies from a 90% chance that the Democratic candidate wins to a 99% chance that the Republican candidate wins. Which party tends to win via inversion also changes. For example, under the assumed realignment in which swing states become more Republican, the probability that inversions favor Republicans increases.

Most importantly, the figure shows that no such change significantly decreases the probability of an inversion in a race decided by within one point. The underlying inversion rate is stable and at least about 40% across the same set of permutations. This range extends up to above 50% across the extreme scenarios considered. Texas (or any other state) shifting its political alignment will not change this fact.

4.3 Robustness to Model and Parameter Uncertainty

The high probability of an inversion in a close election is a result that is robust to various alternative model and sample restrictions. Nonetheless, US presidential elections are rare events, occurring only 25 times per century. There is a serious inferential challenge in estimation with so few data points. To examine sensitivity to this fundamental parameter uncertainty, we

[26] This is a stochastic extension of the (deterministic) "uniform partisan swing" method common in election studies.

next report results that iterate over a grid of exogenously specified variances and covariances in the data-generating process. This hyperparameter approach *assumes* rather than estimates parameters. These models, which include state shocks, regional shocks, and national shocks, are described in full detail in Appendix E.3 and displayed across Figures A4 and A5. The resulting parameter sets span distributions from an underdispersed across-simulation minimum standard deviation of 0.77 popular vote percentage points to a maximum standard deviation of 4.47 percentage points. Across these simulations, the probability of an inversion in a close election is entirely robust.

To visually summarize the impact of model and parameter uncertainty, Figure 3 overlays the simulated election outcomes for about one hundred different data-generating processes of the modern period. Panel A plots Win(NPV). Panel B plots the unconditional distribution of the NPV. A detailed description of every model included in Figure 3 is provided in Table A5. The set includes the hyperparameter approaches. It also includes, for comparison, parametric models (M1, M2, M5), non-parametric bootstrap models (M3, M4, M6), models that omit data from the elections in which an inversion actually occurred (M10, M11, R10), models that extend the modern sample backward to include elections from 1964 to 2016 (M12)—the widest possible timeframe in which "Democrat" and "Republican" are arguably stable identities for our purposes, and further variants.[27]

Among the expanded parametric results displayed in Figure 3 are models that either assign all third-party votes to Democrats prior to estimation or assign all third-party votes to Republicans prior to estimation. Assigning the third party votes in these extreme ways significantly changes the central tendency of the NPV distribution. Despite this, the Inv(NPV) function is indistinguishable between the default handling of third parties and each of these two extremes in Figure 3. (See also Figure A6, which narrowly focuses on just this third party robustness result.)

[27]The M12 model, which covers 1964-2016 and generates a diffuse distribution over the national popular vote, offers a point of comparison with Strömberg (2008). In a paper primarily focused on resource allocation across state races, Strömberg (2008) estimates the unconditional probability of a presidential inversion in a model trained on 1948 to 2004 data, finding about a 4 percent inversion rate. Our unconditional inversion probability, represented by the triangle for M12 in Appendix Figure A1, is 5.9 percent. Our M12 model doesn't perfectly overlap with Strömberg's sample period: We do not extend further back than 1964 to avoid combining election outcomes before and after the partisan realignment of the early 1960s.

Other variants on the parametric results included in Figure 3 are models that make different assumptions regarding turnout. Differences across states in turnout could in principle impact simulated election results because, as we discuss below, these affect the state-specific ratio of citizen votes to EC electors. In practice, however, choices around which numbers to use for relative turnout have no substantive bearing on our results: Figure 3 includes eight variants on the M1 model, assigning turnout according to the actual 1988, 1992, 1996, 2000, 2004, 2008, 2012, and 2016 levels. The probability of an inversion in a close election is remarkably stable across these specifications. (See also Figure A7 for further detail.)

For additional specifications of the bootstrap models, we alter the structure of the bootstrap sampling procedure, varying the extent to which shocks in a simulated election are correlated across states. We do this by tuning the excess probability that state draws come from the same election year in 5% steps from 15% up to 50% as described in Section 3.4. We repeat this procedure for swing states only, sampling non-swing states independently.[28] We also repeat this procedure for "safe" states only, sampling non-safe states independently. Here safe states are those in the top quintile of vote share margin (Democrat- or Republican-leaning) averaged over the sample period. Finally, we also step the M1 and M3 simulated election results left and right along the horizontal NPV axis by adding a deterministic, common shift to the simulated election outcomes. These hybrid models shift the partisan balance mechanically, making the whole country more or less Republican-leaning than estimated, while preserving the correlation structure and estimated variance and covariance parameters that govern election uncertainty. These models help to disentangle whether the Democrat/Republican asymmetry in conditional inversion probabilities is driven by the unconditional probability that the Republican loses the popular vote. It is not. Even as the distributions in Panel B of Figure 3 shift, the Win(NPV) function in Panel A remains fixed. (See also Figure A8 for further detail.)

Across all 109 model variants in Figure 3, the Win(NPV) function in Panel A remains similar. Most importantly, Panel C, which provides the distribution of inversion rates across models, shows that these results establish an informative lower bound for our primary parameter of interest, $\pi(.01)$. The minimum across all models for the probability of inversion in races decided

[28]For this exercise, we follow *Politico* and *FiveThirtyEight* in defining swing states. See Appendix D.5.

by a one point margin or less is 40%. The median across models is 43%.

Despite the similarity in the probability of an EC victory at each level of the NPV and despite the similarity in inversion probabilities in close elections, the models in Figure 3 are substantially different in terms of the simulated elections they produce. Panel B shows that the probability densities over the national popular vote differ. The cross-state correlations—which have been critical in recent innovations in election forecasting—also differ considerably across the models considered. Figure A9—which plots every within-model, across-state correlation term against the model's inversion rate—shows that these models are substantially different by these metrics. Nonetheless, there is little relationship between these cross-state correlation magnitudes (or even signs) and the probability of inversion in a close race.

In sum, Figures 3, A4, A5, and A9 indicate that even if it is not possible to fully identify the data-generating process for presidential elections from the small set of observed elections, our main results are robust to alternative models and parameter sets. Importantly, models with shocks linked by election year, region, racial composition, and educational characteristics produce similar inversion probabilities (e.g., models M2, M7, M8, M9 included in Figures 3 and A1) to models that assume that state shocks are completely independent. This suggests that smaller econometric changes, such as the particular choices around how state demographic variables are parameterized, are unlikely to affect the conclusions here. We illustrate this in Appendix E.4 where we alter the parameterizations of the race and education variables. Figure A10 shows that this has negligible impact on conditional inversion estimates.

4.4 Asymmetry

The probabilities of inversion we estimate are asymmetric across parties. In the past 30 to 60 years, this has favored Republicans: Conditional on an inversion occurring, the probability that it is won by a Republican ranges from 62% to 93% across the 12 modern-era models (Table A8). This range includes models for which the inversion wins for Republicans are dropped from the estimation sample. One can also ask, conditional on winning the Presidency, what is the probability that the victory was generated by an inversion rather than by a popular vote majority? Here there is less model agreement on the precise parameter, though all models show a modern Republican advantage: the probability that any single presidential win arises from a popular vote loss ranges from 6% to 72% across models for Republicans, compared to less than 6% across models for Democrats (Table A7).[29]

Figure 2 shows that the asymmetry in the post-Reconstruction and Modern periods favors the party expected to lose the popular vote. This is not a general property of the EC. Figure A2 shows that the pattern does not hold over the middle of the twentieth century, 1936–1956. Nor does it hold in simulations in the modern period in which we shift the popular vote distribution artificially in order to understand robustness to modeling assumptions regarding third parties. (See Figure A6.) The only sense in which there is a systematic advantage for the popular vote minority is that a party can only win via inversion if it loses the NPV. Because of this, the minority party typically has an *unconditionally* higher probability of winning via inversion simply because it is more likely to lose the popular vote.

In general, partisan asymmetry arises because states are heterogeneous both in EC representation (electors per citizen vote cast) and in partisan alignment. Correlation between these leads to one party or another being advantaged in the EC. The historical and political forces behind this correlation—and therefore asymmetry—have differed over time. So unlike $\pi(.01)$, asymmetry is entirely sensitive to the political context. For example, in the Modern period, Democrats have tended to win large states by large margins and lose them by small margins. In 2016, Clinton won by double digit margins in some of the largest states: 30 points in CA, 22 points in NY, and 17 points in IL. Further, Modern Republicans have been favored, on average, by the disproportionate electoral advantage given to small states by the two Senator-linked electors. In contrast, the statistical asymmetry that favored Democrats in the post-Reconstruction period (Figure 2, Panel F) was in large part due to heterogeneity in turnout, not margins of victory or small state overrepresentation.[30]

The post-Reconstruction asymmetry is instructive in highlighting how turnout shapes

[29] The EC is asymmetric under various other measures of partisan symmetry. For example, graphs of Win(NPV) show that the Electoral College does not meet the standard for partisan symmetry that Katz, King and Rosenblatt (2018) define for legislative elections because, in general, Win(NPV) $\neq 1-$ Win($1-$ NPV).

[30] Although inversions were likely for both parties in the post-Reconstruction Period, and although no Democrat has ever in fact won via an inversion, we find that the *ex ante* probabilities of inversions favored Democrats. See Table A7. Statistics indicating an advantage for Whigs or Democrats in the Antebellum period are sensitive to model choice. See Table A8.

inversion probabilities. The number of citizens who cast votes on election day determines the national popular vote, but the population that determines EC representation is all persons, including non-voters and non-citizens, as measured in the last Census. If the turnout-to-population ratio differs across states in a way that is correlated with states' partisan alignment, it can create a wedge between the probable popular vote and the probable EC outcome. This is exactly what happened in the post-Reconstruction South at the beginning of the Jim Crow era.

Over the 1872–1888 sample period, blacks counted toward the apportionment of EC electors in the South but were disenfranchised. A statistical consequence of the brutal voter suppression was that an EC vote in the South could be won with fewer votes. Because Democrats controlled the South, the typical EC ballot cast for a Democratic candidate during this time was backed by fewer citizen votes. Figure A11 illustrates the point, showing that the number of electors per citizen vote in a state was positively correlated with Democratic partisan alignment of the state in the period. Figure A12 shows that overall and within just the former Confederate states, turnout per population was strongly negatively correlated with the black share of the state population. In other words, by suppressing black votes while benefitting from the apportionment that counted black persons, southern politics delivered an Electoral College advantage to the Democrats relative to the national popular vote.

5 Decomposing Sources of Inversion

5.1 Malapportionment and "Wasted" Votes

It is broadly understood that EC apportionment, which allocates electors to states equal in number to Senators plus Representatives, overweights votes in states with small populations. Today, this malapportionment amounts to more than a three-times difference in EC votes per capita between Wyoming and California. It is also broadly understood that because states award EC votes on a winner-takes-all basis (statewide or—in the case of Maine and Nebraska—districtwide), the aggregation algorithm attaches zero weight to citizen votes cast in excess of the vote needed to generate the slimmest plurality in the voting unit. Large margins in a state amount to many "wasted" votes in the Electoral College.[31] But what is the relative importance of malapportionment versus winner-takes-all state contests in generating inversions and asymmetry? In this section we mechanically alter the EC's aggregation formula and examine how the EC outcome would change, holding fixed the citizen votes. This decomposition sheds new light on the statistical mechanics that underlie inversions.

Holding fixed the set of simulated votes cast by individual voters, we alter the EC system to either (i) eliminate the two electors that each state receives for its Senators, (ii) award each state's EC votes proportionally to the state's popular vote outcome (up to the nearest whole ballot), or (iii) do both simultaneously. Under (i), DC and Wyoming are each apportioned one elector instead of three. Under (ii) a candidate that won 49.99% of the vote in a state with 25 EC electors, such as Gore in Florida in 2000, would win 12 EC votes instead of zero. This exercise is not intended as an evaluation that could account for the endogenous responses of voters or parties to a changing electoral system. Instead, it is meant to illustrate, for example, whether the popular press is correct in asserting that modern Republicans have a statistical advantage due to disproportionately garnering votes in lower-population states.[32] We return to the issue of endogenous response below.

Figure 4 plots inversion probabilities under these alternative aggregation rules over the three periods. We start from an existing set of 100,000 simulated elections and use these alternative rules to aggregate up from simulated votes to an EC winner. In the left panels, we use the same simulation draws as in Figure 2 (models M1, R1, A1). In the right panels, we do the same for the $x2$ family of models that incorporate subnational shocks correlated on the basis of state characteristics.

The alternative aggregation rules variously shrink or shift the range over which inversions are likely. Consistent with popular perception of a modern Republican EC advantage due to small states' alignment, removing two EC electors per state for the Senators shifts the $\text{Inv}(\text{NPV})$ function right in the Modern period. This implies a smaller chance of an inversion that awards the Presidency to the Republican candidate. However, the shift is moderate and merely changes

[31] See Figure A13 for an illustration of vote margins by party in large states over the Modern period.)

[32] See, for example, *The Economist* "American Democracy's Built-in Bias Towards Rural Republicans" (July 12, 2018) and *The New York Times* "Why Trump Had an Edge in the Electoral College" (December 19, 2016).

the partisan balance without markedly reducing the overall inversion probability (the area under the displayed function). For example, in M1, the probability of an inversion within a one percentage point NPV margin changes negligibly from 42.4% to 41.6% with the removal of the two Senator-linked electors, even as $\text{Inv}(\text{NPV})$ moves closer to symmetry.

Because the sources of asymmetry differ over history, these alternative aggregation rules yield different effects in the earlier periods. In the Antebellum-era models, each of the aggregation structures either introduces or exacerbates partisan asymmetry, without much overall reduction in the frequency of inversion. The probability of inversion in a race decided by less than a point is reduced from a baseline of 39% in A1 in Figure 4 to a minimum of 35% across the alternative aggregation rules, while making it much more likely that any inversion is won by a Democrat.

In the post-Reconstruction-era models, removing two electors per state has no effect on asymmetry because Democrats and Republicans tended to split the small states in this period. Perhaps counterintuitively, awarding EC votes proportionally in this period exacerbates partisan asymmetry. We return to why below.

The decomposition in Figure 4 is intended to shed light on which aspects of the aggregation rules in the EC are mechanically contributing to inversions. A related but distinct question involves counterfactuals—how frequently would inversions occur under these alternative rule sets, allowing for politics to endogenously respond to the new system? In Appendix F, we modify our modeling approach to confront this issue. These counterfactuals incorporate a stylized, reduced-form representation of behavioral responses to the changing electoral map, as states move in or out of "battleground" and "safe" status under the counterfactual EC aggregation rules.[33] Results allowing for endogenous responses (Figure A15) tend to generate somewhat higher inversion rates in the counterfactuals, compared to the Figure 4 decomposition, which holds the data-generating process fixed. This suggests that Figure 4 may represent a lower bound on counterfactual inversion rates, as behaviors of agents change to pursue the EC victory condition.

[33]Figure A14 shows how the set of battleground states changes.

5.2 Rounding and Turnout

Popular accounts of what drives EC inversions focus on the plus-two (malapportionment) and winner-takes-all (wasted votes) aggregation rules. Indeed, some proposals for reform would implement exactly these rules.[34] So why would inversions persist under these alternatives? One reason is that apportionment is coarse and infrequent. States' representation in the EC is rebalanced only every ten years. The infrequency implies that as state populations differentially change in the intercensal years, states' per capita electoral representation in the EC drifts out of parity. When the rebalancing occurs after each Census, electors are few in number and indivisible. The small number of electors (which is linked to the size of the US House) generates rounding errors in EC representation.

First, to understand the potential magnitude of rounding effects, Figure 5 considers counterfactual sizes of the US House of Representatives. The present size of the House is 435, which leads to 538 electors. In the figure, we consider how EC aggregation would be affected by various sizes of the House, up to 10,000. We apportion Congressional seats across states in accordance with the 2010 Census population with increasingly fine granularity as the House size increases. We then apply M1's estimates for vote totals to each state. To isolate the role of rounding errors, the plots in Figure 5 consider the aggregation rules that remove two EC electors per state and replace the winner-takes-all rule with awarding EC votes proportionally.[35] Malapportionment and wasted votes can play no role in the figure.

Figure 5 shows that inflating the number of seats in the US House of Representatives, and therefore the size of the Electoral College, would further shrink inversion probabilities.[36] But it would not produce an inversion probability below about 10% in a close race decided by within

[34] Andrew Yang's 2020 proposal for Electoral College reform involves keeping the EC but "making electors determined on a proportional basis," presumably mirroring one of the exercises in Figure 4 exercise. See, for example, https://www.politico.com/2020-election/candidates-views-on-the-issues/elections/electoral-college/.

[35] There are two types of rounding errors simultaneously addressed here: As the house size grows, rounding errors in apportioning US House Districts across states are reduced. And because this exercise assumes that EC votes are awarded proportionally with state popular vote, rounding errors in dividing states' whole EC votes between presidential candidates are reduced. For example, take a House size of 5,000. Florida would have 305 House seats and EC electors in the Figure 5 calculation. Winning Florida 50.005 to 49.995 (as Bush did in 2000) would yield 153 EC votes from Florida for the state winner and 152 for the loser.

[36] House size effects on presidential races have been investigated in the prior literature, though previous studies have examined the impacts assuming a deterministic model of presidential elections. See, e.g., Barthélémy, Martin and Piggins (2014). In such studies, it is not possible to calculate inversion probabilities as a function of House size, which we do in Figure 5.

one percentage point. The remaining source of divergence between the NPV and EC is that electors are apportioned according to population-last-Census, which includes all residents of all ages, measured up to a decade prior to election day. In contrast, election-day turnout includes only some adults of the current population.

Panel A of Figure 6 illustrates the role of population growth in this, using the case of the reapportionments that followed the 1980, 1990, 2000, and 2010 Censuses. The figure shows, for a few of the largest states, that differential population growth can quickly cause EC representation to diverge across states. The figure plots house seats per million current persons in the state. The statistics jump following each Census as reapportionment brings representation across states closer to parity (up to a rounding error due to the small House size). But the disparity reemerges immediately and continues to grow until the next reapportionment. In fact, by the time of the first presidential election following reapportionment, the census count used is already two or four years out of date. Election years ending in zero—just before a reapportionment—are likely to be especially skewed by this measure of representation.[37]

But even resolving this (via some hypothetical just-in-time Census on election day), and even combining a just-in-time Census and apportionment with any of the alternative aggregation rules considered above, it would not be possible to eliminate inversions in the US Electoral College. Panel B of Figure 6 illustrates the final major hurdle: turnout differences across states relative to current state populations. Electors are apportioned according to population, but election-day turnout includes only voters in that election. There is significant variation across states in the ratio of persons to votes cast. This is in part because of differences—both systematic and random (Alvarez, Bailey and Katz, 2008; Fujiwara, Meng and Vogl, 2016)—across states and election years in the turnout of eligible voters. It is in part because of differences across states in the proportion of non-citizens, disenfranchised felons, and other non-voting-eligible adults. And it is in part because of differences across states in the ages of residents. In 2016, among the lowest turnout-to-current-population states was Texas at 29%, which had a median age of 33.6

[37] Apportionment for the 2020 presidential election, for example, was based on populations on the last Census day, April 1st 2010. In the intercensal decade, the population of Texas, with 38 Electoral Votes from the 2010 Census, has grown by about 15%; the populations of swing states Pennsylvania and Ohio have grown by about 1% each; and the population of Illinois, with 20 Electoral Votes, has declined.

in the 2010 Census. The highest was New Hampshire at 55%, which had a median age of 41.1 in the 2010 Census.

The voter-turnout-to-current-population heterogeneity was wider at other points in history with different political forces and institutions. In 1888, the ratio ranged from 7% in South Carolina, which had the largest black population share in the 1880 Census (61% black), to 26% in Colorado, where the 1880 black population share was 1%. In the Antebellum period, slaves counted towards the apportionment-relevant population at a rate of three-fifths and cast no votes.[38] Accordingly, for these time periods we find less convergence between the popular vote and EC outcome under the alternatives in Figure 4, all of which ignore the discrepancy between apportionment-relevant persons and voters.[39]

In the post-Reconstruction era, as Section 4.4 noted, the suppression of black votes can be understood in terms of this turnout-to-population ratio. Blacks counted fully toward apportionment, but their disenfranchisement meant that an EC vote was controlled by fewer voters in states with large black populations. This advantaged the Democrats, who controlled the South. Indeed, there is a strong negative relationship over this time period in EC votes per person and the black share of the state population: Even focusing just on former Confederate states with large black populations, the states with the largest black populations could control an EC vote with fewer (white) citizen votes. (See Figure A12 and Appendix E for further details.)

Malapportionment, winner-takes-all awarding of EC votes, coarse apportionment of House seats, and turnout heterogeneity all contribute to inversion possibilities. In some historical periods these phenomena reinforce each other in generating partisan asymmetry. In others, including the post-Reconstruction era, they act as counterbalancing forces. Indeed, correcting one source of inversion without addressing the others need not improve agreement between the NPV and the EC outcome. In the case of the post-Reconstruction-era, removing the distor-

[38]The extreme turnout-to-current-population ratio in California in 1852 in Figure 6 reflects the fact that California's population in the mid-19th-century experienced rapid change in size and composition due to the gold rush. Our calculation of the 1852 population for the purpose of this figure is based on exponential interpolation between the 1850 and 1860 decennial Censuses—the same procedure done for all states for all non-Census years. But the extreme growth between Censuses in California was front-loaded in the decade, so that the 1852 interpolation is an underestimate, generating an overestimate of the turnout-to-current-population ratio.

[39]Figures A16 through A18 show the interaction between inflating the House and other changes, over each of the study periods.

tion caused by winner-takes-all without also addressing turnout heterogeneity (due to voter suppression) merely makes inversions more asymmetrical towards Democrats. In summary, there is no guarantee that any change to the Electoral College system, short of implementing a national popular vote, will reduce the probability of inversion or of asymmetry.

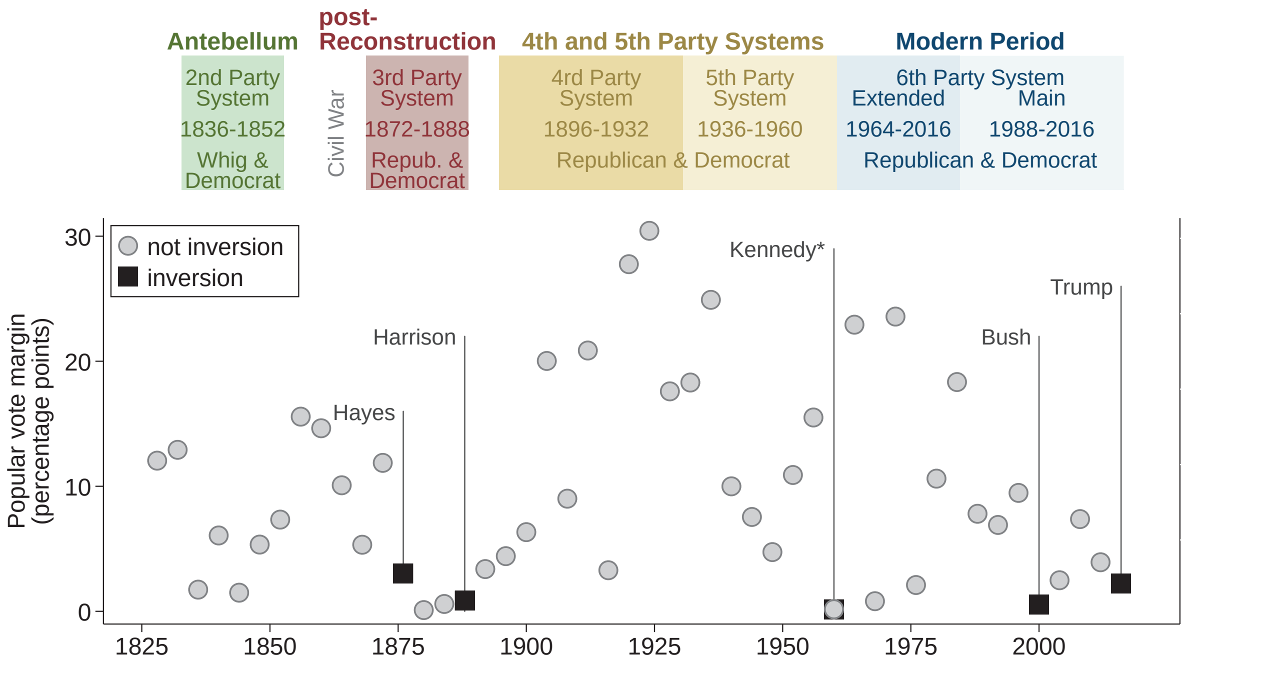
6 Conclusion

A robust finding of every model considered here is that inversions are likely in close elections—where "close" includes elections with popular vote margins in the millions. A game-theoretic equilibrium for two-party competition (Downs, 1957) is a close election, which may be why US presidential popular vote margins have often been small in stable party systems. Recent decades have resulted in particularly close elections relative to most of the twentieth century (see Figure 1). Our findings imply that if elections continue to remain close, frequent inversions are likely.

Ultimately, the EC system adds random—though not mean-zero—noise to the popular vote outcome. Feasible policy changes shrink the variance of—but do not eliminate—this noise, reducing the range over which inversions are likely, though at some margins actually increasing the probability of inversions.

Our paper shows that mismatches between the EC and the NPV have been historically rare events only because presidential elections—and, in particular, close presidential elections—have been historically rare events. We conclude that electoral inversions are not statistical flukes but are enduringly fundamental to the US Electoral College system. No tweak of election rules short of moving to a national popular vote will prevent a chance of inversions in close elections.

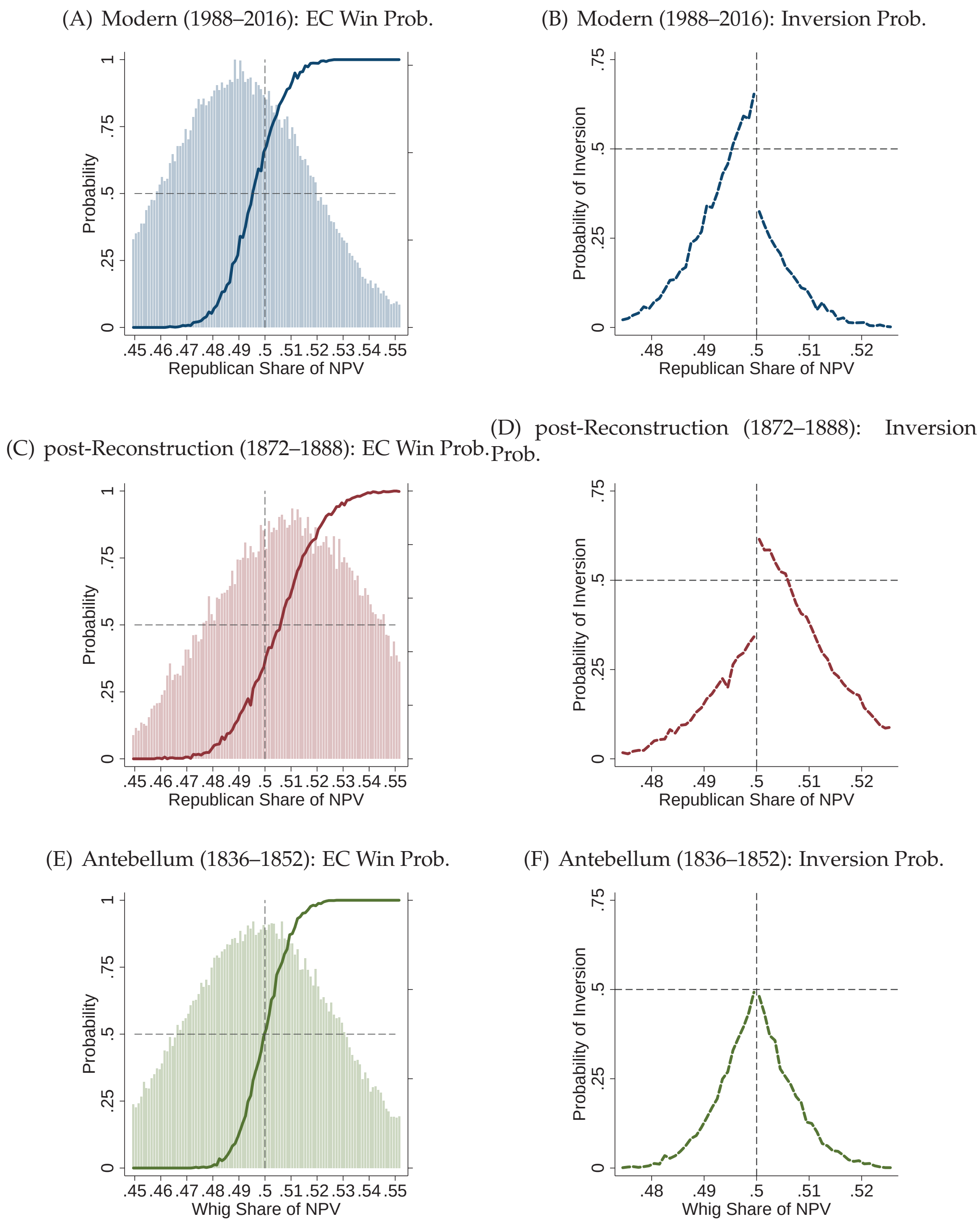
Figure 1: Background: Parties, Victory Margins, and Inversions in US Presidential Elections



Note: Timeline shows the periods of stable "party systems." The plotted points indicate the national popular vote margin for each US Presidential election from 1828 to 2016. The margin is measured as the difference in vote shares of the two major parties competing in the election. These shares are calculated as the fraction of the two-party vote total won by each of the two parties. States-years with no citizen vote for President do not contribute to the national popular vote statistics. There are 4 widely acknowledged inversions: 1876, 1888, 2000, and 2016. All were won by Republicans.

*In the 1960 election, Kennedy arguably lost the popular vote to Nixon despite winning the Electoral College; see Gaines (2001) and Appendix B.

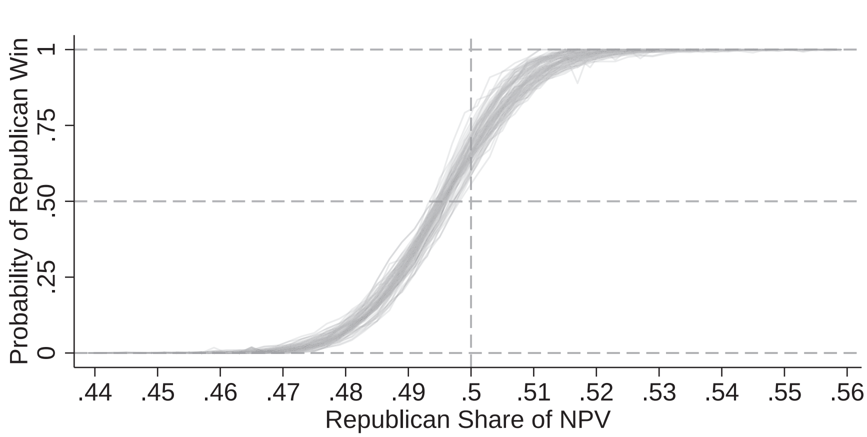
Figure 2: Simulated NPV Distributions and Inversion Probabilities, Baseline Estimates



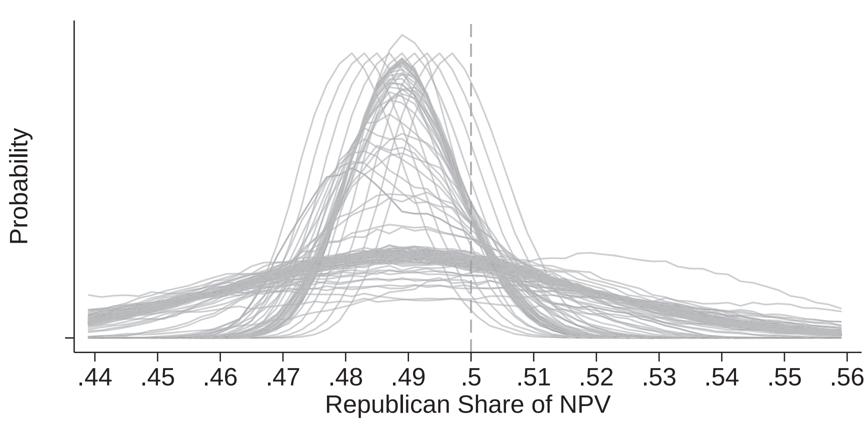
Note: Figure shows inversion probabilities and probability distributions over national popular vote (NPV) outcomes implied by the parametric estimates of the baseline model (M1, R1, A1). Rows correspond to different historical periods, as indicated. Each panel consists of 100,000 simulated election draws. The Whig and Republican national popular vote shares run along the horizontal axes. The solid lines in the left panels (A, C, E) trace the conditional probability of a Whig/Republican electoral win at each level of the Whig/Republican vote share. In the left panels, win rates greater than zero for Whig/Republican vote shares $< 50\%$ indicate inversions in favor of the Whig/Republican candidate. Win rates less than one for Whig/Republican vote shares $> 50\%$ indicate inversions in favor of the Democrat candidate. The right panels (B, D, F) plot the inversion probabilities at each level of the vote share.

Figure 3: The Conditional Probability of Inversion is Invariant to Model & Parameter Uncertainty

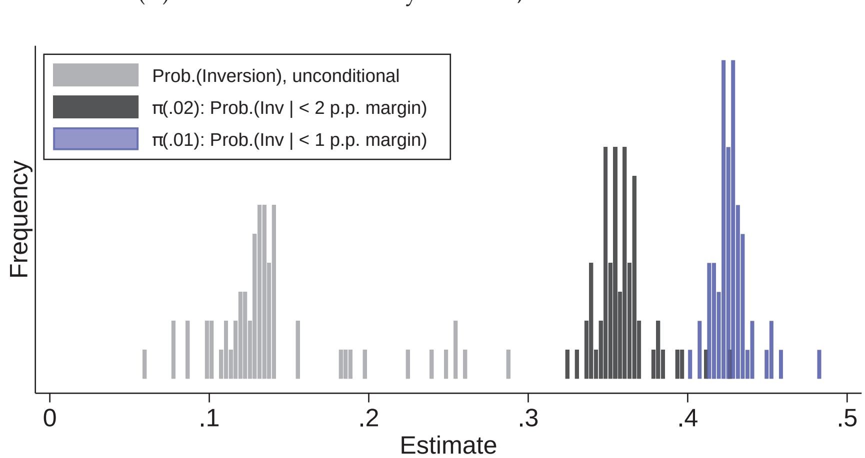
(A) Probability of Republican EC Win at Each NPV, Across 109 Models



(B) National Popular Vote Distributions, Across 109 Models

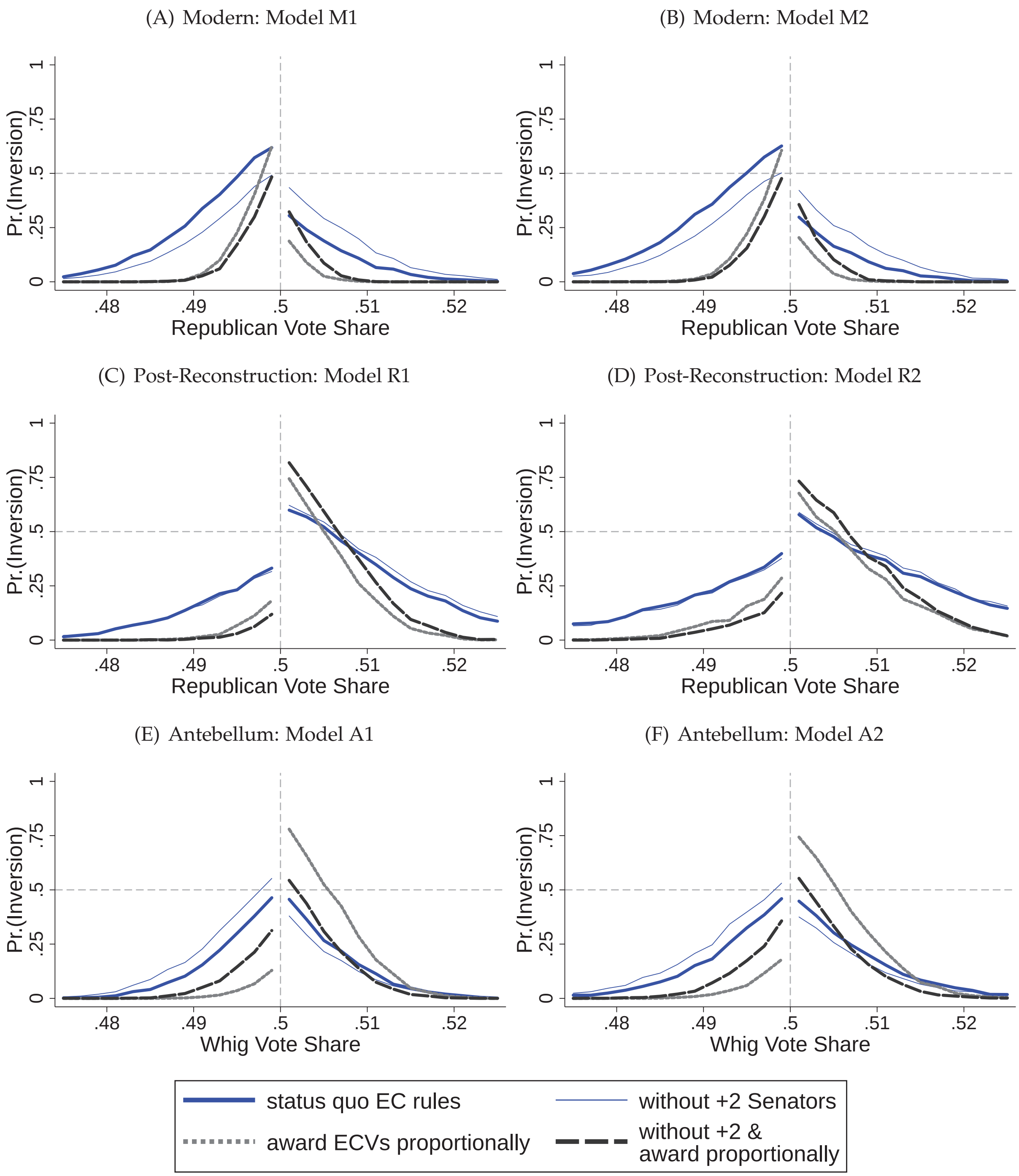


(C) Estimates of Summary Statistics, Across 109 Models



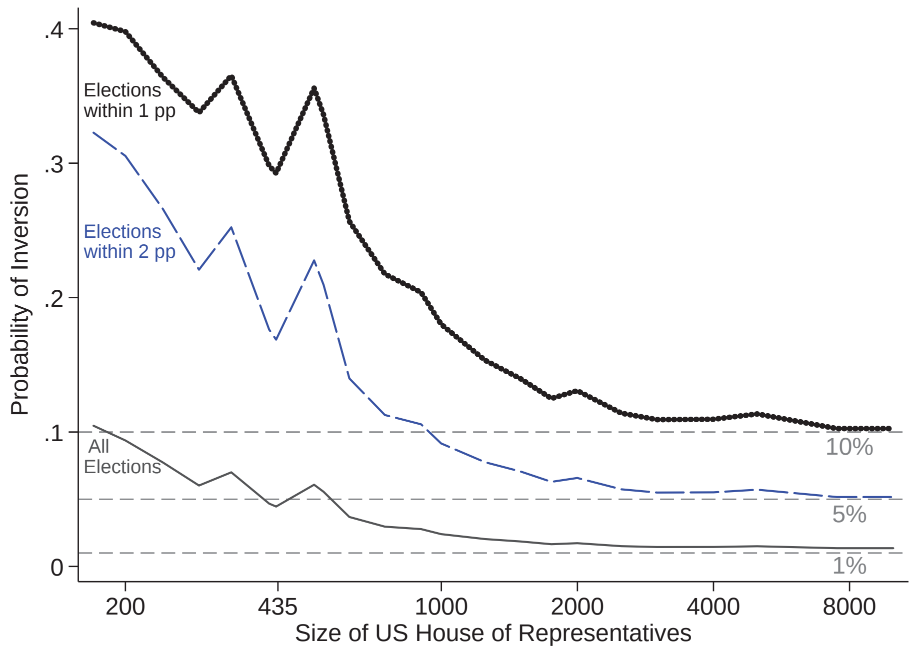
Note: Figure shows statistics under various modeling restrictions and approaches. NPV is national popular vote. All models are for the Modern sampling frame. See Table A5 for a detailed list of each model included. For each model, 100,000 simulated elections are drawn. Panel C displays histograms of summary statistics across these models: the unconditional inversion probability and the probability conditional on a close election decided by a margin of less than one or two percentage points.

Figure 4: Inversion Probabilities Under Alternative EC Aggregation Rules



Note: Figure illustrates inversions under alternative EC aggregation rules that translate citizen votes into a presidential winner. The alternative that removes the two Senator-derived EC electors assigns each state electors equal in number to the size of the state's US House delegation. The alternative that removes the winner-takes-all condition awards state EC votes (ECVs) according to each candidate's popular vote share in the state, up to a rounding error.

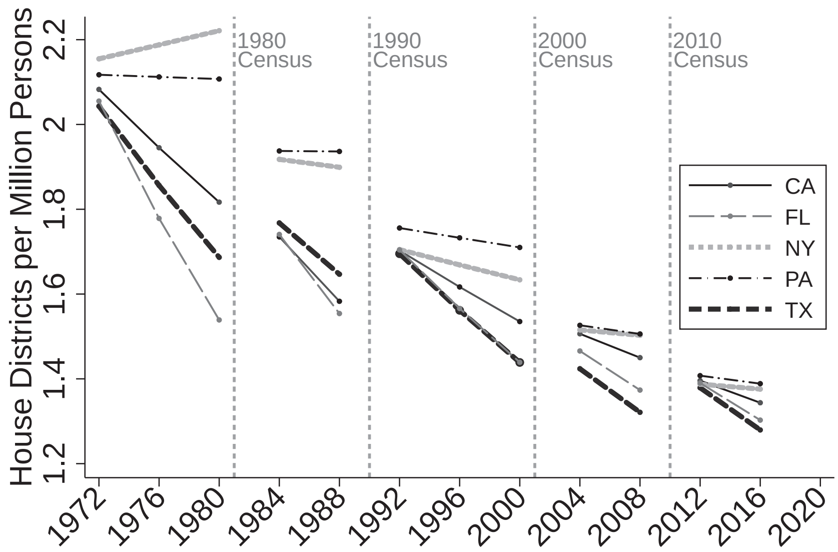
Figure 5: An Arbitrarily Large House Does Not Eliminate Inversions



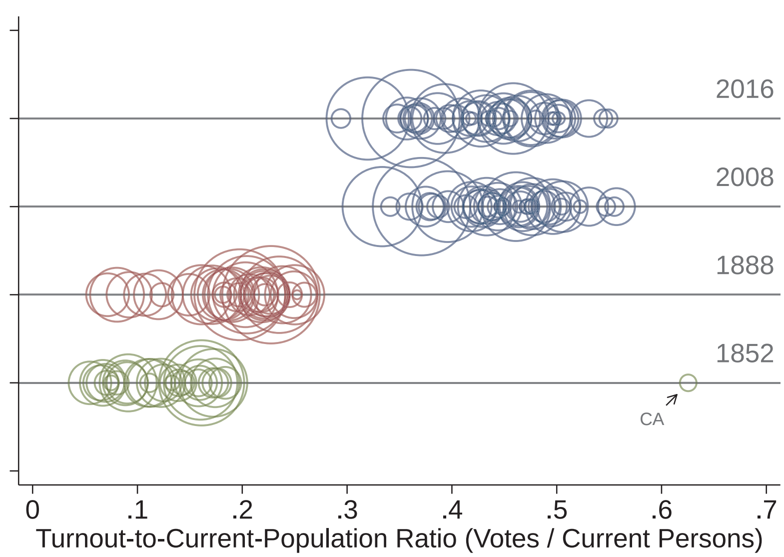
Note: Figure shows how inflating the size of the US House of Representatives affects inversion probabilities in our simulated election outcomes. The lines trace the probability of inversion unconditionally and conditionally on a one or two percentage point (pp) margin, $\pi(.01)$ and $\pi(.02)$. As the House grows, rounding errors are reduced in allocating whole House Seats across states. In addition to inflating the House, the exercise applies the aggregation rule set from Figure 4 that removes two EC electors per state (corresponding to Senators) and awards EC votes proportionally within a state. Therefore, as the House grows, rounding errors are also reduced in how (whole) state electoral votes are awarded across the candidates. For the exercise, Washington DC is apportioned electors equal in number to the the number apportioned to the smallest population state.

Figure 6: Differential Population Growth and Turnout Create Disparities in EC Representation

(A) Population Growth Between Censuses Leads to Disproportionate EC Representation



(B) Turnout as a Fraction of Current Population Differs Across States



Note: Figure shows how EC representation per voter is heterogeneous across states due to population growth between Census counts and due to turnout differences across states. Panel A plots House districts per million persons over time. House districts per person are rebalanced after each Census reapportionment. Panel B shows the extent to which the ratio of turnout (voters on election day) to Census persons differs across states. Each marker is a state. Marker sizes proportional to population. The various horizontal lines in Panel B correspond to different election years as indicated.

References

Allcott, Hunt, and Matthew Gentzkow. 2017. "Social media and fake news in the 2016 election." *Journal of Economic Perspectives*, 31(2): 211–36.

Alvarez, R Michael, Delia Bailey, and Jonathan N Katz. 2008. "The effect of voter identification laws on turnout." California Institute of Technology Social Science Working Paper 1267R.

Ball, William J, and David A Leuthold. 1991. "Estimating the likelihood of an unpopular verdict in the electoral college." *Public Choice*, 70(2): 215–224.

Banzhaf III, John F. 1968. "One man, 3.312 votes: a mathematical analysis of the Electoral College." *Vill. L. Rev.*, 13: 304.

Barthélémy, Fabrice, Mathieu Martin, and Ashley Piggins. 2014. "The architecture of the Electoral College, the House size effect, and the referendum paradox." *Electoral Studies*, 34: 111–118.

Blair, Douglas H. 1979. "Electoral College reform and the distribution of voting power." *Public Choice*, 34(2): 201–215.

Boxell, Levi, Matthew Gentzkow, and Jesse M Shapiro. 2017. "Greater Internet use is not associated with faster growth in political polarization among US demographic groups." *Proceedings of the National Academy of Sciences*, 201706588.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2008. "Bootstrap-Based Improvements for Inference with Clustered Errors." *The Review of Economics and Statistics*, 90(3): 414–427.

Cascio, Elizabeth U, and Ebonya Washington. 2014. "Valuing the vote: The redistribution of voting rights and state funds following the voting rights act of 1965." *The Quarterly Journal of Economics*, 129(1): 379–433.

Cascio, Elizabeth U, and Shenhav Na'ama. 2020. "A Century of the American Women Voter: Sex Gaps in Political Participation, Preferences, and Partisanship Since Women's Enfranchisement." NBER WP 26709.

Cullen, Julie Berry, Nicholas Turner, and Ebonya Washington. 2018. "Political Alignment, Attitudes Toward Government, and Tax Evasion." Working Paper.

de Mouzon, Olivier, Thibault Laurent, Michel Le Breton, and Dominique Lepelley. 2018. "The theoretical Shapley–Shubik probability of an election inversion in a toy symmetric version of the US presidential electoral system." *Social Choice and Welfare*, 1–33.

Downs, Anthony. 1957. "An economic theory of political action in a democracy." *Journal of Political Economy*, 65(2): 135–150.

Edwards, George C. 2011. *Why the Electoral College is bad for America.* Yale University Press.

Fujiwara, Thomas, Kyle Meng, and Tom Vogl. 2016. "Habit formation in voting: Evidence from rainy elections." *American Economic Journal: Applied Economics*, 8(4): 160–88.

Gaines, Brian J. 2001. "Popular myths about popular vote-electoral college splits." *PS: Political Science and Politics*, 34(1): 71–75.

Gallup. 2011. "Americans Would Swap Electoral College for Popular Vote."

Gallup. 2016. "Gallup Vault: Rejecting the Electoral College."

Gallup. 2019. "Americans Split on Proposals for Popular Vote."

Garand, James C, and T Wayne Parent. 1991. "Representation, swing, and bias in US presidential elections, 1872-1988." *American Journal of Political Science*, 1011–1031.

Gelman, Andrew, and Gary King. 1990. "Estimating the electoral consequences of legislative redistricting." *Journal of the American Statistical Association*, 85(410): 274–282.

Gelman, Andrew, and Gary King. 1994. "A Unified Method of Evaluating Electoral Systems and Redistricting Plans." *American Journal of Political Science*, 38(2): 514–554.

Gelman, Andrew, and Pierre-Antoine Kremp. 2016. "The Electoral College magnifies the power of white voters." *Vox*, December 16.

Gelman, Andrew, Nate Silver, and Aaron Edlin. 2012. "What is the probability your vote will make a difference?" *Economic Inquiry*, 50(2): 321–326.

Gentzkow, Matthew, Nathan Petek, Jesse M Shapiro, and Michael Sinkinson. 2015. "Do newspapers serve the state? Incumbent party influence on the US press, 1869–1928." *Journal of the European Economic Association*, 13(1): 29–61.

Gerber, Alan, Mitchell Hoffman, John Morgan, and Collin Raymond. 2019. "One in a Million: Field Experiments on Perceived Closeness of the Election and Voter Turnout." *Forthcoming, American Economic Journal: Applied Economics*.

Kaniovski, Serguei, and Alexander Zaigraev. 2018. "The probability of majority inversion in a two-stage voting system with three states." *Theory and Decision*, 84(4): 525–546.

Katz, Jonathan N, Andrew Gelman, and Gary King. 2004. "Empirically evaluating the electoral college." In *Rethinking the Vote: The Politics and Prospects of American Election Reform*. , ed. Jon Krosnick, JM Miller, MP Tichy, AN Crigler, MR Just and EJ McCaffery. Oxford University Press.

Katz, Jonathan N, Gary King, and Elizabeth Rosenblatt. 2018. "Theoretical Foundations and Empirical Evaluations of Partisan Fairness in District-Based Democracies." Caltech working paper.

Kikuchi, Kazuya. 2017. "The likelihood of majority inversion in an indirect voting system." *SSRN*.

Kuziemko, Ilyana, and Ebonya Washington. 2018. "Why did the Democrats lose the south? Bringing new data to an old debate." *American Economic Review*, 108(10): 2830–67.

Leip, Dave. 2018. "David Leip's atlas of U.S. Presidential Elections, Datasets."

Lepelley, Dominique, Vincent R Merlin, Jean-Louis Rouet, Laurent Vidu, et al. 2014. "Referendum paradox in a federal union with unequal populations: the three state case." *Economics Bulletin*, 34(4): 2201–2207.

May, Kenneth. 1948. "Probabilities of certain election results." *The American Mathematical Monthly*, 55(4): 203–209.

Merrill, Samuel. 1978. "Empirical estimates for the likelihood of a divided verdict in a presidental election." *Public Choice*, 33(2): 127–133.

Miller, Nicholas R. 2012. "Election inversions by the US Electoral College." In *Electoral Systems*. 93–127. Springer.

Peirce, Neal R, and Lawrence D Longley. 1968. *The people's President: the electoral college in American history and the direct-vote alternative.* Simon and Schuster New York.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek. 2018. "IPUMS USA: Version 8.0 [dataset]."

Silver, Nate. 2016. "A User's Guide To FiveThirtyEight's 2016 General Election Forecast." *FiveThirtyEight.com*. https://fivethirtyeight.com/features/a-users-guide-to-fivethirtyeights-2016-general-election-forecast/. Accessed: 2019-6-13.

Sterling, Carleton W. 1978. "The Electoral College biases revealed: the conventional wisdom and game theory models notwithstanding." *Western Political Quarterly*, 31(2): 159–177.

Strömberg, David. 2008. "How the Electoral College influences campaigns and policy: the probability of being Florida." *American Economic Review*, 98(3): 769–807.

The Economist, Andrew Gelman, and Merlin Heidemanns. 2020. "Forecasting the US Elections." *Economist.com*. https://projects.economist.com/us-2020-forecast/president/how-this-works. Accessed: 2020-9-16.

Vogl, Tom S. 2014. "Race and the politics of close elections." *Journal of Public Economics*, 109: 101–113.

ONLINE APPENDIX for "Inversions in US Presidential Elections: 1836-2016"

by Geruso, Spears, and Talesara

A Background

A.1 The Party Systems in the 19th and 20th Centuries

Our earliest sampling frame consists of Antebellum elections from 1836 to 1852. This range includes all years in which Democrats and Whigs were the predominant political parties in national politics. Political scientists typically classify the range 1828 to 1854 as the Second Party System and consider 1852 to be the last presidential election year prior to the Civil War in which the parties were stable. The 1832 election does not easily fit with our two-major-party procedure, nor does it fit with the rest of the Antebellum period, as the presidential candidates earned Electoral College votes from parties other than the Whigs and Democrats (National Republican, Nullifier, Anti-Masonic). We start in 1836, when the major parties were Whigs and Democrats and after all states (other than South Carolina) began allowing their citizens to vote in presidential elections.

Our second sampling frame consists of the post-Reconstruction Era, 1872-1888. Political scientists typically classify the range 1854 to 1892 as the Third Party System. We drop the Civil War years and elections before 1872 as these were characterized by a multiplicity of competing parties that earned EC ballots as well as Republican landslide victories. There was also a changing roster of states gradually rejoining the union in this postwar period. We end our post-Reconstruction sample period at 1892 because the next election in 1896 represented a major political realignment. The realignment is typically recognized as the end of the Third Party System and the beginning of the Fourth Party System.

Other time periods are less useful in providing identifying variation in electoral outcomes. For example, the period 1896 to 1932—the Fourth Party System—mostly yielded presidential landslide victories. These create less useful variation for the purposes of understanding inversion probabilities in close elections. We nonetheless present results for the early twentieth century in Appendix Section E.5 and Figure A2.

A.2 Related Systems Around the World

Inversion are possible in the US because of the Electoral College's tiered system of voting, in which citizens cast votes for electors, who in turn elect the President. Even absent the possibility of faithless electors, the national popular vote (NPV) and the EC outcome can diverge for a host of reasons that we detail in Section 5, where we examine the aggregation mechanics of the EC.

A useful way to understand EC-NPV mismatch is that it can occur when electoral ballots at the second tier can be captured by different numbers of citizen votes at the first tier. In the US, states are heterogeneous in EC representation, which results in elector ballots cast that are disproportionate to citizen votes. Further, that heterogeneity can be (and often is) correlated with partisan alignment across states. In Westminster-style Parliamentary systems, voting units that elect the Prime Minister (Parliamentary districts) are typically similar to each other in representation because they tend to be similarly-sized. All else equal, this tends towards fewer inversions. But some countries, including India and Norway, intentionally introduce malapportionment when electing MPs, who then serve as electors in the second tier that elects the Prime Minister. For example, Norway's system upweights rural counties and India weights regional votes using the regional populations of 1971, which were very different from the geographic distribution of population today.

Even without intentional malapportionment, inversions still occur. For example, Canada's 1979 federal election resulted in a loss for the Liberals and for their incumbent Prime Minister Pierre Trudeau: Despite capturing more citizen votes than the Progressive Conservatives (L: 40% v. PC: 36%), the Liberals were elected to fewer seats in parliament (L: 114 v. PC: 136) and Trudeau was therefore defeated.[40]

B Further Discussion of the Related Literature

Because of the considerable importance of the EC to US politics—as well as the importance of multi-tiered elections to democratic systems worldwide—the EC has received extended attention in the literature. However, because no prior study has investigated the same question we ask here, which is about the fundamental statistical nature of EC inversions, no prior paper has used the same materials and methods. Here we detail how our approach is distinguished from prior literature that: (i) studies empirical facts about the EC other than about inversions in the EC (B.1) or (ii) studies properties of inversions other than their conditional and unconditional probabilities across historical periods (B.4).

In one striking example of the richness of the EC literature, political scientists and historians have even debated which elections should count as an inversion—a debate that is possible because of the complexity, and therefore ambiguity, of the implementation of the EC in practice across states, parties, and centuries (Kallina, 1985; Rakove, 2004; Estes, 2011). Gaines (2001), for example, argues that the 1960 election should be counted as an inversion because over 175,000 popular votes in Alabama (a number in excess of Kennedy's national popular vote margin of victory) were for Democratic electors who were opposed to Kennedy.

A more recent literature considers potential advantages and disadvantages of a national popular vote compact (DeWitt and Schwartz, 2016; Koza, 2016; de Mouzon et al., 2019). Because these studies often either take a normative or legal focus or do not use empirical data, and because they consider aspects of presidential elections other than the probability of inversion (such as the probability or difficulty of a recount, or incentives for strategic voting), we do not consider them further here.

B.1 Empirical facts about the EC, but not about inversions

One of the oldest empirical literatures about the EC documents empirical facts about the distribution of electoral votes across states. In particular, much of this literature describes the allocation of average electoral influence (in the sense of EC ballots per popular vote or EC ballots per person) across states or across population groups. For example, Blair (1979) computes that, by such metrics, whites have more average voting power than blacks. Warf (2009) maps differences in average voting power across states.

Another category of descriptive analysis computes facts about average "voting power" in a way that is distinct from the mere probability of being pivotal, which is the focus of the next

[40] Miller (2012) further discusses inversions in Westminster-modeled parliamentary systems, including in the United Kingdom, New Zealand, and Canada.

section. Banzhaf III (1968), for example, makes computations that compare the size of each state with its number of electoral votes, in order to compute a state-specific index of voting power. A follow-up literature has considered properties of Banzhaf's index and proposed alternatives (Owen, 1975; Dubey and Shapley, 1979).

B.2 Pivotal voters and the EC

A long literature in political science and economics considers the relative costs and benefits of voting, in particular focusing on the probability of being pivotal in deciding the election (Riker and Ordeshook, 1968; Gelman, Katz and Tuerlinckx, 2002). Several papers have applied these ideas to the Electoral College context, including Gelman, King and Boscardin (1998); Gelman, Silver and Edlin (2012); Miller (2013). Our paper is not concerned with the probability that a voter, or a voter in a particular state, or a voter in an election of particular closeness, will be pivotal.

B.3 Inversion analyses using "uniform partisan swing" method

An important feature of our analysis is the modeling of election uncertainty. This distinguishes our work from the many studies in political science that characterize presidential elections deterministically, such as via "uniform partisan swing analysis," and therefore cannot assess the *probability* of an inversion (e.g., Garand and Parent, 1991; Grofman, Koetzle and Brunell, 1997).

Uniform partisan swing analysis—originating in Butler (1951) and Gudgin and Taylor (1979)—has become a standard tool for understanding the relationship between electorate votes and election outcomes, such as congressional seats. The method takes an observed election outcome and, in the classic application, "swings" all legislative districts by the same common vote share. By varying the vote share in a deterministic way in small increments, the method can trace when seats flip and so can trace the relationship between swings in the common, across-district component of votes and the aggregated election outcomes. Primarily applied to estimating seats-votes curves in legislative elections such as for the US Congress (e.g., in Gelman and King, 1990), the method has been ported to analyzing EC. In particular, several studies map the relationship between electorate votes and EC ballots (Garand and Parent, 1991; Miller, 2012). The important differentiator of our study is the incorporation of uncertainty. In uniform swing analysis, there is no probability distribution over the aggregate vote share. In addition, there is no uncertainty in the way that contests across states (or legislative districts) resolve differently. They are assumed to comove perfectly. Therefore, these studies—which do not estimate probability distributions—do not address the goals of our paper, which are the computation of a set of important conditional and unconditional probabilities.

B.4 Inversions: Theoretical computations and election-specific predictions

Our paper uses data from many elections in the 19th through 21st centuries to estimate the unconditional and conditional probability of an inversion, abstracting away from the features of any particular pair of parties or candidates. The wide set of methods that we employ has not previously been applied to this question, and no prior set of estimates of these probabilities exists in the literature.

One of the richest existing literatures about the EC, from the game theory and formal political science literatures, theoretically computes the probability of inversions in mathematical models that abstract away from any data about the actual EC (Kikuchi, 2017; de Mouzon et al., 2018). Many of these papers, like ours, are focused on the stochastic properties of electoral systems. But unlike ours they are not grounded in voting data—for example, how partisan alignment and voting patterns in New York differ from those in Texas.

Another set of papers considers the probability of an inversion in one or more particular elections. Here, we have been able to build upon the methods of prior studies focused on single-election predictions or postdictions. For example, our M1 model is structurally analogous to the model that Katz, Gelman and King (2004) apply to specific years,[41] and our M2 model is similar to the model that Silver (2016) used to predict the distribution of potential outcomes prior to the 2016 presidential election. Some papers in this election-specific literature consider counterfactual policies, as in our Figure 3, but without a probabilistic approach. Cervas and Grofman (2019), for example, apply a set of counterfactuals to determine whether they would have yielded an inversion in several actual historical inversions, assuming that vote totals were the same as what historically occurred.

Among the literature that considers the statistical properties of presidential elections in particular time periods, two of the papers closest to ours in methodology are Merrill (1978) and Ball and Leuthold (1991), which are in dialogue with one another. Neither paper computes or discusses the probability of a close election, which plays a central role in our analysis. The interpretation of these papers is somewhat limited by the details and specificity of their modeling choices. Their sample selection differs from ours and from one another: Merrill, in the mathematics literature, pools elections from 1900 to 1976 (which ignores the mid-20th-century partisan realignment, and therefore ignores the fact that a vote for a Democrat in the time of Wilson had different economic, geographic, and racial correlates than a vote for a Democrat in the time of Carter); Ball and Leuthold (1991) (like Katz, Gelman and King, 2004) compute statistics for each of a series of years from 1920 to 1984, but also pool problematically across distinct periods of partisan realignment (e.g., their 1984 estimates pool data from 1944 to 1984). Methodologically, each paper makes analytic computations, assuming a single parametric form which specifies that each state shares the same distribution: a symmetric normal distribution in the case of Merrill (1978) and a parameterized beta distribution in the case of Ball and Leuthold. Neither paper explores robustness to these assumptions—Ball and Leuthold suggest that a non-parametric approach, such as we use in M3 and M4, would be "difficult to conceptualize." Despite these limitations, these papers are important for their early anticipation that state-indexed models could be used to describe statistical properties of presidential elections.

Finally, Bakthavachalam and Fuentes (2017) in a short note report results on inversion probabilities for a period overlapping with our Modern period. Similar to our results, they conclude that inversion probabilities are high in close elections. In contrast with our results, they conclude that there is no partisan asymmetry. The note does not provide enough technical detail to compare and contrast methods or findings in depth.

[41] Thomas et al. (2013) use essentially the same model as Katz, Gelman and King (2004) to estimate partisan bias in the EC in 14 specific elections but do not estimate the probability of inversions. Partisan bias is indeed important to quantify but is distinct from the probability of an inversion: for example, a two-tiered system that added high-variance, mean-zero noise to election outcomes would generate zero ex ante partisan bias but would yield a high probability of inversion.

C Data and Estimation

C.1 Data

C.1.1 Data Sources

For data on state populations, we use IPUMS extracts from decennial Censuses (Ruggles et al., 2018). For intercensal election years, we follow the standard practice of exponentially interpolating state populations.

The key inputs to our analysis are the historical election returns by state for each presidential election year. For each presidential election, we assemble data on vote tallies for each candidate in each state, as well as data on EC elector ballots cast for each candidate by the EC delegation from each state. Data on state-level election returns and on EC ballots cast come from the Leip (2018) compilation of state returns. We check these data against Federal Election Commission records. In the few state $\times$ election year instances of disagreement, we rely on state government election records, where available. Further details on election data cleaning are documented below.

We use state $\times$ year data on education and race in some models for the Modern period; these are from published summary tables of the American Community Survey. Race data by state from the 19th century come from published Census reports.

C.1.2 Third Parties

For most simulations, we retain information on only the two major parties—Democrats and Whigs from 1836 to 1852 and Democrats and Republicans for the later periods we examine. This normalization, which is standard in the literature (see, e.g., Gelman and King, 1994) does not substantively impact our analysis of inversion probabilities, as third-party candidates won no EC ballots over our study periods.[42,43] When we scale popular vote outcomes by turnout, we include third-party voters in our measure of total state-level turnout.

Of course, a third-party candidate could be pivotal in determining which major party candidate wins a state $\times$ year. The building blocks of our estimation and Monte Carlo exercise are actual state $\times$ year election outcomes. We primarily take these outcomes as basic data and make no assumptions on how a state return might have differed if not for a third-party candidate. Thus, most of our statistics describe a typical election outcome over our sample period, rather than elections in which we counterfactually remove or change the influence of third parties. However, in Figure A6 we assess sensitivity to two extreme and opposite assumptions on the impact of third parties. First we reestimate our baseline model reassigning all third-party votes in each state $\times$ year to the Democratic candidate. Then we reestimate our baseline model reassigning all third-party votes in each state $\times$ year to the Republican candidate.

[42] In particular, one could relabel the horizontal axes in our figures below to center on the state $\times$ year specific threshold, with ticks on the axis indicating distance from that state-election-specific threshold. In the Florida 2000 example, Bush and Gore won 48.847% and 48.838% of votes respectively, with 2.315% going to other candidates. The state $\times$ year specific threshold for a Republican victory in this case would be 0.488425 ($= (1 - 0.02315)/2$).

[43] The last third-party candidate to win a single pledged Electoral College vote was Wallace in 1968, which predates our primary modern sampling frame. The Antebellum and post-Reconstruction elections produced no third-party EC ballots, other than via faithless electors.

C.2 Data Cleaning and Restrictions

Here we catalogue our handling of various special cases and anomalies that arise in the election data:

- We ignore the few historical instances of faithless electors, who cast EC ballots for a candidate other than the candidate to whom they were pledged. In the cases of faithless electors, we award Electoral College ballots as they would have been awarded based on state popular vote results.
- We exclude third parties from our analysis. No third-party candidate won EC ballots in any of the election years we use for our main sampling periods (except through faithless electors). See also Section C.1.2.
- In each election year, we drop states where EC ballots were allocated by state legislatures rather than by the state popular vote. This includes South Carolina in 1836 to 1852 and Colorado in 1876.
- In 1836, the Whig party ran multiple candidates across the country. All states that held a citizen vote for President (as opposed to awarding EC ballots via the state legislature) had one Whig candidate on their ballot, and no states had more than one. We treat all Whig candidates as one candidate in the 1836 election.
- We start the post-Reconstruction era in 1872 because not all of the Confederate states had rejoined the Union by 1868.
- In 1872, Horace Greeley, the Democratic candidate, died after the general election but before electors had formally cast votes. Officially, no EC ballots were allowed for Greeley. We use popular vote data from the general election and award EC ballots as if Greeley had not died.
- In 1872, the electors of Arkansas and Louisiana were not certified by Congress. We use the popular vote outcomes in these states in 1872 to award their EC ballots to Ulysses S. Grant.
- We end the post-Reconstruction era in 1888 because there was a major third party in 1892 (Populists). Additionally, not all states had both major party candidates on their ballots.
- In the Modern period, Maine (since 1972) and Nebraska (since 1992) have split their EC ballots between the state popular vote winner and congressional district popular vote winners. In practice, both states have only split their EC ballots once each. We ignore this rule and allocate Maine and Nebraska's electoral votes by a winner-takes-all rule.

In a few instances, the Leip tallies of individual citizen votes do not align with the Federal Election Commission tallies. We investigate these cases and use states' election commissions to resolve the disputed numbers. The differences are often small, within a mere handful of votes. The cases for which we update Leip's tallies on the basis of state election commission records are:

- 2016 California
- 2016 Minnesota

- 2016 New York
- 2016 Ohio
- 2012 North Dakota
- 2012 Ohio
- 2004 Rhode Island
- 1996 South Carolina
- 1992 Nebraska
- 1988 Louisiana
- 1988 Massachusetts
- 1988 Nebraska

C.3 Sampling Frames

Figure 1 of the main text indicates the periods in US history that we study. Political scientists have identified several stable Party Systems, characterized by competition between a fixed pair of parties with stable political properties.[44] We take these groupings as a starting point for our sample definitions. We further restrict attention to spans of years that include electoral victories for both parties because consecutive landslide victories of a single party do not generate useful variation for our purposes of studying inversion probabilities in close elections. Given these criteria, we study the Second, Third, and Sixth Party Systems, corresponding to the Antebellum, post-Reconstruction, and Modern periods, as indicated in the figure.

Our earliest study period consists of elections between 1836 and 1852. This range includes all years in which Democrats and Whigs were the predominant political parties in national politics and spans through the last presidential election year prior to the Civil War in which the parties were stable. In the post-Reconstruction Era, we study years 1872-1888. Like today, the parties during this period were Republicans and Democrats, though the political alignment of states was rather different. Democrats dominated in the Southeastern US; Republicans dominated in the North, West and Mid-West.[45]

Finally, we treat 1988 to 2016 as our baseline modern period, although model M12 (included in Figures 3 and A1) demonstrates that our estimates of the probability of an inversion conditional on a close election are robust to extending the modern sampling frame further back to the 1960s.

We do not additionally focus on elections between 1900 and 1960 because over this period there was little usable variation for our purposes. With the exception of Woodrow Wilson's terms, Republicans won landslide victories from 1900 to 1928. This was followed by consecutive Democratic landslide victories (four of them by Franklin Roosevelt) beginning in 1932, and then Republican landslide victories again in the 1950s. Sampling or estimating from periods

[44] See Appendix A.1 for further discussion of how our sampling frames align with conventional treatments of the historical US party systems in political science.

[45] Although the Third Party System includes the 1892 election, we exclude it from our analysis as the two major parties were not on the ballot in every state.

of consecutive landslide victories of one party generates landslide counterfactuals, leading to degenerate distributions with little to no probability density around the 50% national popular vote share, which is our threshold of interest. Nonetheless, in Appendix E.5, we show results for the 1916–1956 timeframe for completeness.

D Additional Details on Methods

D.1 Parametric Analysis

Table A1 reports the maximum likelihood estimates of the parameters in Equation 1 of the main text. Estimates in the table are grouped by period. Within each period, the first model (M1) corresponds to the baseline estimate, following the Gelman and King (1994) "unified method of evaluating electoral systems." It includes a national shock and independent state shocks, with state shocks drawn from a common distribution. The parameters of particular interest are the variances of the national and state shocks, σ_{γ}^2 and σ_{ϕ}^2. Either 31, 38, or 51 expected state vote share parameters, $\overline{\alpha}_s$, are also estimated, depending on the data period.[46]

Other columns compute estimates for alternative samples or model restrictions. Column 2 (M2) estimates additional state covariance terms on the basis of geographic region, race, and education. States within a region receive a common, independent shock. The race term multiplies the fraction of each state that is nonwhite by a random, common coefficient drawn from a mean-zero t distribution. The education term multiples the fraction of each state's adult population that is college-educated by a random, common coefficient drawn from a mean-zero t distribution. Data on these demographics come from published, state-level summary statistics of the American Community Survey. The M2 model closely follows the Silver (2016) approach to modeling uncertainty in election forecasting.[47] The next column (M5) estimates the model assuming no national shocks, counter to the stylized facts from the elections literature about the importance of a common, national component to the uncertainty. Columns 7 through 9 add race and education covariance terms singly and together. M10 drops from the sampling frame the two historical instances of inversions in 2000 and 2016. The last model in the modern period (M12) extends the sample to 1964, which walks the data period backward to the partisan realignment of the North and South in the early 1960s (Kuziemko and Washington, 2018). Additional columns repeat these estimates for the Antebellum and post-Reconstruction eras. Model R10 drops the inversion instances (1876 and 1888) from the post-Reconstruction period; there were no inversions in the Antebellum period. The requisite data for estimating the demographic covariance terms (models M7, M8, M9) exist only for the Modern period.

In order to convert state vote shares, V_{st}, into a national popular vote tally, it is necessary to scale V_{st} by voter turnout. Although the national level of turnout is irrelevant to our statistics of

[46] Depending on the data period, some states were not present for all election years within the sample frame or did not use a statewide citizen vote to determine EC votes. Either 25, 37, or 51 expected state vote share parameters, $\overline{\alpha}_s$, are estimated by joint maximum likelihood. For other states, including Colorado in the post-Reconstruction frame (which was not a state in 1872 and which did not hold a popular vote in 1876), $\overline{\alpha}_s$ parameters are estimated separately as means and do not contribute to estimating variance parameters.

[47] FiveThirtyEight's probability distributions over elections account for three potential types of error and uncertainty, relative to the best mean predicted vote share in each state: a common national error, a set of demographic and regional errors, and independent state-specific errors. For demographic and regional errors: "The following characteristics are considered in the simulations: religion (Catholic, mainline Protestant, evangelical, Mormon, other, none); race (white, black, Hispanic, Asian, other); region (Northeast, South, Midwest, West); party (Democrat, Republican, independent); and education (college graduate or not)." (Silver, 2016)

interest, the *relative* turnout across states could meaningfully impact simulated election results. Additionally, EC representation changed within each sample period. For example, Florida had 21 EC ballots in 1988 and 29 in 2016. Unless otherwise noted, we use the actual turnout and EC apportionment from the last election of each sample period. For example, we use 2016 turnout for the modern period. In practice, choices around which numbers to use for turnout and EC apportionment have little bearing on our results. In Figure A7, we rerun the M1 model eight times, in each case assigning different turnout and EC apportionment. Lines in the overlay plot correspond to setting turnout and EC representation to 1988, 1992, 1996, 2000, 2004, 2008, 2012, and 2016. Across these specifications, the probability of an inversion conditional on a margin of victory within one percentage point varies only slightly—from 41% to 42%.

D.2 Bootstrap Monte Carlo: Turnout and the Varying State Roster

One practical consideration that arises when sampling and combining state election outcomes from different years is that the raw vote counts of later years tend to be larger, reflecting population growth. This creates a problem when summing citizen votes across states to yield a national popular vote. We address this by scaling the each party's vote tally in a state by that state's turnout in some common reference year before summing across states. Unless otherwise noted, we use turnout from the last election of each sample period. For example, we use 2016 turnout for the modern period. In practice, using alternative reference years for the turnout weights make almost no difference to the simulation results. See Figure A7. A related consideration is that EC apportionment can vary across election years. As with turnout, we assign EC apportionment according to the last election year of the relevant sample period.

Finally, in our earliest historical period, the Union itself was changing: There were 25 states in 1836 but 31 states by 1852. Therefore, when performing bootstrap Monte Carlo simulations in this period, the sampling procedure generates some draws of state $\times$ years for which the state did not exist. These null draws do not contribute to the simulated NPV or to the simulated EC outcome in these simulated elections.

D.3 Bootstrap Monte Carlo: Generating Correlation Between State Outcomes

A downside of the independent sampling in our baseline bootstrap is that the lack of a common election-year component to the variation leads to under-dispersion relative to the actual span of election outcomes. To better capture the fact that national sentiment (or the characteristics of a particular pair of candidates) tends to moves states together in a given election year, we generate a variant in which we sample state outcomes with probability weights that attach extra probability mass to being drawn from the same election year. In particular, for each simulation we first draw a focal year, y^*, uniformly, independently, with replacement. Let M denote the excess probability that the outcome from the y^* election is sampled for each state in a given simulated national election, t. Increasing M increases the within-year, across-state correlation in voting patterns without imposing parametric assumptions on the distribution of the shocks.

This addition to the bootstrap procedure brings the dispersion closer to the actual dispersion of observed elections. For model M4 in Figure A1, we set the excess probability of drawing from the same focal year at 50%. In Figure 3, we vary M in 0.05 steps from 0.15 to 0.50. In our baseline sample, which contains 8 elections from 1988 to 2016, an equal-probability draw would put 0.125 weight on each year.

D.4 Bootstrap Monte Carlo: Wild Pooled Error Sampling

Models M6, R6, A6 create a larger pool of empirical error terms for bootstrap drawing, consisting of all state deviations from their period means over all elections in the period. Each state is first assigned its empirical sample-period mean two-party vote share. Then for each state, there is an independent wild bootstrap draw from this common pool, so that the ϕ_{st} term is drawn identically across states. The "wild" here is in the sense of Cameron, Gelbach and Miller (2008). It refers to multiplying each draw by a random 1 or -1, effectively doubling the sampling frame and imposing symmetry on the empirical distribution.

D.5 Bootstrap Monte Carlo: Swing State and Safe State Correlations

In Section 4.3, we discuss how we include models in Figure 3 that vary the implied correlation structure of the bootstrap procedure, tuning the excess probability that state draws come from the same election in 5% steps from 15% up to 50%. We do this overall, as well as for swing states separately and "safe" states separately. We define swing states following recent convention: Colorado, Florida, Iowa, Michigan, Minnesota, Nevada, New Hampshire, North Carolina, Ohio, Pennsylvania, Virginia, and Wisconsin.[48] If, for example, the 1992 outcome is drawn with excess probability mass for Colorado, then the 1992 outcomes are also drawn with the same excess mass for the other 11 swing states. In this approach, simulated elections are meant to come closer to the true equilibrium processes by which campaigns are making joint decisions on allocating investments across swing states as they anticipate factors like voter responsiveness to advertising and candidate visits. The primary source of variation in this set of simulations is the margin by which reliably red or reliably blue states are won (according to state-specific historical variability).

Conversely, when we sample "safe" states from the same election with excess probability, we primarily vary the state victor and the margin by which swing states and other potentially contestable states are won. Safe states, in the context of this analysis include the top quartile of states (12) in terms of the average Democrat or Republic margin of the victory over the sample period.

E Supplementary Results and Robustness

E.1 Extended results for 25 models

Figure A1 reports key summary statistics from several parametric and bootstrap models, each with different assumptions and constraints on the data-generating process. The first row in the figure corresponds to M1 (Panels A and B from Figure 2). Subsequent rows correspond to alternative models. The middle graphical panel displays the main findings: inversion probabilities conditional on a close race occurring.

The $x1$ and $x2$ models and their variants ($x7$ through $x10$) track the main approaches to modeling election uncertainty from the political science literature and election forecasting professionals, though such models have generally not been applied backwards to the periods of US history we study. The NPV distributions implied from these models' estimates are

[48] *Politico* published a swing state list leading up to the 2016 election that included: Colorado, Florida, Iowa, Michigan, Nevada, New Hampshire, North Carolina, Ohio, Pennsylvania, Virginia and Wisconsin. FiveThirtyEight adds Minnesota to this list to generate a list of "traditional swing states."

wider than in the ad hoc bootstraps of $x3$. This is because the assumed structure in $x1$ and $x2$ allows for a common, national component to deviations from state-level expectations. The vote share distribution is more diffuse in the $x2$ models (which align with Silver, 2016) than in $x1$ (which align with the Gelman and King, 1994 "unified method") because $x2$ incorporates additional correlated shocks linked across states within the same region and across states with similar demographics. The $x2$ set also draws from a fatter-tailed distribution, as described in Section 3.2. The independent sampling inherent in the bootstrap models ($x3$) tends to generate NPV distributions that are under-dispersed relative to historical data. This is true also for the parametric models ($x5$) that include no national common component to the shock. Other models are described elsewhere in the text in detail. Despite these significant differences in the predicted NPV distributions, Figure A1 shows that the conditional probability of an inversion in a close election is stable.

E.2 Asymmetry in Inversion Probabilities

Table A8 reports, conditional on an inversion occurring, which party was likely to have won the EC (and to have lost the popular vote). Table A7 reports the probability that an inversion accounts for the expected wins of each party. All models agree that, for the Modern period, inversion favored Republicans. Across all 12 Modern-period models, the probability that an inversion was won by a Republican ranges from 62% to 93%. For the post-Reconstruction period, Democrats were favored in inversions according to the standard, parametric models—though models based on bootstrap draws disagree. In the Antebellum period, there is no consensus across models as to whether Whigs or Democrats were favored.

In Section 4.4, we describe how the asymmetry that favored Democrats during the post-Reconstruction era arose in large part from the suppression of black votes in the US South, where Democrats were dominant. Here we present the analysis supporting that claim. At the core of our claim is the observation that blacks were numerous in the South and counted towards apportionment but were kept from voting in commensurate numbers. This is a consequence of the widely documented suppression of black rights, including voting rights, in the South at the turn of the twentieth century (e.g., Epperly et al., 2019 and Lichtman and Kazin, 2010).

The left panels of Figure A11 plot citizen votes per EC ballot against the Republican share of the two-party vote total for the post-Reconstruction period. The positive slope indicates that states likely to vote for Democrats (to the left within each plot) tended to have lower turnout-to-EC elector ratios. Voters in such states could control an EC ballot with fewer citizen votes. For comparison, we create analogous plots for the Modern period in the right panels of Figure A11. In the Modern period, Republicans were advantaged in the sense of EC ballots per citizen votes, so the slope is negative, opposite to the post-Reconstruction period. But the sources of imbalance were different over time. In the Modern period, part of the Republican advantage was tied to over-representation of small-population states due to the two senator-linked electors. To show this, in panels C and D we subtract two EC ballots from each state?s apportionment and recompute the plots. This adjustment negates any relationship in the Modern period between the partisan alignment of states and EC representation (Panel D). But the same is not true for the post-Reconstruction period where subtracting two Electors from each state does not neutralize the relationship (Panel C).

Figure A12 completes the picture: Whether examining all states or restricting attention to former Confederate states with large black populations, voter turnout per population and voter turnout per EC ballot were both strongly negatively correlated with the black share of the state population. For example, Panel A shows that a state whose population was 5% black would be expected to have about 20% turnout as a fraction of the (apportionment-relevant) total population. A state whose population was 60% black would be expected to have just 11% turnout. Therefore, states with the largest black populations could control an EC ballot with fewer white votes. Winning low-turnout states, all else equal, helps a party win EC inversions because electoral votes are won with few popular votes.

In the Antebellum period, the disproportionate electoral ballots per citizen voter allocated to the South did not confer a partisan advantage to any party because Whigs and Democrats split the South.

These results illustrate that heterogeneity in turnout across states is an important source of potential inversions. In the post-Reconstruction, this heterogeneity was the result of deliberate policy. But even random variation in turnout—such as due to weather or to differences in the intercensal rate of population growth–could cause inversions, especially in the close elections that we study. This turnout-to-representation ratio is important across the world as well. Notably, India has what has been called a "crisis of representation,"[49] because the apportionment of Parliamentary seats today is based on population counts in the 1971 Census. As population growth has trended differentially in the north and south of India over the last 50 years, the votes (or registered voters or even residents) that can elect a member of Parliament have diverged dramatically across regions of India, leading to skewed representation that favors populations in low-growth states. (In the US context, India's situation would be as if Texas had not gained in representation in the EC or US Congress over the last several decades despite its explosive population growth over that period).

E.3 Robustness to Gridded Parameter Values

With only a few elections per party system, it is impossible to be confident that estimates of the true parameter values underlying the data-generating process are precise. To examine the extent to which our main results could be sensitive to errors in these estimates, we calculate our outcomes of interest under a set of exogenously specified variances and correlations. In these simulations, we take only the state historical means of vote shares as data. Uncertainty around these means is assumed to follow $\gamma_t \sim N(0, \sigma_\gamma)$ and $\phi_{st} \sim N(0, \sigma_\phi)$ as in our baseline models (M1, R1, A1). But here we cycle over a grid of values for σ^2_γ and σ^2_ϕ, rather than relying on estimates.

Figure A4, which is described in the main text, presents results from this procedure. The procedure generates 96 unique, assumed parameter combinations in each period. In addition to iterating over national and state variances, each combination is used while including or omitting a shared shock by geographic region.

In Figure A5, we present supplementary detail for a subset of the assumed parameter combinations in Figure A4. These simulations include state and national shocks. The variance of the national shock increases along the horizontal axis in each panel. The variance of the state shocks are traced in several contour lines in each panel, as indicated. In the panels on the left, we report the probability of close elections within a 2 percentage point margin. In the panels on the right, we report inversion probabilities, conditional on close elections within the same margins.

[49] See Carnegie Endowment for International Peace: "India's Emerging Crisis of Representation" https://carnegieendowment.org/2019/03/14/india-s-emerging-crisis-of-representation-pub-78588

The slopes of contours in Panels A, C and E indicate that the probability of a close election outcome is sensitive in each period to the gridded parameter values. In particular, it is sensitive to the variance of the common, national shock. However, in all cases the inversion probabilities (Panels B, D, and F) remain high. In the Modern period, the probability of an inversion—conditional on a margin less than 2 percent—never drops below 35%, regardless of the parameters exogenously set. The graph thus traces the same lower envelope on inversion probabilities for the Modern period as Figure A1.

We can summarize Figures A1, 3, A4, and A5 as indicating that our finding of high inversion probabilities in close elections is robust to: (i) parametric approaches that vary the assumptions on the data-generating process across those adopted by the political science literature and election forecasting practitioners, (ii) non-parametric bootstrap approaches that include both independent and highly correlated sampling of state outcomes, (iii) approaches that omit from estimation or bootstrap sampling the actual historical instances of electoral inversions, and (iv) searching over a wide grid of potential parameters, including parameters that are likely to be outside of the true parameter space.

E.4 Alternative Parameterizations State Demographic Characteristics

In Section 4.3, we note that models with shocks linked by election year, region, racial composition, and educational characteristics produce similar inversion probabilities to models that assume that state shocks are completely independent, and it is therefore unlikely that smaller tweaks will affect our main findings. Here we demonstrate this.

For the model plotted in Figure A10, we allow for race-linked shocks to multiply an $\mathbf{X}$ vector that includes region indicators, % non-hispanic white, % non-hispanic black, % hispanic, % college degree, and % high school completion in the state. This contrasts with M2, where $\mathbf{X}$ includes only % non-white and % college degree. In the figure we overlay a plot of this more flexible model with M2. The two are statistically indistinguishable in terms of the conditional probability of an inversion they imply (right panel).

E.5 Results for the Fourth and Fifth Party Systems (1916-1956)

Our main analysis samples do not include elections in the first half of the twentieth century, which was characterized by landslide victories for both Democrats and Republicans. For completeness, we estimate inversion probabilities for this time period here. We divide the timeframe according to a standard typology of party systems. We analyze separately elections in the Fourth Party System (1896–1932) and the Fifth Party System (1936–1956). For the Fifth Party System, we do not include 1960, because doing so would add the complication that it would be the *only* election in this span during which Alaska or Hawaii were states. We also omit 1948, when Strom Thurmond received 39 electoral votes for the States' Rights Democratic Party. For the Fourth Party System, we begin in 1916 in order to generate a stable set of states over the sample period, and we drop 1912 and 1924 because a third party won EC ballots in each of these election years.

Figure A2 presents results for the $x1$ and $x2$ class of models over the Fourth and Fifth Party System periods. These models apply the same structural assumptions and estimation procedures used for M1 and M2 in the Modern period (see Figure A1). The characteristic Win(NPV) curves are similar to other periods. Further, the ex ante probabilities of an inversion in a close election are high in these models. An important difference between these results and results from the Antebellum, post-Reconstruction, and Modern periods is that the probability of a close election was much lower, making the unconditional probability of an inversion lower. It is notable that the conditional inversion probabilities in Figure A2—the primary results of interest—are very similar to the corresponding statistics for our main sample periods, even though the other statistical properties of these times are so different.

F Counterfactuals that Account For Endogenous Behavior

In Section 5, we reference a modification to our statistical model that incorporates endogenous responses to counterfactual EC aggregation rules. Here we describe the procedure in detail. The counterfactual rule sets we consider are the same as in Figure 4—minus two ballots, awarding ballots proportionally, or both changes simultaneously.

Our approach begins with identifying how the set of potentially pivotal states (i.e., "swing" or "battleground" states) would change under the counterfactuals. Denote the probability that state s is pivotal in an election with $\hat{Q}_s$, which we sometime refer to as "swingy-ness" below. Whereas a state like CA has essentially no chance of being pivotal under the status quo and present political alignment towards Democratic candidates, a counterfactual in which its 55 elector ballots are distributed in proportion to the state vote tally brings CA into play and opens the possibility that an EC ballot from CA could be decisive in the election. Call the status-quo swingy-ness of a state $\hat{Q}^0_s$, and call the corresponding quantity under the counterfactual $\hat{Q}^{CF}_s$. After estimating $\hat{Q}^{CF}_s$ for each state, we allow for the behavior of campaigns and voters to influence the data-generating process—in a stylized manner precisely described below.

The approach is grounded in the idea that changes to campaign investment, voter attentiveness, and other political inputs influencing the election are likely to track the changes in the set of potentially pivotal states. Beyond the consistency of this approach with the folk wisdom that *only swing states matter* in a presidential contest in terms of investments like campaign spending, our focus on the changing electoral map aligns with the expert consensus. In economics for example, the model in Strömberg (2008) shows that in equilibrium, campaign resources will be spent (symmetrically by both parties) exactly in proportion to the probability that a state is pivotal in swinging the election ($\hat{Q}_s$). Likewise, statisticians and political scientists have long been focused on the empirical, state-specific probability that a vote cast is decisive in swinging the election (e.g., Gelman, Katz and Tuerlinckx, 2002 and Gelman, Silver and Edlin, 2012). The consensus view is that these probabilities are focal for campaigns assessing where to invest in turning out or persuading voters.

Despite wide consensus around $\hat{Q}_s$ as an object of interest, the literature is mixed on exactly how a change in $\hat{Q}_s$ could alter a race. Would an exogenous shock that increased $\hat{Q}_s$ tighten the race in the state? Or cause turnout to climb in the state? Or reduce election-day uncertainty around the expectation, as parties lock in their voters and convert the undecided ahead of election day? The literature offers no singular guidance.

We therefore model three plausible but substantively different types of endogenous responses as states gain or lose battleground status. Our intent is to span the range of plausible endogenous responses with stylized mechanisms, without taking a position on the correct behavioral model, which is unknown to social science. We assume, in turn, either that (i) the margins will tighten in new battleground states, (ii) variances of potential voting outcomes will shrink in new battleground states, or (iii) turnout will increase in new battleground states. Each of these reduced-form adjustments makes no assumption about the exact mechanisms underlying the net effects. For example, Enos and Fowler (2018) show that an aggregate effect of large-scale campaigning in 2012 was to increase voter turnout by several percentage points relative to the counterfactual in the most highly targeted states. Our turnout counterfactual nests that phenomenon. But our reduced-form turnout adjustment would also nest the case in which turnout effects were instead arising from the (correct) perception among voters in newly minted battleground states that they have an increased potential to impact the election outcome. We model the turnout change (for example), not the ultimate cause of it.

F.1 Methods

Separately, for each counterfactual rule set, we perform the following steps:

1. **Determine Q_s (Swingy-ness) Under Counterfactual Rules.** We begin by using the main simulation results to calculate a probability that each state would be pivotal under the counterfactual EC rules. For each state, separately for each of the original 100,000 simulation runs in Model M1, we re-assign 0.5% of statewide votes from the state winner to the state loser. We calculate the fraction of simulated elections in which this reassignment of votes would flip the EC outcome under the counterfactual EC aggregation rules being considered. Even without changing the simulated voting outcomes, different EC rules will produce different probabilities of each state being pivotal. For example, in the status quo, a high $\hat{Q}_s^0$ would require both that the simulated state outcome has the losing candidate within 0.5% and that the national EC margin of victory is no larger than two times the state's EC ballots. But for other rule sets, including proportional allocation of a state's EC ballots, this is not the case. Call the normalized probability that a state is pivotal under current rules $Q_s^0 \left(= \frac{\hat{Q}_s^0}{\sum_{j=1}^{51} \hat{Q}_j^0} \right)$ and the analogous normalized probability that a state is pivotal under counterfactual rules Q_s^{CF}. For each state and each counterfactual CF, we compute the ratio $\Lambda_s^{CF} \equiv Q_s^{CF} / Q_s^0$, which is greater than one if a state gains battleground importance and less than one if it loses relative importance.

2. **Adjust Data-Generating Process.** We next introduce an endogenous response, altering the state mean, variance, or turnout as a function of Λ_s^{CF} calculated in step 1. In three separate counterfactuals, we:

   (a) Shrink the expected partisan alignment in proportion to Λ_s^{CF}. This shifts expectations toward 50/50 in states that become more swingy and away from 50/50 in states that become safer, where the trailing party loses incentive to compete. In these simulations, we adjust the log-odds ratio of a Republican victory at the state level by setting a new state mean: $\alpha_s^{CF} = \alpha_s^0 \cdot (\Lambda_s^{CF})^{-1}$, where α^0 is the original estimate of the log-odds of Republican victory in the state. The shifted α_s^{CF} becomes the new constant in the log-odds-transformed vote share process: $V_s = \alpha_s + \epsilon_s$ from Equation 1. Note that $(\Lambda_s^{CF})^{-1}$ is less than one when a state's battleground importance increases under counterfactual rules. For example, as California goes from being a safe state under the status quo to having some of its elector ballots in play under the proportional rules, the adjustment to α_s^0 will move the expectation of the log-odds ratio $\left(ln\left(\frac{\text{Pr.(R)}}{\text{Pr.(D)}}\right) \right)$

closer to zero. In other words, it moves California closer to a 50/50 vote share.[50] In this way, the race endogenously tightens in newly generated battlegrounds and loosens in newly generated safe states, where $(\Lambda_s^{CF})^{-1}$ is greater than one.

(b) Shrink the variance term in the state's data-generating process in proportion to Λ_s^{CF}. This reduces uncertainty in the new battleground states and increases it in states that become safer. In these simulations, we set $\sigma_s^{CF} = \sigma_s^0 \cdot \Lambda_s^{CF}$. Recall that σ_s^2 is the variance of the state uncertainty term.

(c) Inflate voter turnout in proportion to changes in $\hat{Q}_s^{CF}$. This increases turnout in the new battleground states and reduces it in states that become safer. In these simulations, we multiply the turnout that would otherwise occur in a state by $\left(1 + 0.1 \times \frac{\hat{Q}_s^{CF} - \hat{Q}_s^0}{\max_r \left(\hat{Q}_r^{CF} - \hat{Q}_r^0\right)}\right)$, where the denominator is the maximal difference across states for a specified counterfactual set of rules, so that turnout increases by 10% for the state with the greatest increase in swingy-ness and increases or decreases by other amounts for other states, depending on their relative change in swingy-ness.

In each case above, we bottom-code Q_s at the 33rd centile across the states prior to calculating Λ_s^{CF}. This is to avoid large ratios due to tiny probabilities in the denominator. Intuitively, there is little practical difference in exactly how safe Kansas and Massachusetts are in the Modern period, but noise in small values of Q_s among extremely safe states could explode the ratio Λ_s^{CF}.

3. **Rerun the Simulations Using the Adjusted DGP.** Starting from the adjusted parameters to the data-generating process, we generate 100,000 new simulation draws. From these counterfactual-specific simulations, we calculate inversion rates under each counterfactual rule set. This generates probabilistic voting outcomes that incorporate endogenous responses to the changing electoral map.

As a check on our process, we first calculate the Q_s^0 probabilities (swingy-ness) assuming the status quo EC rules, rather than a counterfactual. The validation check produces a familiar list. In descending order of Q_s^0, the pivotal states over the 1988–2016 period are FL, PA, OH, MI, and VA. We then calculate analogous probabilities under each of the counterfactual rule sets, Q_s^{CF}. We show these plotted against Q_s^0 in Figure A14. Panel A shows that removing two elector ballots from each state changes little in the list of swing and safe states. For example, PA gains in relative importance compared to OH, but the effect is small. In contrast, under the proportional ballot counterfactual in Panel B, the relative importance of PA, OH, MI, and VA fall, while CA, TX, and NY rise to join FL among the most important battleground states.[51] Note that under the counterfactuals that include proportional allocation (Panels B and C), states' probabilities of swinging the election are less differentiated. Everywhere becomes more in-play, so the distribution of Q_s^{CF} (across the vertical axis) is tighter.

An important caveat here is that the exercise is most likely to be informative for counterfactuals that are closest to the kinds of marginal changes to states' battleground importance that exist in the estimation sample. For example, removing the two electors tied to each state's

[50] In the logit equation $V_s = \alpha_s \cdot (\Lambda_s^{CF})^{-1} + \epsilon_s$, and the Republican vote share is $e^{V_s}/(1+e^{V_s})$.

[51] A "battleground" state is conceptually different when EC ballots are awarded proportionally rather than as winner-takes-all. Under proportional allocation, CA, TX, and NY gain in relative importance because swaying half a percent of votes in these states (e.g., by campaigns appealing to state-specific concerns) corresponds to more voters.

Senators is a counterfactual similar in both kind and magnitude to the effects of decennial reapportionment. Consider the 1990 re-apportionment, which our Modern period spans. Following re-apportionment, New York lost three EC ballots; Pennsylvania, Ohio, Illinois, and Michigan each lost two ballots; New Jersey, Massachusetts, Louisiana, Kentucky, Iowa, Montana, Kansas each lost one ballot; while California, Texas, and Florida each gained several EC ballots.[52] In contrast, proportional ballot allocation in place of a winner-takes-all rule at the state level is a counterfactual that is far out of sample.

Further, we view these counterfactuals primarily as directionally informative. The appropriate size of the adjustment—e.g., by just how much does turnout increase?—involves parameters for which there exist no credible estimates to our knowledge. Therefore, while the exercise potentially forms bounds by revealing whether accounting for endogeneity increases or decreases the estimated probabilities of inversion in a counterfactual, we do not claim it to be informative of the precise magnitudes of effects.

F.2 Results

With those caveats, we present results in Figure A15 and Table A9. Panel A repeats results from Figure 4 for reference. In Panel A, there is no adjustment to the data-generating process for possible endogenous responses. The simulations behind Panel B assume that the race tightens in states that increase in battleground importance and loosens in states that become safer. To put the magnitude of the changes this procedure generates in context, when considering the counterfactual of proportional ballot allocation, states in the top quintile of movement towards Republicans move an average of 6.7 percentage points in expectation. States in the top quintile of movement towards Democrats move an average of 7.0 percentage points. Interestingly, comparing Panel B to Panel A indicates that the effect of accounting for the endogenous responses of voters and campaigns in this way is to amplify the resulting partisan imbalances under the counterfactuals and to increase the probability of inversions under the counterfactuals.

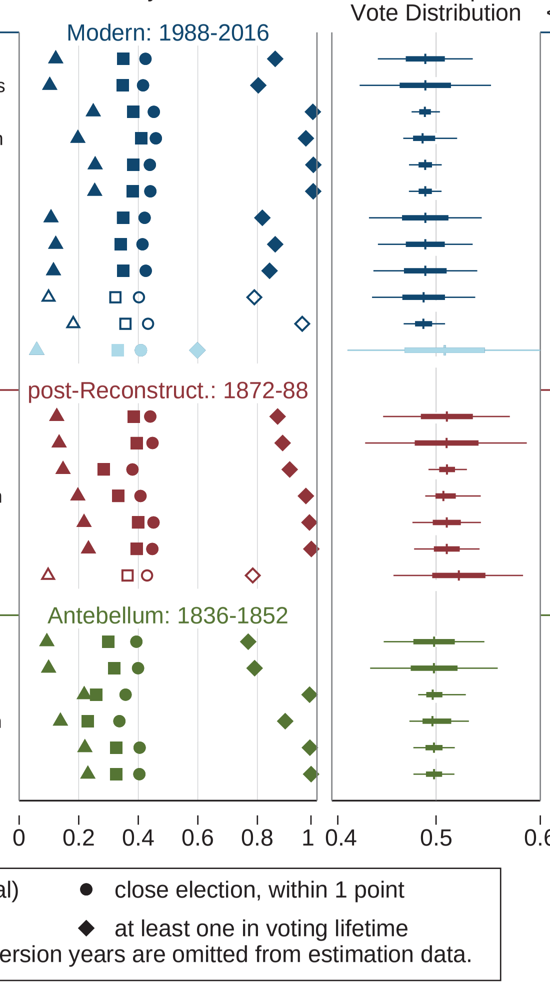
The remaining panels in Figure A15 report the analogous results for the assumptions that gaining battleground status reduces variance (Panel C) or increases turnout (Panel D). Again, the simulated effects of the electoral map shifting are large. In the simulations behind Panel D, the national composition of turnout across states shifts towards the most politically salient states.

The impacts of endogenizing turnout and variance are less stark than the impacts of endogenizing tightness. But it remains the case, as we show in Table A9, that each of these endogenous-response counterfactuals either has almost no impact on the probability of an inversion in a close race or increases it relative to the baseline assumption of no endogenous response. On this basis, we conclude that the baseline decompositions in Figure 4, which hold the statistical model fixed when calculating the impact of the policy changes, are likely to represent a lower bound on inversion probabilities under the counterfactuals.

[52] Likewise, inflating the size of the House of Representative to reduce the rounding errors inherent in allocating a small number of US House Districts (today 435 with voting members) across states is similar in kind (though not magnitude, depending on the particular inflation factor) to changes that have occurred in US history: The size of the US House has grown through various Apportionment Acts of Congress.

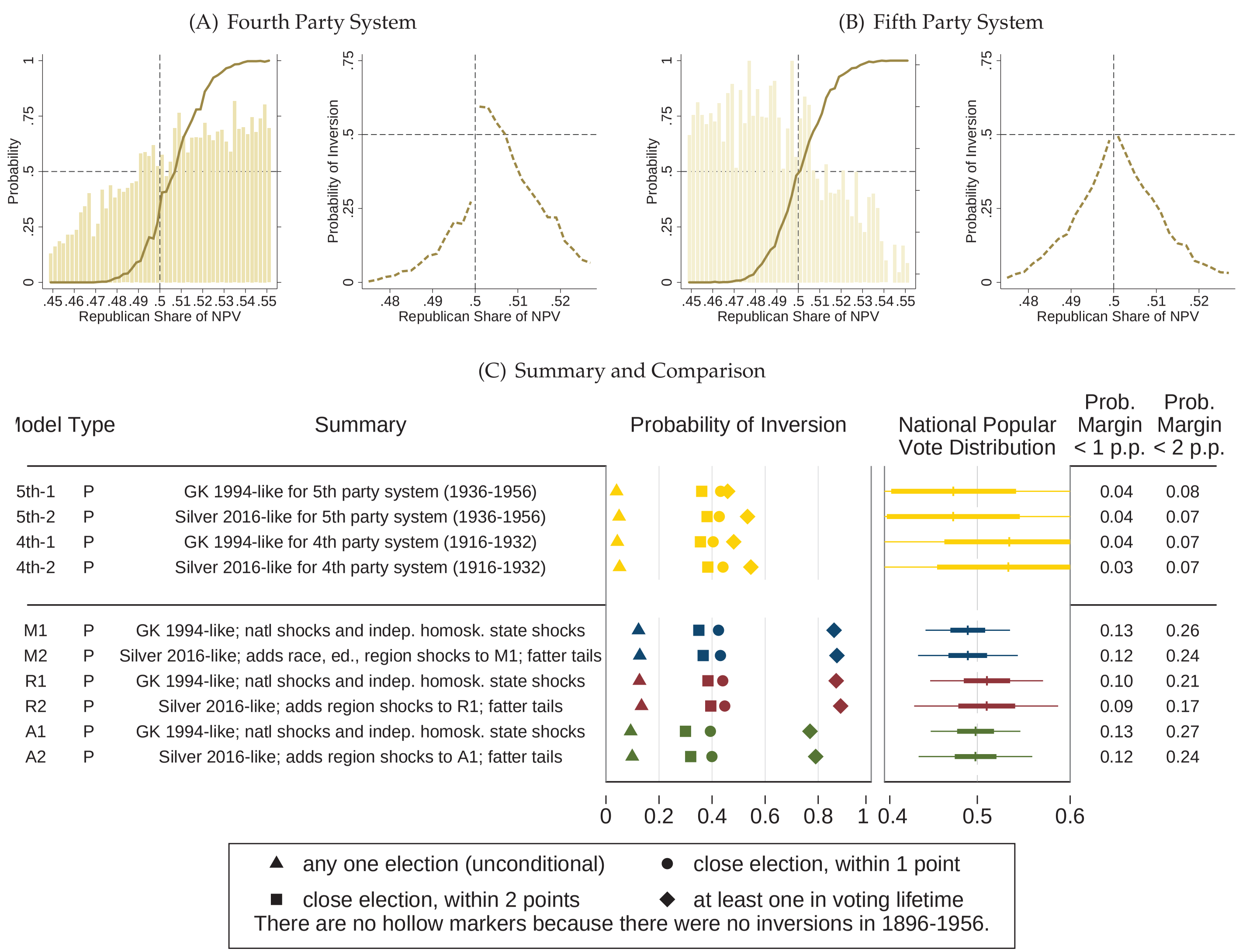
Figure A1: Inversion Probabilities Across Models, Methodologies, and Time Periods

Model	Type	Summary	Prob. Margin < 1 p.p.	Prob. Margin < 2 p.p.
Modern: 1988-2016				
M1	P	GK 1994-like; natl shocks and indep. homosk. state shocks	0.13	0.26
M2	P	Silver 2016-like; adds race, ed., region shocks to M1; fatter tails	0.11	0.21
M3	NP	Bootstrap; indep. draws from state-specific (heterosk.) history	0.22	0.45
M4	NP	M3, with increased prob. (50%) of sampling from same election	0.17	0.33
M5	P	M1 without common natl shocks	0.23	0.47
M6	NP	Bootstrap; wild sampling of pooled (homosk.) state shocks	0.23	0.47
M7	P	M1 adding race shocks	0.12	0.23
M8	P	M1 adding educ shocks	0.13	0.26
M9	P	M1 adding race & educ shocks	0.12	0.25
M10	P	M1 dropping inversions (2000, 2016)	0.12	0.24
M11	NP	M3 dropping inversions (2000, 2016)	0.19	0.39
M12	P	[Extended M1: 1964-2016]	0.07	0.14
post-Reconstruct.: 1872-88				
R1	P	GK 1994-like; natl shocks and indep. homosk. state shocks	0.10	0.21
R2	P	Silver 2016-like; adds region shocks to R1; fatter tails	0.09	0.17
R3	NP	Bootstrap; indep. draws from state-specific (heterosk.) history	0.23	0.46
R4	NP	R3, with increased prob. (50%) of sampling from same election	0.30	0.53
R5	P	R1 without common natl shocks	0.17	0.34
R6	NP	Bootstrap; wild sampling of pooled (homosk.) state shocks	0.18	0.35
R10	P	R1 dropping inversions (1876, 1888)	0.09	0.17
Antebellum: 1836-1852				
A1	P	GK 1994-like; natl shocks and indep. homosk. state shocks	0.13	0.27
A2	P	Silver 2016-like; adds region shocks to A1; fatter tails	0.12	0.24
A3	NP	Bootstrap; indep. draws from state-specific (heterosk.) history	0.31	0.56
A4	NP	A3, with increased prob. (50%) of sampling from same election	0.17	0.37
A5	P	A1 without common natl shocks	0.32	0.58
A6	NP	Bootstrap; wild sampling of pooled (homosk.) state shocks	0.33	0.61



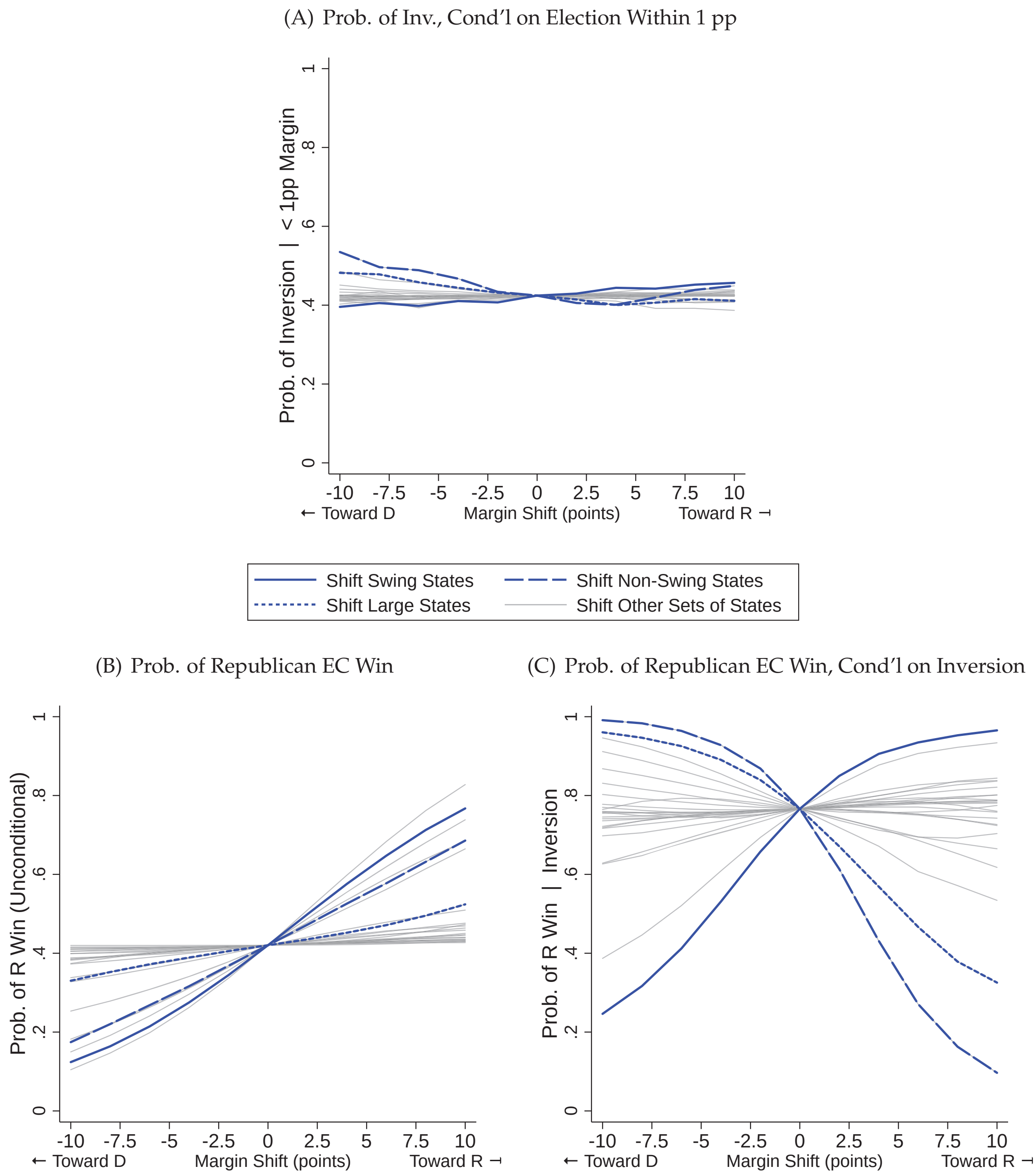
Note: Figure reports summary statistics from 25 election models. Repeated numerals in model names indicate that the same specification is used across time periods—as in M1, R1, A1. The popular vote distribution presents 5th, 25th, 50th, 75th, and 95th centiles of the simulated elections. P/NP denotes parametric/non-parametric. Parametric models are estimated by maximum likelihood; non-parametric (i.e., "bootstrap") models resample past election outcomes for each state to generate a national outcome. P(mar. <1pp) and P(mar. <2pp) report the probability that the popular vote margin is within 1 and 2 percentage points, respectively. "At least one in voting lifetime" indicates the probability of experiencing an inversion for a voter who votes in 15 presidential elections over a 60-year voting lifetime and faces the unconditional distribution of voting outcomes described by the model.

Figure A2: Results for the Fourth and Fifth Party Systems: 1916–1956



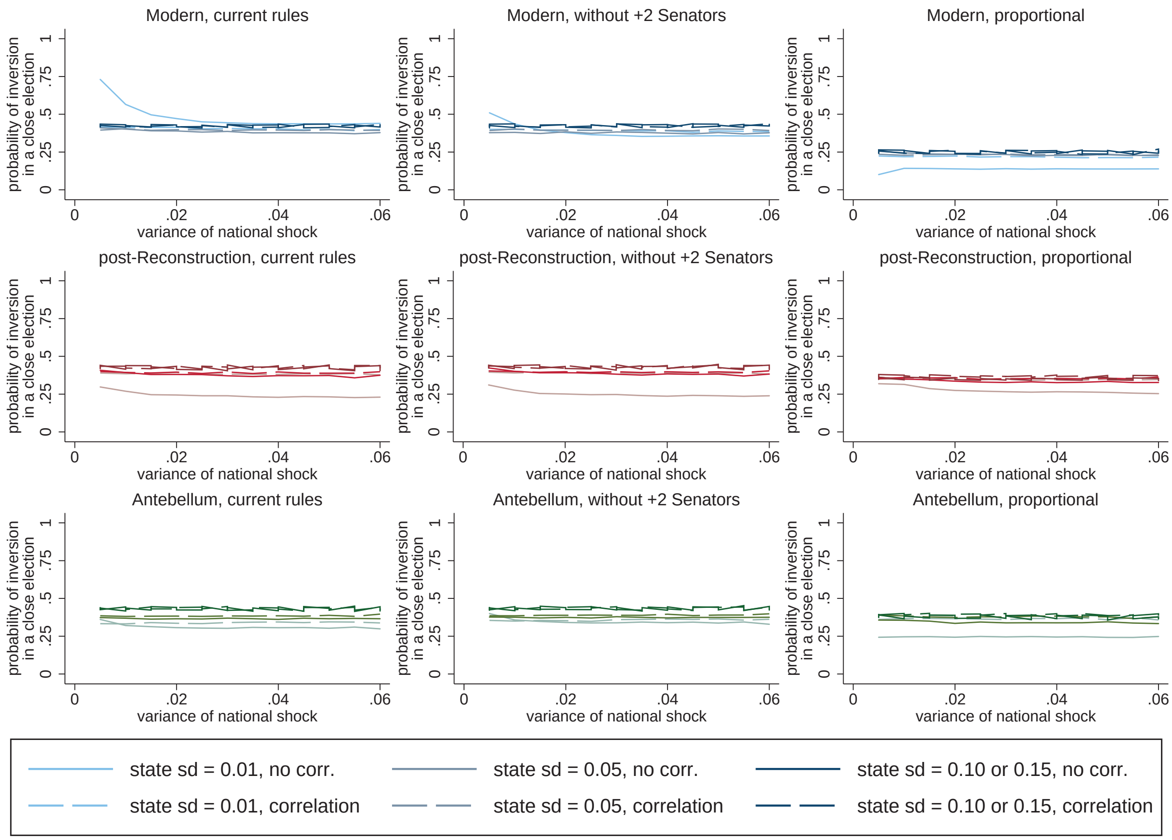
Note: Figure shows inversion probabilities and probability distributions over national popular vote outcomes implied by the parametric estimates of the baseline $x1$ and $x2$ family of models estimated for sample periods in the first half of the twentieth century—the sand-colored regions of the party system history described in Figure 1. Each model simulation consists of 100,000 simulated election draws. The Republican share of the national popular vote runs along the horizontal axis in the top panels. See Figure 2 and Figure A1 for additional notes on Figure elements. See Appendix E.5 for a full description of the exercise.

Figure A3: Alternative Partisan Alignment Scenarios



Note: Figure shows the impacts on key statistics of simulating a range of potential changes to the partisan alignment of states. Each point in each panel corresponds to a partisan shift relative to the $M1$ model estimates for the indicated set of states. Moving left to right across the horizontal axis steps through 10 such counterfactuals, where the partisan alignment of the indicated states is shifted in 2 percentage point increments from a 10 point margin shift toward Republican to a 10 point margin shift toward the Democrat. Panel A shows the impacts of these shifts on the conditional probability of an inversion in an election decided by within one percentage point, $\pi(.01)$. Panel B shows the impacts on the unconditional probability of a Republican EC victory. Panel C shows the impacts on the probability of a Republican EC victory, conditional on an inversion occurring. All corresponding estimates used to generate the figure are displayed in Tables A2, A4, and A3.

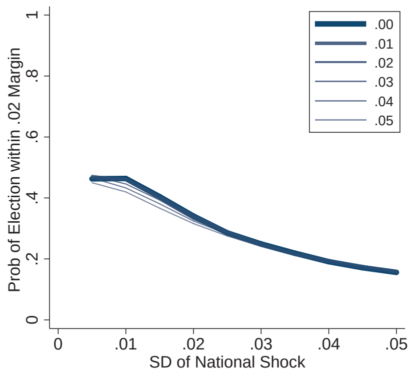
Figure A4: Robustness: Iterating Over a Grid of Exogenously Set Parameters



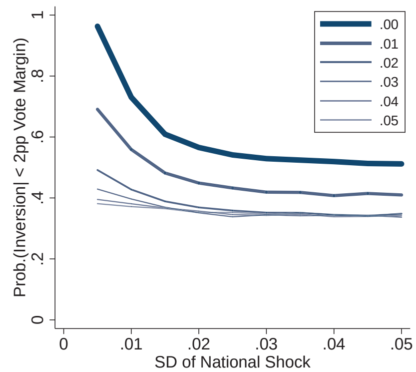
Note: Figure calculates inversion probabilities under a set of exogenously specified variances and correlations. The error terms from the data-generating process in Equation 1 of the main text are assumed to follow $\gamma_t \sim N(0, \sigma_\gamma)$ and $\phi_{st} \sim N(0, \sigma_\phi)$ as in the baseline models (M1, R1, A1). Specifically, these models implement Eq (1) as a random national shock (with magnitude as indicated along the horizontal axes) plus a random state-specific shock (at a standard deviation of 1, 5, 10, or 15 percentage points as indicated), plus (for the models labelled with "correlation") a random regionally correlated shock at a standard deviation of 5 percentage points. Each panel plots the probability of an inversion conditional on a 1.55 percentage point popular vote margin or less (which corresponds to 2 million popular votes at 2016 turnout). Note that these 288 models use only state-specific means from past election data. The variances are exogenously specified as hyperparameters. The second and third columns of plots expand on the decomposition described in Section 5: "Without +2 Senators" allocates each state a number of electors equal to its number of Representatives, without two electors for Senators. "Proportional" divides the whole number of electors per state between parties.

Figure A5: Robustness: Iterating Over a Grid of Exogenously Set Parameters (further detail)

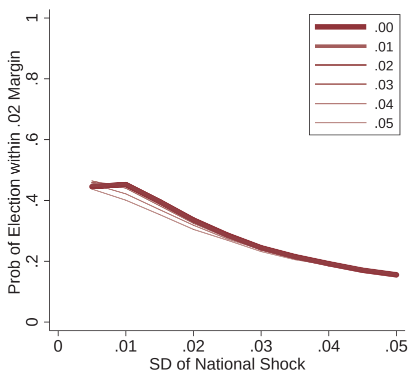
(A) Modern: Prob.(< 2 pp margin)



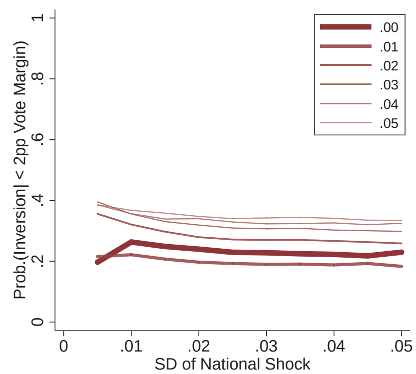
(B) Modern: Prob.(Inv. | < 2 pp margin)



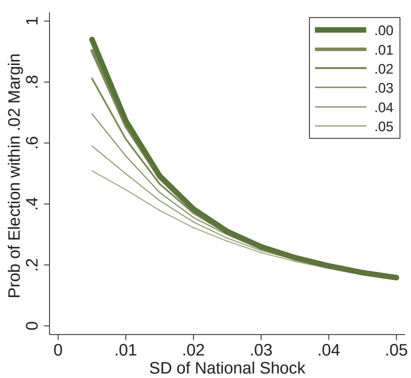
(C) Post-Reconstruction: Prob.(< 2 pp margin)



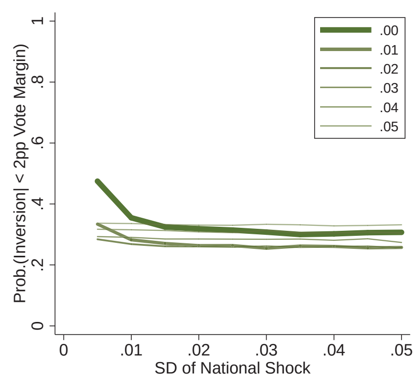
(D) Post-Recon.: Prob.(Inv. | < 2 pp margin)



(E) Antebellum: Prob.(< 2 pp margin)



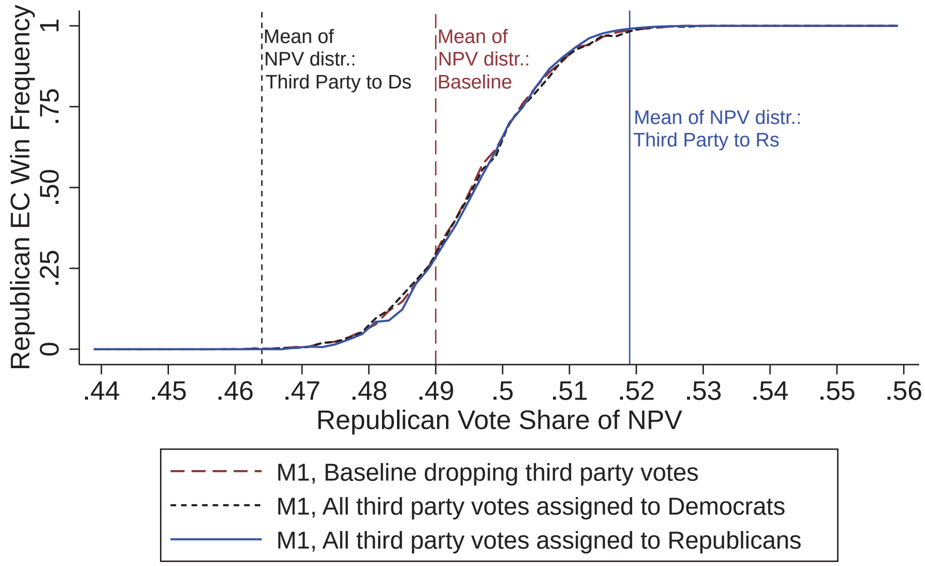
(F) Antebellum: Prob.(Inv. | < 2 pp margin)



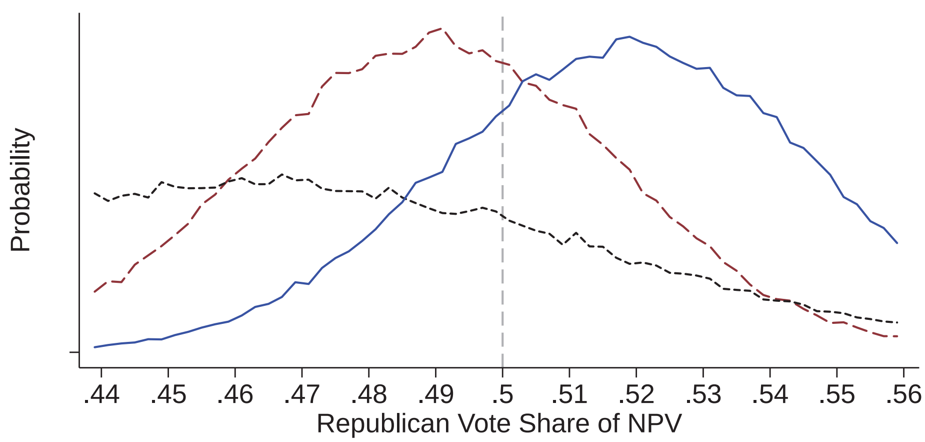
Note: Figure calculates inversion probabilities under a set of exogenously specified variances and correlations. The error terms from the data-generating process in Equation 1 of the main text are assumed to follow $\gamma_t \sim N(0, \sigma_\gamma)$ and $\phi_{st} \sim N(0, \sigma_\phi)$ as in the baseline models (M1, R1, A1). We cycle over a grid of values for σ^2_γ and σ^2_ϕ, rather than relying on estimates. The variance of the national shock increases along the horizontal axis in each panel. The variance of the state shocks are traced in several contour lines in each panel, as indicated. In the panels on the left, we report the probability of close elections. In the panels on the right, we report inversion probabilities, conditional on close elections within the same margins.

Figure A6: Results Are Robust to Any Treatment of Third-Party Votes

(A) Probability of Republican EC Win at Each NPV Under Various Third-Party Treatments

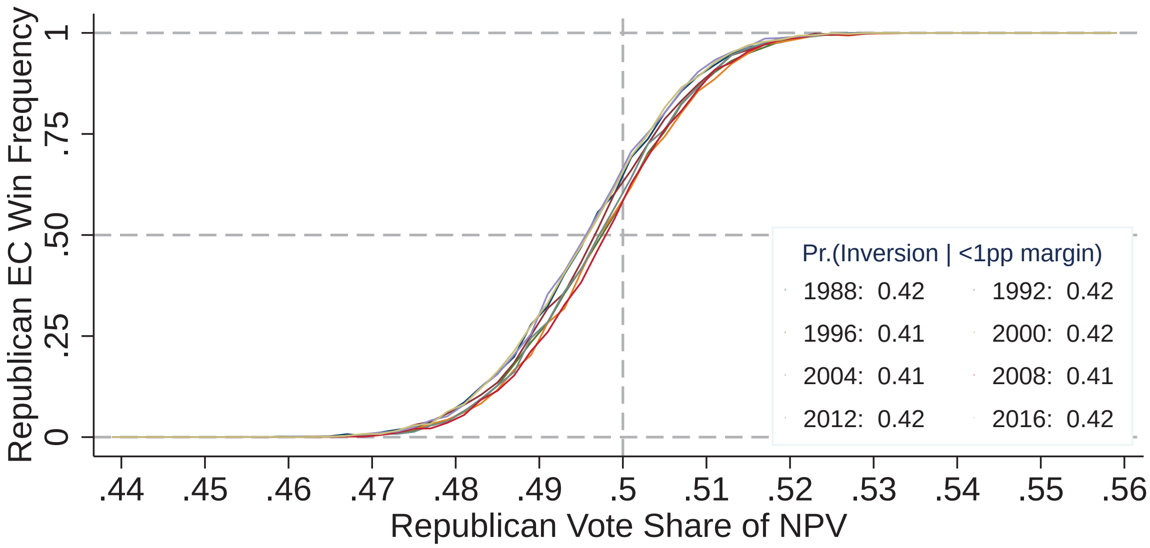


(B) NPV Distribution Under Various Third-Party Treatments



Note: Figure demonstrates robustness of estimates to extreme treatments of third-party votes, plotting the conditional win function for three models that treat third-party votes differently. The results here estimate the M1 model after either assigning all third-party votes to the Democratic candidate or after assigning all third-party votes to the Republican candidate. The original M1 model, which ignores third-party votes, is also plotted for reference. As the vertical lines show, this counterfactual assignment makes a large difference to the central tendency of the distribution of popular votes. However, it does not change the object of interest: the conditional probability of winning the EC as a function of the popular vote outcome.

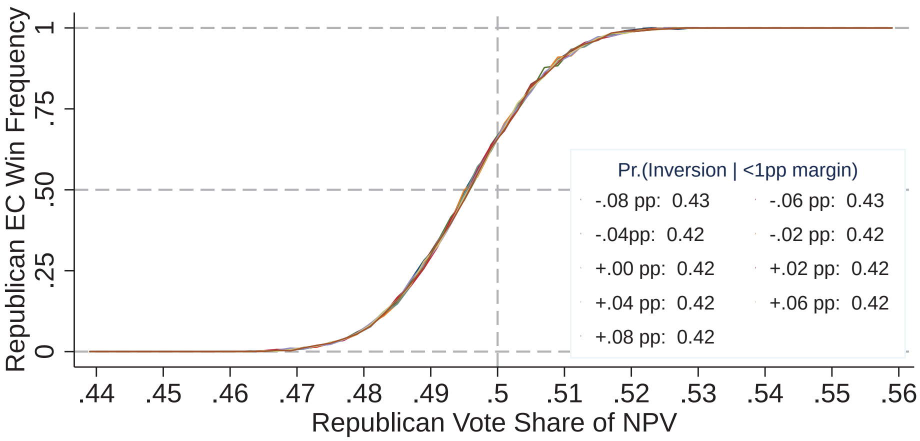
Figure A7: Results Are Robust to Alternative Turnout Weights



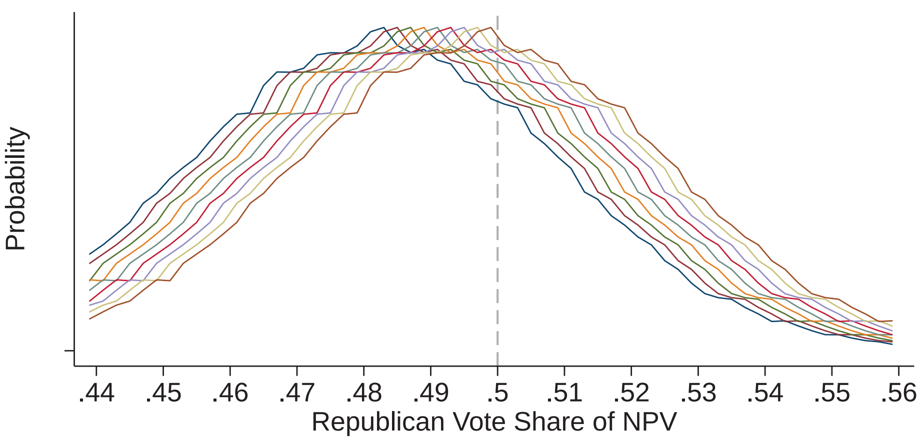
Note: Figure demonstrates robustness of estimates to various assumptions regarding turnout. In order to convert state vote shares, V_{st}, into a national popular vote tally, it is necessary to scale V_{st} by voter turnout. The figure replicates the conditional win function for the M1 model (modern period) eight times, in each case assigning different turnout and EC apportionment when tallying the popular vote and EC ballots across states. Lines in the overlay plot correspond to setting turnout and EC representation to 1988, 1992, 1996, 2000, 2004, 2008, 2012, and 2016. The box lists probabilities of an inversion in each model, conditional on an NPV victory margin within one percentage point.

Figure A8: Results Are Robust to Hybrid Models that Shift Means of Estimated Distributions

(A) Conditional Win Rate Function Under Various Shifts to the Estimated Distribution



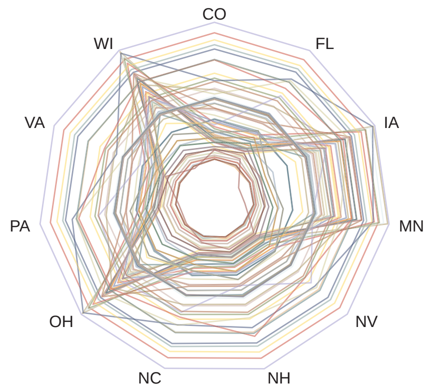
(B) NPV Distribution Under Various Shifts to the Estimated Distribution



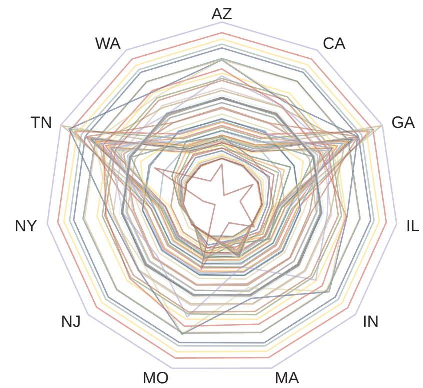
Note: Figure demonstrates robustness of estimates to hybrid models that first estimate parameters for model M1 and then mechanically shift each state mean left or right by a uniform percentage point margin. Lines in the overlay plot correspond to setting the shift at: $\{-.8, -.6 - .4, -.2, 0, +.2, +.4, +.6, +.8\}$. The distribution of the national popular vote is displayed in panel A, and the conditional win function is shown in panel B. The box lists probabilities of an inversion in each model, conditional on an NPV victory margin within one percentage point.

Figure A9: Robustness: The Conditional Probability of Inversion is Invariant to Model & Parameter Uncertainty, Despite That Other Statistics Are Not (Same 109 Models as Figure 3)

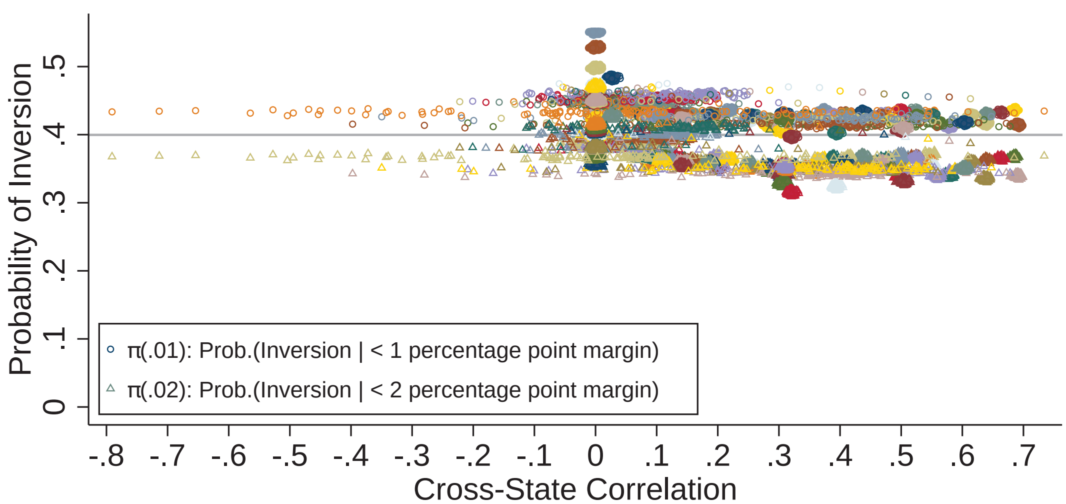
(A) Within-Election Correlation: MI with Other Swing States (Same 109 Models)



(B) Within-Election Correlation: TX with Other Large States (Same 109 Models)

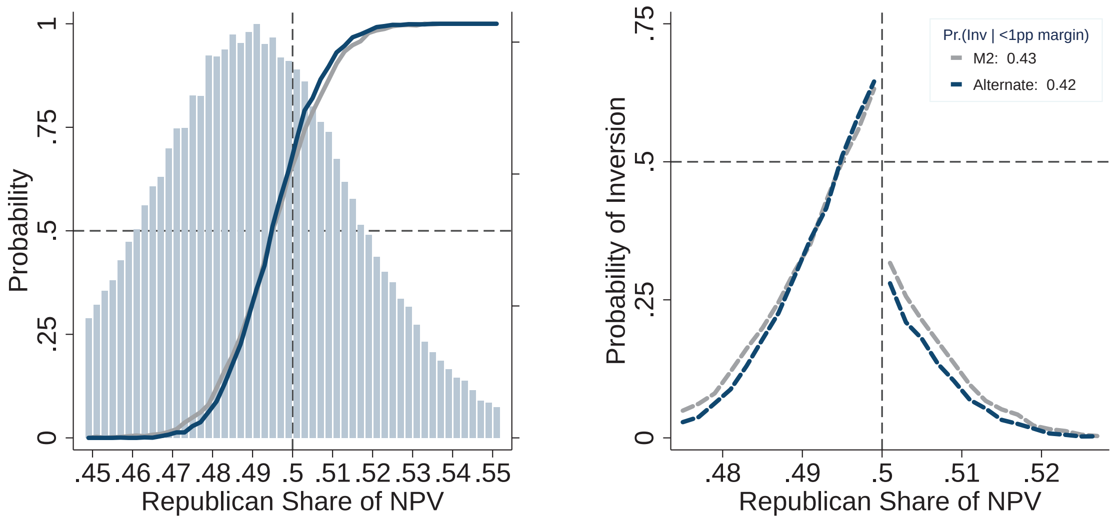


(C) Inversion Probabilities are Independent of the Cross-State Correlation Structure (Same 109 Models)



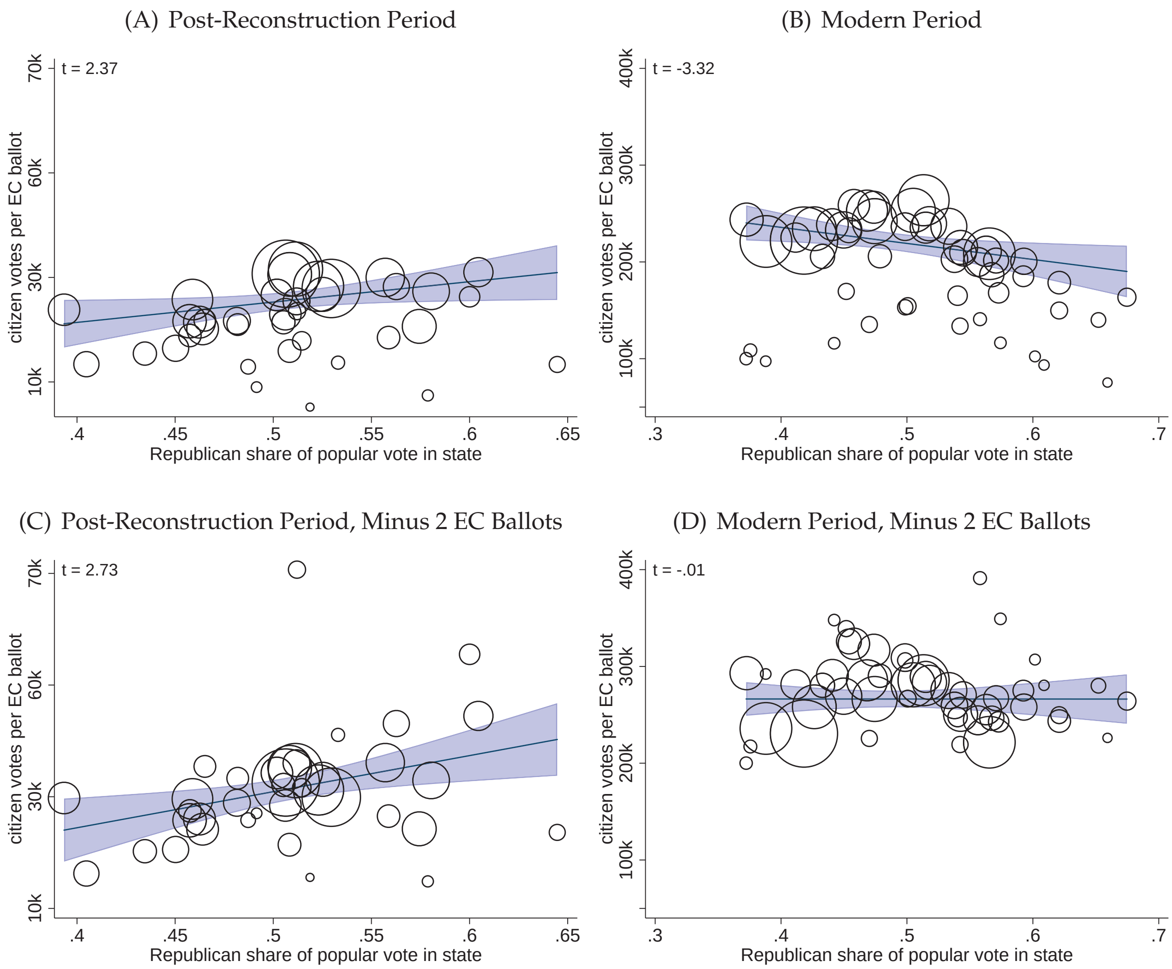
Note: Figure shows additional statistics corresponding to the models displayed in Figure 3. See Table A5 for a detailed listing of each model included. Radar plots in Panels A and B compare the within-year, across-state correlations between Michigan and 11 other swing states and between Texas and 11 other large, non-swing states. Panel C plots inversion rates against the across-state correlations in the voting outcomes for each model. Each model generates 2,550 points in Panel C: 2 statistics per model times the 51 state lower triangular correlation matrix (1275 correlations) for each model. Points in Panel C are jittered.

Figure A10: Robustness to Alternative Parameterization of Race- and Education-Linked Shocks



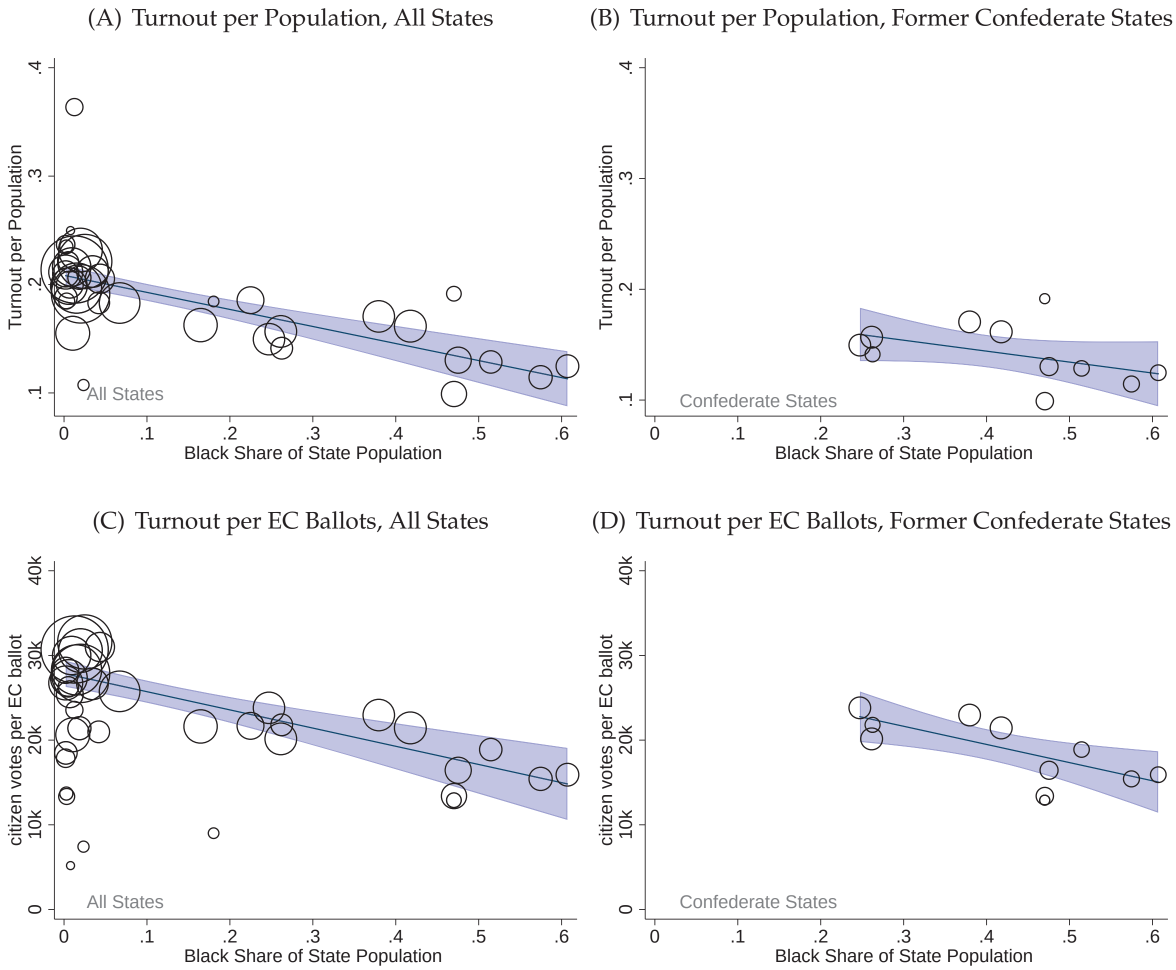
Note: Figure shows inversion probabilities and probability distributions over national popular vote outcomes implied by the parametric estimates of model M2 compared to an alternate model that changes how state characteristics are parameterized in the shock term. The alternate model allows for race-linked shocks to multiply an **X** vector that includes region indicators, % non-hispanic white, % non-hispanic black, % hispanic, % college degree, and % high school completion in the state. This contrasts with M2, where **X** includes only % non-white and % college degree. Each model simulation consists of 100,000 simulated election draws. The M2 model is plotted in gray for reference behind the alternate model in blue. The histogram corresponds to the alternate model. The Republican share of the national popular vote runs along the horizontal axis. The solid blue line is the conditional probability of a Republican electoral win at each level of the national popular vote share. The box lists probabilities of an inversion in each model, conditional on an NPV victory margin within one percentage point. See Figure 2 and Appendix E.4 for additional notes.

Figure A11: Asymmetry in the Post-Reconstruction and Modern Periods



Note: Figure shows the relationship between partisan alignment and EC representation of voters in the post-Reconstruction and Modern eras. Each circle is a state, with size proportional to turnout. Voter turnout and EC electors per state are based on averages over the indicated sample periods. The Republican share (horizontal axes) is the state mean over the indicated period. OLS lines and 95% confidence intervals are displayed. The vertical axes plot the number of citizen votes in a presidential election divided by the number of EC electors apportioned to the state. In the post-Reconstruction period, Democratic alignment was correlated with EC ballots being controlled by fewer citizen votes. See Appendix E.2 for additional detail.

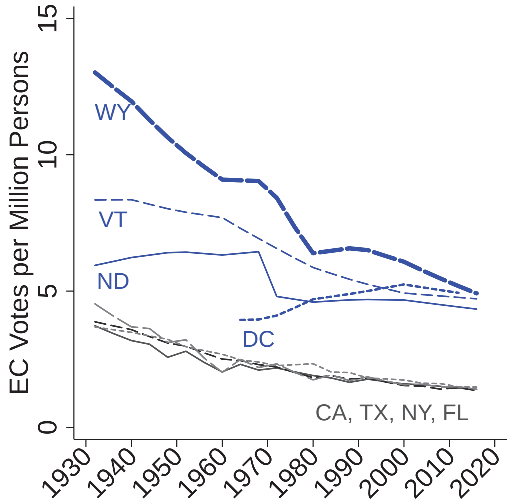
Figure A12: During the Post-Reconstruction Era, Turnout Was Depressed and EC Representation of White Voters Was Inflated Because Black Votes Were Suppressed



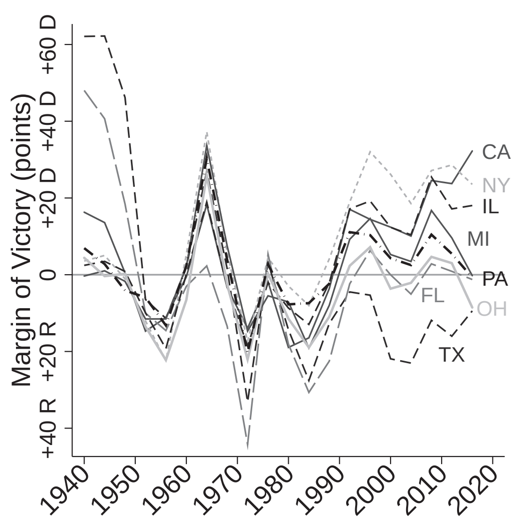
Note: Figure shows that turnout per population was strongly negatively correlated with the black share of the state population in the post-Reconstruction period. States with larger black populations could control an EC ballot with fewer citizen votes. Populations are calculated from the 1880 Census. Voter turnout and EC electors per state are based on averages over the 1872–1888 period. In Panels A and B, the vertical axes plot turnout per total population (white and black). In Panels C and D the vertical axes plot the number of citizen votes in a presidential election divided by the number of EC electors apportioned to the state. Left panels include all states; right panels include former Confederate states. OLS lines and 95% confidence intervals are displayed. See Appendix E.2 for additional detail.

Figure A13: Possible Sources of Inversion in the Electoral College

(A) Small States Disproportionately Represented

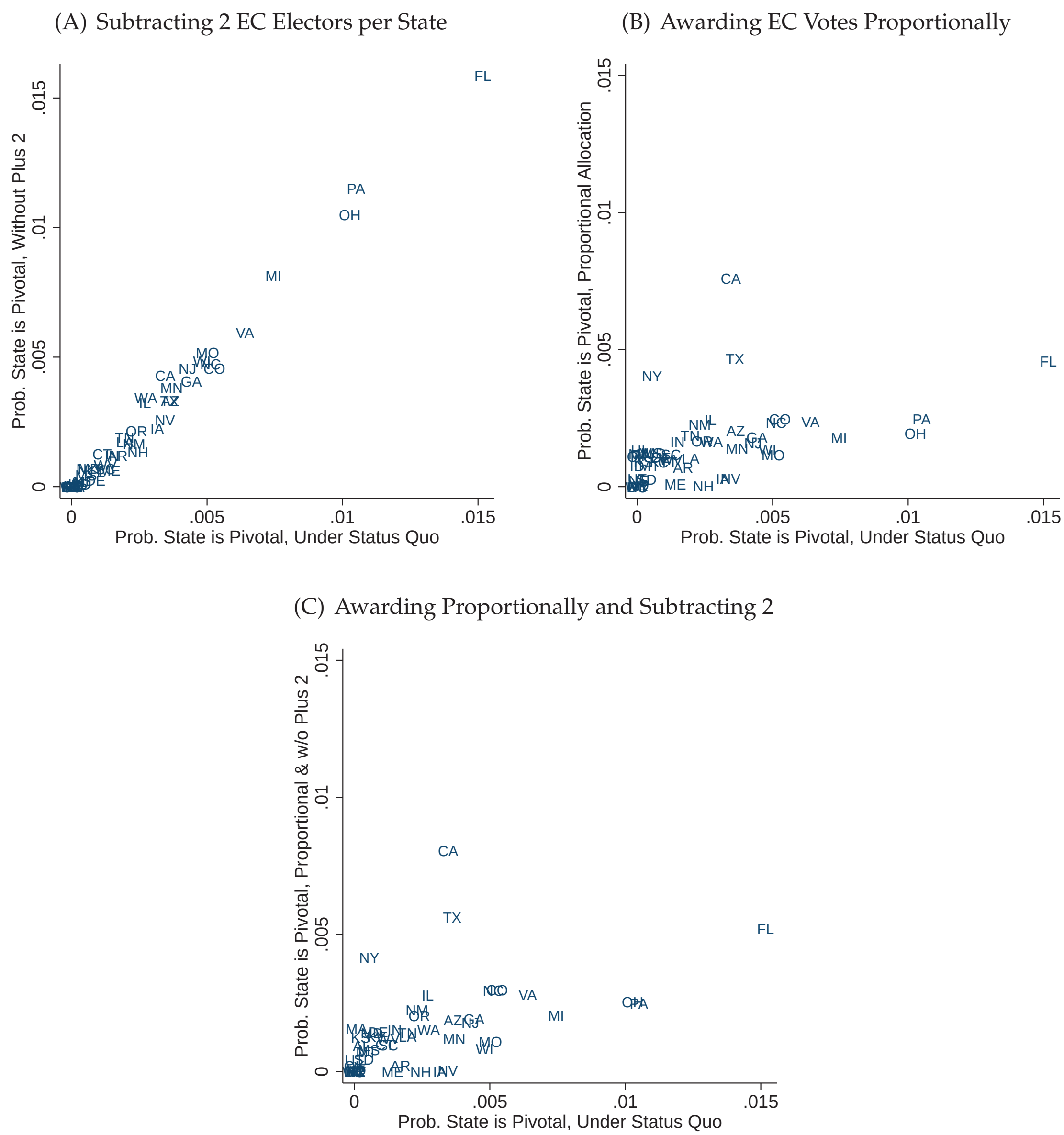


(B) Margins of Victory Often Differ by Party in the Largest States



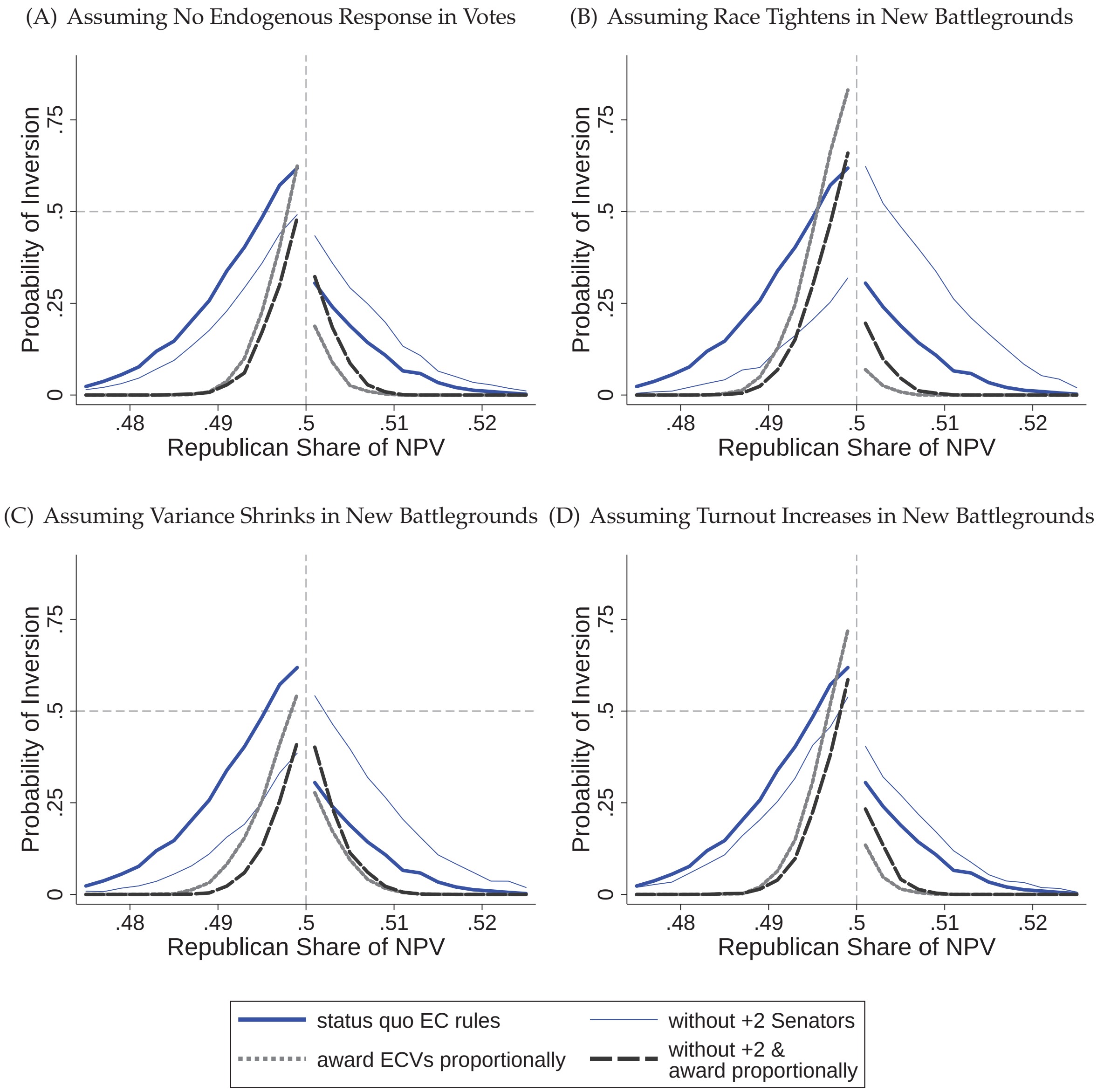
Note: Panel A plots, for the four largest and four smallest states by today's population, the state's apportionment of EC electors divided by its population. Panel B plots the average vote margins over time by Democrat and Republican candidates for the largest states.

Figure A14: New Battlegrounds: Probability of Being Pivotal Under Various EC Counterfactuals



Note: Figure plots the probability that a state is pivotal under status quo EC rules ($\hat{Q}^0_s$, horizontal axes) and various counterfactual EC rules ($\hat{Q}^{CF}_s$, vertical axes). The counterfactuals considered are indicated in the panel headings. To calculate $\hat{Q}^{CF}_s$, we re-assign 0.5% of statewide votes from the state winner to the state loser for each of the original 100,000 simulation runs in Model M1. We then calculate the fraction of simulated elections in which this reassignment of votes would flip the EC outcome. See Appendix F for additional detail.

Figure A15: Counterfactuals that Allow for Endogenous Responses to a Changing Electoral Map

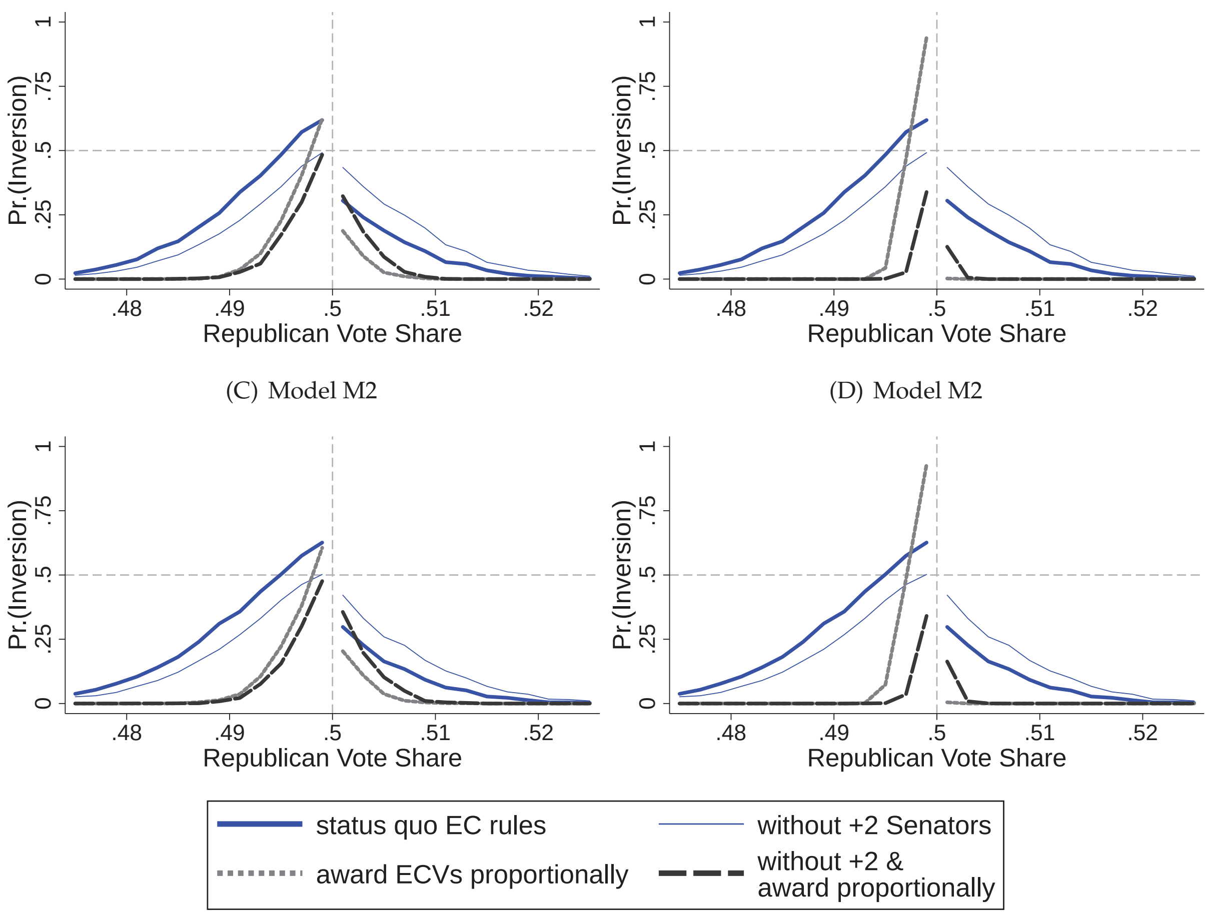


Note: Figure shows how the counterfactuals for the M1 model in Panel A (repeated from Panel A of Figure 4) change under simulations that allow for endogenous response to the counterfactuals. We incorporate a stylized, reduced-form representation of behavioral responses to the changing electoral map, as states move in or out of "battleground" and "safe" status under the counterfactual EC aggregation rules. The panels show results for different assumptions. In Panel B, we assume that margins will tighten in new battleground states. In Panel C, we assume that variances of potential voting outcomes will shrink in new battleground states. In Panel D, we assume that turnout will increase in new battleground states. See Appendix F for additional detail.

Figure A16: Electoral Inversions Under Alternative Aggregation: Modern Period

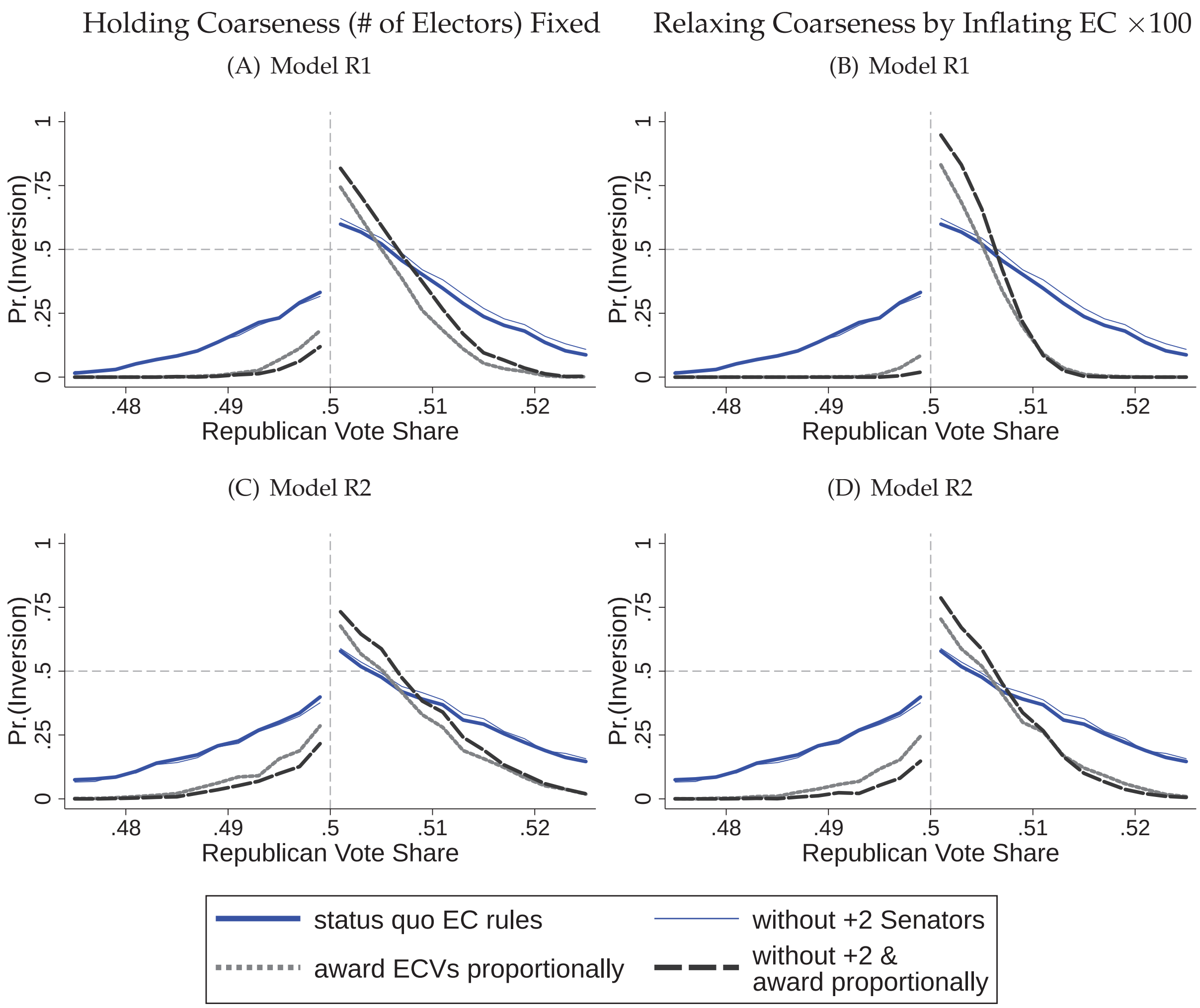
Holding Coarseness Fixed at 538 Electors | Relaxing Coarseness by Inflating EC ×100

(A) Model M1 (B) Model M1



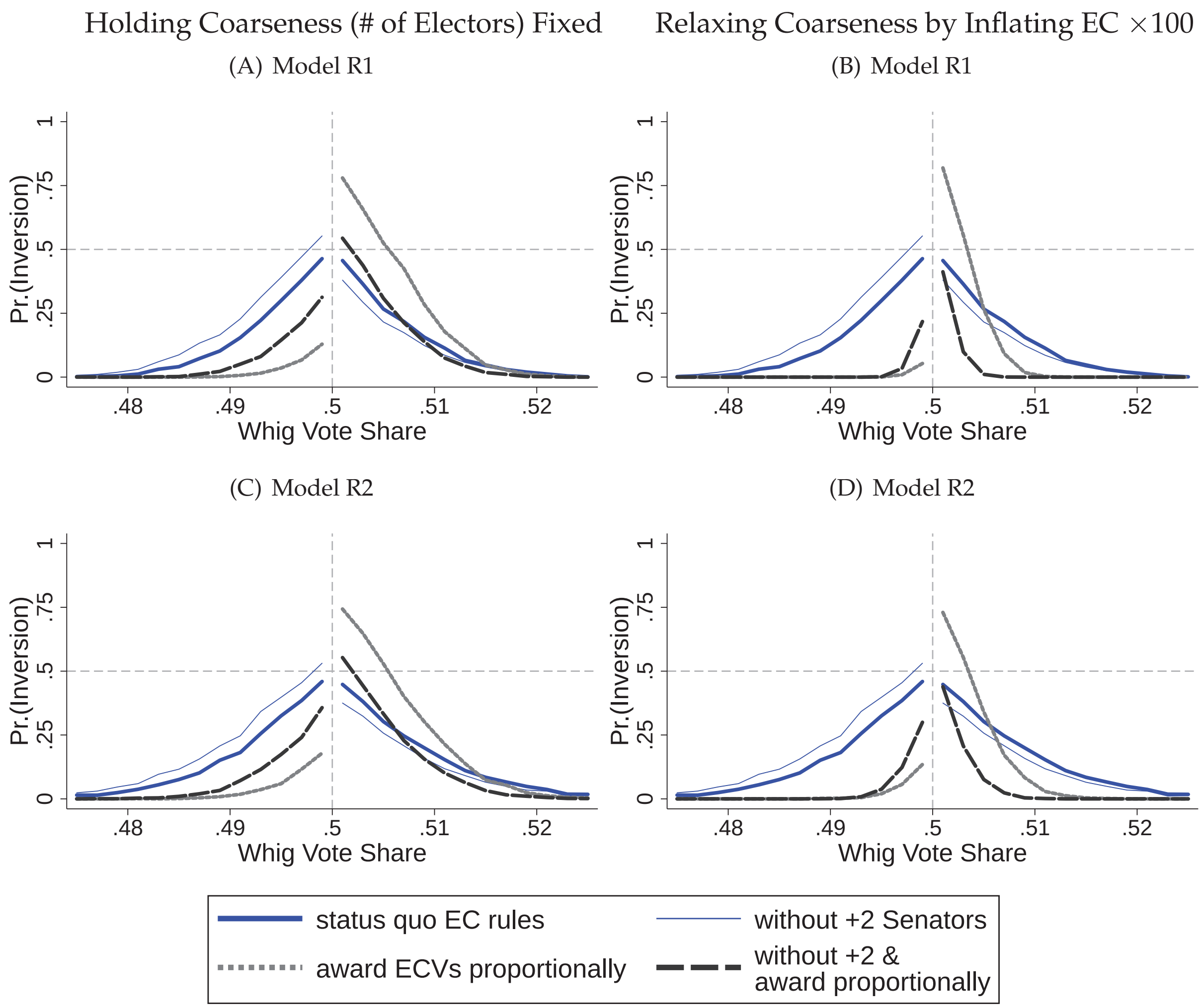
Note: Figure illustrates inversions under alternative EC aggregation rules that translate citizen votes into a presidential winner. The alternative that removes the two Senator-derived EC electors assigns each state electors equal in number to the size of the state's US House delegation. The alternative that removes the winner-takes-all condition awards state EC votes (ECVs) according to each candidate's popular vote share in the state, up to a rounding error. The left panels holds the congressional apportionment fixed at 435 House members and 2 Senators per state. The right panel inflates the congressional delegation size by 100 times to examine the impact of relaxing the rounding error ("coarseness") constraint. In these, each state receives 200 electors corresponding to Senators and these 200 are removed in the "without Senators" alternatives.

Figure A17: Electoral Inversions Under Alternative Aggregation: Post-Reconstruction Period



Note: Figure illustrates inversions under alternative EC aggregation rules that translate citizen votes into a presidential winner. The alternative that removes the two Senator-derived EC electors assigns each state electors equal in number to the size of the state's US House delegation. The alternative that removes the winner-takes-all condition awards state EC votes (ECVs) according to each candidate's popular vote share in the state, up to a rounding error. The left panels holds the congressional apportionment fixed at 435 House members and 2 Senators per state. The right panel inflates the congressional delegation size by 100 times to examine the impact of relaxing the rounding error ("coarseness") constraint. In these, each state receives 200 electors corresponding to Senators and these 200 are removed in the "without Senators" alternatives.

Figure A18: Electoral Inversions Under Alternative Aggregation: Antebellum Period



Note: Figure illustrates inversions under alternative EC aggregation rules that translate citizen votes into a presidential winner. The alternative that removes the two Senator-derived EC electors assigns each state electors equal in number to the size of the state's US House delegation. The alternative that removes the winner-takes-all condition awards state EC votes (ECVs) according to each candidate's popular vote share in the state, up to a rounding error. The left panels holds the congressional apportionment fixed at 435 House members and 2 Senators per state. The right panel inflates the congressional delegation size by 100 times to examine the impact of relaxing the rounding error ("coarseness") constraint. In these, each state receives 200 electors corresponding to Senators and these 200 are removed in the "without Senators" alternatives.

Table A1: ML Parameter Estimates for Variance Terms

	Modern Period: 1988-2016								post-Reconstruction: 1872-1888				Antebellum: 1836-1852		
	State and National Shocks	State, National, & Other Shocks	State Shocks Only	State, National, & Other Shocks	State, National, & Other Shocks	State, National, & Other Shocks	Omitting Inversions 2000 & 2016	Extended M1: 1964-2016	State and National Shocks	State, National, & Other Shocks	State Shocks Only	Omitting Inversions 1876 & 1888	State and National Shocks	State, National, & Other Shocks	State Shocks Only
	M1	M2	M5	M7	M8	M9	M10	M12	R1	R2	R5	R10	A1	A2	A5
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
σ_y	0.012	0.014		0.015	0.012	0.013	0.014	0.053	0.018	0.003		0.018	0.013	0.011	
σ_s	0.027	0.013	0.039	0.022	0.018	0.016	0.022	0.051	0.123	0.075	0.140	0.131	0.020	0.018	0.032
Other shocks															
region		0.005								0.051				0.002	
race		0.198		0.240		0.129									
education		1.396			2.208	1.738									
# of α_s estimates (states)	51	51	51	51	51	51	51	51	37	37	37	37	25	25	25
										(estimation omits CO)				(omits CA, FL, IA, TX, WI)	

Note: Table reports maximum likelihood estimates of the parameters in Equation (1) of the main text under various model specifications, as indicated in the column headers. Model estimates in the table are grouped by period. Either 25 (because we only use states that were present throughout the Antebellum period for estimating variances), 37, or 51 expected state vote share parameters, $\overline{\alpha}_s$, are also estimated by joint maximum likelihood but not reported here. Depending on the data period, some states were not present for all election years within the sample frame or did not use a statewide citizen vote to determine EC votes. These, including Colorado in the post-Reconstruction frame, are simply assigned their mean for $\overline{\alpha}_s$ and do not contribute to estimating variance parameters.

Table A2: Stability: Probability of Inversion in Races Decided by <1pp Margin ($\pi(.01)$), After Imposing Alignment Shifts

Cells Report π(.01): Probability of Inversion in Elections Decided by Within 1 Point

Assumed Shift, in Vote Margin Points

Set of shifted states	Shift to More Democratic-Leaning								Shift to More Republican-Leaning		
	-10	-8	-6	-4	-2	Baseline	2	4	6	8	10
All swing states	0.40	0.41	0.40	0.41	0.41	0.42	0.43	0.44	0.44	0.45	0.46
All non-swing states	0.53	0.50	0.49	0.47	0.43	0.42	0.41	0.40	0.42	0.44	0.45
Other large states	0.48	0.48	0.46	0.44	0.43	0.42	0.41	0.40	0.41	0.42	0.41
States won by Trump 2016	0.40	0.41	0.39	0.41	0.42	0.42	0.43	0.42	0.43	0.43	0.44
States won by Clinton 2016	0.49	0.46	0.46	0.44	0.43	0.42	0.41	0.41	0.39	0.39	0.39
Swing states won by Trump 2016	0.41	0.40	0.40	0.41	0.41	0.42	0.43	0.43	0.44	0.44	0.44
Swing states won by Clinton 2016	0.41	0.41	0.41	0.42	0.42	0.42	0.43	0.43	0.43	0.43	0.44
Swing county fraction	0.42	0.43	0.42	0.43	0.42	0.42	0.42	0.42	0.41	0.41	0.41
Swing states (individually)											
Colorado	0.42	0.42	0.42	0.42	0.42	0.42	0.43	0.43	0.42	0.42	0.43
Florida	0.41	0.41	0.42	0.42	0.42	0.42	0.42	0.43	0.43	0.43	0.43
Iowa	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.43	0.43	0.43
Michigan	0.43	0.43	0.42	0.42	0.42	0.42	0.43	0.43	0.42	0.43	0.43
Minnesota	0.43	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.43
Nevada	0.42	0.42	0.42	0.42	0.42	0.42	0.43	0.43	0.43	0.43	0.43
New Hampshire	0.42	0.42	0.42	0.42	0.42	0.42	0.43	0.43	0.43	0.43	0.43
North Carolina	0.41	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.43	0.42
Ohio	0.41	0.42	0.42	0.42	0.42	0.42	0.42	0.43	0.43	0.43	0.43
Pennsylvania	0.42	0.42	0.43	0.42	0.42	0.42	0.42	0.42	0.43	0.43	0.43
Virginia	0.41	0.42	0.41	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42
Wisconsin	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.43	0.43
Other large states (individually)											
California	0.45	0.44	0.44	0.43	0.43	0.42	0.42	0.42	0.41	0.41	0.41
Texas	0.42	0.42	0.42	0.43	0.42	0.42	0.42	0.42	0.42	0.41	0.41
New York	0.44	0.44	0.43	0.43	0.43	0.42	0.42	0.42	0.42	0.41	0.41
Illinois	0.43	0.43	0.43	0.42	0.43	0.42	0.42	0.42	0.42	0.42	0.42
Georgia	0.42	0.42	0.42	0.42	0.43	0.42	0.42	0.42	0.42	0.42	0.42

Note: Rows specify the sets of states that is shifted for the exercise. Column headers indicate the size of the imposed shift in the states' vote share distribution. For example, row 1 shifts the state-specific sampling distribution for all swing states left (Democratic) or right (Republican). The shift spans -10 points to $+10$ points in two-point increments. Other rows repeat the exercise for all non-swing states, other large states, states won by Trump in 2016, states won by Clinton in 2016, etc., as indicated. Swing county fraction, from (Cullen, Turner and Washington, 2018), is the fraction of counties in a state that changed presidential party in elections from 1988 to 2008. (Alaska, which Cullen, Turner and Washington, 2018 omit, is assigned the national median.) This fraction multiplies the indicated shift, so that a state made up entirely of swing counties receives the full shift, and a state with no swing counties receives no shift. Each permutation cell in the tables represents 100,000 simulation draws.

Table A3: Stability: Probability of Republican EC Win, After Imposing Alignment Shifts

	Cells Report Unconditional Probability of Republican EC Victory										
	Assumed Shift, in Vote Margin Points										
Set of shifted states	Shift to More Democratic-Leaning									Shift to More Republican-Leaning	
	-10	-8	-6	-4	-2	Baseline	2	4	6	8	10
All swing states	0.12	0.16	0.21	0.28	0.34	0.42	0.50	0.58	0.65	0.71	0.77
All non-swing states	0.17	0.22	0.27	0.32	0.37	0.42	0.47	0.53	0.58	0.63	0.69
Other large states	0.33	0.35	0.37	0.39	0.41	0.42	0.44	0.45	0.47	0.50	0.52
States won by Trump 2016	0.10	0.15	0.20	0.26	0.34	0.42	0.51	0.60	0.68	0.76	0.83
States won by Clinton 2016	0.25	0.28	0.31	0.34	0.38	0.42	0.47	0.51	0.56	0.61	0.66
Swing states won by Trump 2016	0.18	0.22	0.26	0.31	0.36	0.42	0.48	0.54	0.59	0.64	0.68
Swing states won by Clinton 2016	0.33	0.34	0.36	0.38	0.40	0.42	0.44	0.46	0.48	0.49	0.51
Swing county fraction	0.15	0.19	0.24	0.30	0.36	0.42	0.49	0.56	0.62	0.68	0.74
Swing states (individually)											
Colorado	0.40	0.40	0.41	0.41	0.42	0.42	0.43	0.43	0.43	0.44	0.44
Florida	0.34	0.35	0.37	0.39	0.40	0.42	0.43	0.45	0.46	0.46	0.47
Iowa	0.41	0.41	0.41	0.41	0.42	0.42	0.42	0.43	0.43	0.43	0.44
Michigan	0.40	0.40	0.40	0.41	0.41	0.42	0.43	0.44	0.45	0.46	0.46
Minnesota	0.41	0.41	0.41	0.41	0.42	0.42	0.42	0.43	0.43	0.44	0.45
Nevada	0.41	0.41	0.41	0.41	0.42	0.42	0.42	0.43	0.43	0.43	0.43
New Hampshire	0.41	0.41	0.41	0.42	0.42	0.42	0.42	0.42	0.43	0.43	0.43
North Carolina	0.38	0.39	0.40	0.41	0.41	0.42	0.43	0.43	0.43	0.43	0.43
Ohio	0.37	0.38	0.39	0.40	0.41	0.42	0.43	0.44	0.45	0.45	0.46
Pennsylvania	0.39	0.39	0.40	0.40	0.41	0.42	0.43	0.44	0.45	0.47	0.48
Virginia	0.39	0.39	0.40	0.41	0.41	0.42	0.43	0.43	0.43	0.44	0.44
Wisconsin	0.40	0.41	0.41	0.41	0.42	0.42	0.43	0.43	0.44	0.44	0.45
Other large states (individually)											
California	0.41	0.41	0.42	0.42	0.42	0.42	0.43	0.43	0.44	0.46	0.47
Texas	0.37	0.39	0.40	0.41	0.42	0.42	0.42	0.42	0.43	0.43	0.43
New York	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.43	0.43
Illinois	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.43	0.43	0.44	0.45
Georgia	0.38	0.39	0.40	0.41	0.42	0.42	0.42	0.43	0.43	0.43	0.43

Note: Rows specify the sets of states that is shifted for the exercise. Column headers indicate the size of the imposed shift in the states' vote share distribution. For example, row 1 shifts the state-specific sampling distribution for all swing states left (Democratic) or right (Republican). The shift spans -10 points to $+10$ points in two-point increments. Other rows repeat the exercise for all non-swing states, other large states, states won by Trump in 2016, states won by Clinton in 2016, etc., as indicated. Swing county fraction, from (Cullen, Turner and Washington, 2018), is the fraction of counties in a state that changed presidential party in elections from 1988 to 2008. (Alaska, which Cullen, Turner and Washington, 2018 omit, is assigned the national median.) This fraction multiplies the indicated shift, so that a state made up entirely of swing counties receives the full shift, and a state with no swing counties receives no shift. Each permutation cell in the tables represents 100,000 simulation draws.

Table A4: Stability: Probability of Republican EC Win Conditional on Inversion Occurring, After Imposing Alignment Shifts

	Cells Report Probability of Republican EC Victory, Conditional on Inversion Occurring										
	Assumed Shift, in Vote Margin Points										
Set of shifted states	Shift to More Democratic-Leaning									Shift to More Republican-Leaning	
	-10	-8	-6	-4	-2	Baseline	2	4	6	8	10
All swing states	0.25	0.32	0.41	0.53	0.66	0.77	0.85	0.91	0.93	0.95	0.97
All non-swing states	0.99	0.98	0.96	0.93	0.87	0.77	0.61	0.43	0.27	0.16	0.10
Other large states	0.96	0.95	0.93	0.89	0.84	0.77	0.67	0.57	0.47	0.38	0.33
States won by Trump 2016	0.76	0.76	0.75	0.74	0.75	0.77	0.79	0.80	0.82	0.84	0.84
States won by Clinton 2016	0.95	0.92	0.89	0.85	0.81	0.77	0.72	0.67	0.61	0.57	0.53
Swing states won by Trump 2016	0.39	0.45	0.52	0.61	0.69	0.77	0.83	0.88	0.91	0.92	0.93
Swing states won by Clinton 2016	0.63	0.65	0.68	0.71	0.73	0.77	0.79	0.81	0.83	0.83	0.84
Swing county fraction	0.80	0.79	0.78	0.78	0.77	0.77	0.77	0.76	0.75	0.75	0.74
Swing states (individually)											
Colorado	0.74	0.74	0.75	0.75	0.76	0.77	0.77	0.78	0.78	0.78	0.78
Florida	0.63	0.66	0.69	0.72	0.74	0.77	0.78	0.78	0.78	0.78	0.76
Iowa	0.76	0.76	0.76	0.76	0.76	0.77	0.77	0.78	0.78	0.79	0.79
Michigan	0.77	0.76	0.76	0.76	0.76	0.77	0.78	0.79	0.80	0.81	0.82
Minnesota	0.78	0.77	0.77	0.76	0.76	0.77	0.77	0.78	0.79	0.79	0.80
Nevada	0.74	0.74	0.75	0.75	0.76	0.77	0.77	0.78	0.78	0.78	0.79
New Hampshire	0.75	0.75	0.75	0.76	0.76	0.77	0.77	0.77	0.78	0.78	0.78
North Carolina	0.72	0.74	0.75	0.76	0.76	0.77	0.76	0.76	0.75	0.74	0.72
Ohio	0.70	0.71	0.72	0.73	0.75	0.77	0.78	0.79	0.79	0.79	0.79
Pennsylvania	0.76	0.75	0.75	0.75	0.75	0.77	0.78	0.80	0.82	0.83	0.84
Virginia	0.72	0.73	0.74	0.75	0.76	0.77	0.77	0.77	0.77	0.76	0.76
Wisconsin	0.76	0.76	0.76	0.76	0.76	0.77	0.77	0.78	0.79	0.80	0.80
Other large states (individually)											
California	0.91	0.89	0.86	0.83	0.80	0.77	0.74	0.71	0.69	0.69	0.70
Texas	0.76	0.79	0.79	0.79	0.78	0.77	0.74	0.72	0.69	0.65	0.62
New York	0.87	0.85	0.83	0.81	0.79	0.77	0.74	0.72	0.70	0.68	0.66
Illinois	0.83	0.82	0.80	0.79	0.78	0.77	0.76	0.76	0.76	0.76	0.77
Georgia	0.72	0.74	0.75	0.76	0.76	0.77	0.76	0.76	0.75	0.74	0.73

Note: Rows specify the sets of states that is shifted for the exercise. Column headers indicate the size of the imposed shift in the states' vote share distribution. For example, row 1 shifts the state-specific sampling distribution for all swing states left (Democratic) or right (Republican). The shift spans -10 points to $+10$ points in two-point increments. Other rows repeat the exercise for all non-swing states, other large states, states won by Trump in 2016, states won by Clinton in 2016, etc., as indicated. Swing county fraction, from (Cullen, Turner and Washington, 2018), is the fraction of counties in a state that changed presidential party in elections from 1988 to 2008. (Alaska, which Cullen, Turner and Washington, 2018 omit, is assigned the national median.) This fraction multiplies the indicated shift, so that a state made up entirely of swing counties receives the full shift, and a state with no swing counties receives no shift. Each permutation cell in the tables represents 100,000 simulation draws.

Table A5: Model List for Figure 3

Model Index	Parametric/ Bootstrap?	Empirical/ Hyperparam./ Hybrid	Description
1	Parametric	Empirical	M1 (see paper)
2	Parametric	Empirical	M2 (see paper)
3	Bootstrap	Empirical	M3 (see paper)
4	Bootstrap	Empirical	M4 (see paper)
5	Parametric	Empirical	M5 (see paper)
6	Bootstrap	Empirical	M6 (see paper)
7	Parametric	Empirical	M7 (see paper)
8	Parametric	Empirical	M8 (see paper)
9	Parametric	Empirical	M9 (see paper)
10	Parametric	Empirical	M10 (see paper)
11	Bootstrap	Empirical	M11 (see paper)
12	Parametric	Empirical	M12 (see paper)
13	Parametric	Empirical	M1-like; sets state turnout and EC allocations according to 1988
14	Parametric	Empirical	M1-like; sets state turnout and EC allocations according to 1992
15	Parametric	Empirical	M1-like; sets state turnout and EC allocations according to 1996
16	Parametric	Empirical	M1-like; sets state turnout and EC allocations according to 2000
17	Parametric	Empirical	M1-like; sets state turnout and EC allocations according to 2004
18	Parametric	Empirical	M1-like; sets state turnout and EC allocations according to 2008
19	Parametric	Empirical	M1-like; sets state turnout and EC allocations according to 2012
20	Parametric	Empirical	M1-like; assigns all third party votes to Republicans before estimation
21	Parametric	Empirical	M1-like; assigns all third party votes to Democrats before estimation
22	Bootstrap	Empirical	M3-like; for all states, place additional weight (15%) on bootstrap draws from same year
23	Bootstrap	Empirical	M3-like; for all states, place additional weight (20%) on bootstrap draws from same year
24	Bootstrap	Empirical	M3-like; for all states, place additional weight (25%) on bootstrap draws from same year
25	Bootstrap	Empirical	M3-like; for all states, place additional weight (30%) on bootstrap draws from same year
26	Bootstrap	Empirical	M3-like; for all states, place additional weight (35%) on bootstrap draws from same year
27	Bootstrap	Empirical	M3-like; for all states, place additional weight (40%) on bootstrap draws from same year
28	Bootstrap	Empirical	M3-like; for all states, place additional weight (45%) on bootstrap draws from same year
29	Bootstrap	Empirical	M3-like; for all states, place additional weight (50%) on bootstrap draws from same year
30	Bootstrap	Empirical	M3-like; for swing states only, place additional weight (15%) on bootstrap draws from same year
31	Bootstrap	Empirical	M3-like; for swing states only, place additional weight (20%) on bootstrap draws from same year
32	Bootstrap	Empirical	M3-like; for swing states only, place additional weight (25%) on bootstrap draws from same year
33	Bootstrap	Empirical	M3-like; for swing states only, place additional weight (30%) on bootstrap draws from same year
34	Bootstrap	Empirical	M3-like; for swing states only, place additional weight (35%) on bootstrap draws from same year
35	Bootstrap	Empirical	M3-like; for swing states only, place additional weight (40%) on bootstrap draws from same year
36	Bootstrap	Empirical	M3-like; for swing states only, place additional weight (45%) on bootstrap draws from same year
37	Bootstrap	Empirical	M3-like; for swing states only, place additional weight (50%) on bootstrap draws from same year
38	Bootstrap	Empirical	M3-like; for safe states only, place additional weight (15%) on bootstrap draws from same year
39	Bootstrap	Empirical	M3-like; for safe states only, place additional weight (20%) on bootstrap draws from same year
40	Bootstrap	Empirical	M3-like; for safe states only, place additional weight (25%) on bootstrap draws from same year
41	Bootstrap	Empirical	M3-like; for safe states only, place additional weight (30%) on bootstrap draws from same year
42	Bootstrap	Empirical	M3-like; for safe states only, place additional weight (35%) on bootstrap draws from same year
43	Bootstrap	Empirical	M3-like; for safe states only, place additional weight (40%) on bootstrap draws from same year
44	Bootstrap	Empirical	M3-like; for safe states only, place additional weight (45%) on bootstrap draws from same year
45	Bootstrap	Empirical	M3-like; for safe states only, place additional weight (50%) on bootstrap draws from same year
46	Parametric	Empirical	M1-like, but using t distribution with 7 d.o.f. instead of normal distribution for shocks
47	Parametric	Empirical	M1-like, but with state-specific coefficients multiplying the national component of the shock
48	Parametric	Hybrid	Shifts all state means in M1 left by 0.2 NPV percentage points post estimation
49	Parametric	Hybrid	Shifts all state means in M1 left by 0.4 NPV percentage points post estimation
50	Parametric	Hybrid	Shifts all state means in M1 left by 0.6 NPV percentage points post estimation

Note: Table continues on next page.

Table A6: Model List for Figure 3 (Cont.)

51	Parametric	Hybrid	Shifts all state means in M1 left by 0.8 NPV percentage points post estimation	
52	Parametric	Hybrid	Shifts all state means in M1 right by 0.2 NPV percentage points post estimation	
53	Parametric	Hybrid	Shifts all state means in M1 right by 0.4 NPV percentage points post estimation	
54	Parametric	Hybrid	Shifts all state means in M1 right by 0.6 NPV percentage points post estimation	
55	Parametric	Hybrid	Shifts all state means in M1 right by 0.8 NPV percentage points post estimation	
56	Parametric	Hybrid	Shifts all state means in M3 left by 0.2 NPV percentage points post estimation	
57	Parametric	Hybrid	Shifts all state means in M3 left by 0.4 NPV percentage points post estimation	
58	Parametric	Hybrid	Shifts all state means in M3 left by 0.6 NPV percentage points post estimation	
59	Parametric	Hybrid	Shifts all state means in M3 left by 0.8 NPV percentage points post estimation	
60	Parametric	Hybrid	Shifts all state means in M3 right by 0.2 NPV percentage points post estimation	
61	Parametric	Hybrid	Shifts all state means in M3 right by 0.4 NPV percentage points post estimation	
62	Parametric	Hybrid	Shifts all state means in M3 right by 0.6 NPV percentage points post estimation	
63	Parametric	Hybrid	Shifts all state means in M3 right by 0.8 NPV percentage points post estimation	
64	Parametric	Hyperpar.	Common national shock SD set to 0.5 pp; state shock SD set to 3.0 pp; no region shock	
65	Parametric	Hyperpar.	Common national shock SD set to 1.0 pp; state shock SD set to 3.0 pp; no region shock	
66	Parametric	Hyperpar.	Common national shock SD set to 1.5 pp; state shock SD set to 3.0 pp; no region shock	
67	Parametric	Hyperpar.	Common national shock SD set to 2.0 pp; state shock SD set to 3.0 pp; no region shock	
68	Parametric	Hyperpar.	Common national shock SD set to 2.5 pp; state shock SD set to 3.0 pp; no region shock	
69	Parametric	Hyperpar.	Common national shock SD set to 3.0 pp; state shock SD set to 3.0 pp; no region shock	
70	Parametric	Hyperpar.	Common national shock SD set to 3.5 pp; state shock SD set to 3.0 pp; no region shock	
71	Parametric	Hyperpar.	Common national shock SD set to 4.0 pp; state shock SD set to 3.0 pp; no region shock	
72	Parametric	Hyperpar.	Common national shock SD set to 4.5 pp; state shock SD set to 3.0 pp; no region shock	
73	Parametric	Hyperpar.	Common national shock SD set to 0.5 pp; state shock SD set to 4.0 pp; no region shock	
74	Parametric	Hyperpar.	Common national shock SD set to 1.0 pp; state shock SD set to 4.0 pp; no region shock	
75	Parametric	Hyperpar.	Common national shock SD set to 1.5 pp; state shock SD set to 4.0 pp; no region shock	
76	Parametric	Hyperpar.	Common national shock SD set to 2.0 pp; state shock SD set to 4.0 pp; no region shock	
77	Parametric	Hyperpar.	Common national shock SD set to 2.5 pp; state shock SD set to 4.0 pp; no region shock	
78	Parametric	Hyperpar.	Common national shock SD set to 3.0 pp; state shock SD set to 4.0 pp; no region shock	
79	Parametric	Hyperpar.	Common national shock SD set to 3.5 pp; state shock SD set to 4.0 pp; no region shock	
80	Parametric	Hyperpar.	Common national shock SD set to 4.0 pp; state shock SD set to 4.0 pp; no region shock	
81	Parametric	Hyperpar.	Common national shock SD set to 4.5 pp; state shock SD set to 4.0 pp; no region shock	
82	Parametric	Hyperpar.	National SD = 2.1 pp; state SD = 3.0 pp; region shock set to match M1 in overall NPV dispersion	
83	Parametric	Hyperpar.	National SD = 2.1 pp; state SD = 4.0 pp; region shock set to match M1 in overall NPV dispersion	
84	Parametric	Hyperpar.	National SD = 2.1 pp; state SD = 5.0 pp; region shock set to match M1 in overall NPV dispersion	
85	Parametric	Hyperpar.	National SD = 2.2 pp; state SD = 3.0 pp; region shock set to match M1 in overall NPV dispersion	
86	Parametric	Hyperpar.	National SD = 2.2 pp; state SD = 4.0 pp; region shock set to match M1 in overall NPV dispersion	
87	Parametric	Hyperpar.	National SD = 2.2 pp; state SD = 5.0 pp; region shock set to match M1 in overall NPV dispersion	
88	Parametric	Hyperpar.	National SD = 2.3 pp; state SD = 3.0 pp; region shock set to match M1 in overall NPV dispersion	
89	Parametric	Hyperpar.	National SD = 2.3 pp; state SD = 4.0 pp; region shock set to match M1 in overall NPV dispersion	
90	Parametric	Hyperpar.	National SD = 2.3 pp; state SD = 5.0 pp; region shock set to match M1 in overall NPV dispersion	
91	Parametric	Hyperpar.	National SD = 2.4 pp; state SD = 3.0 pp; region shock set to match M1 in overall NPV dispersion	
92	Parametric	Hyperpar.	National SD = 2.4 pp; state SD = 4.0 pp; region shock set to match M1 in overall NPV dispersion	
93	Parametric	Hyperpar.	National SD = 2.4 pp; state SD = 5.0 pp; region shock set to match M1 in overall NPV dispersion	
94	Parametric	Hyperpar.	National SD = 2.6 pp; state SD = 3.0 pp; region shock set to match M1 in overall NPV dispersion	
95	Parametric	Hyperpar.	National SD = 2.6 pp; state SD = 4.0 pp; region shock set to match M1 in overall NPV dispersion	
96	Parametric	Hyperpar.	National SD = 2.6 pp; state SD = 5.0 pp; region shock set to match M1 in overall NPV dispersion	
97	Parametric	Hyperpar.	National SD = 2.8 pp; state SD = 3.0 pp; region shock set to match M1 in overall NPV dispersion	
98	Parametric	Hyperpar.	National SD = 2.8 pp; state SD = 4.0 pp; region shock set to match M1 in overall NPV dispersion	
99	Parametric	Hyperpar.	National SD = 2.8 pp; state SD = 5.0 pp; region shock set to match M1 in overall NPV dispersion	
100	Parametric	Hyperpar.	National SD = 2.1 pp; state SD = 4.0 pp; region shock set to match M1 NPV dispersion; all draws	*t*
101	Parametric	Hyperpar.	National SD = 2.2 pp; state SD = 4.0 pp; region shock set to match M1 NPV dispersion; all draws	*t*
102	Parametric	Hyperpar.	National SD = 2.3 pp; state SD = 4.0 pp; region shock set to match M1 NPV dispersion; all draws	*t*
103	Parametric	Hyperpar.	National SD = 2.4 pp; state SD = 4.0 pp; region shock set to match M1 NPV dispersion; all draws	*t*
104	Parametric	Hyperpar.	National SD = 2.5 pp; state SD = 4.0 pp; region shock set to match M1 NPV dispersion; all draws	*t*
105	Parametric	Hyperpar.	National SD = 2.6 pp; state SD = 4.0 pp; region shock set to match M1 NPV dispersion; all draws	*t*
106	Parametric	Hyperpar.	National SD = 2.7 pp; state SD = 4.0 pp; region shock set to match M1 NPV dispersion; all draws	*t*
107	Parametric	Hyperpar.	National SD = 2.8 pp; state SD = 4.0 pp; region shock set to match M1 NPV dispersion; all draws	*t*
108	Parametric	Hyperpar.	National SD = 2.9 pp; state SD = 4.0 pp; region shock set to match M1 NPV dispersion; all draws	*t*
109	Parametric	Hyperpar.	National SD = 3.0 pp; state SD = 4.0 pp; region shock set to match M1 NPV dispersion; all draws	*t*

Note: Table lists details for each of the variants on the models included in Figure 3.

Table A7: Asymmetry: Probability that an Observed EC Win Was Caused by an Inversion

	Probability that Inversion Occurrs, Conditional on a Republican/Whig Win								Probability that Inversion Occurrs, Conditional on a Democrat Win						
		Conditional on Margin of Victory					Any			Conditional on Margin of Victory					Any
	Model	<0.5pp	<1pp	<2pp	<3pp	<4pp	Margin		Model	<0.5pp	<1pp	<2pp	<3pp	<4pp	Margin
Republicans and Democrats	M1	0.47	0.45	0.41	0.36	0.32	0.22	Republicans and Democrats	M1	0.43	0.37	0.25	0.18	0.13	0.05
	M2	0.47	0.44	0.40	0.36	0.32	0.18		M2	0.44	0.37	0.26	0.18	0.13	0.04
	M3	0.54	0.59	0.66	0.70	0.71	0.72		M3	0.32	0.21	0.08	0.05	0.03	0.03
	M4	0.52	0.52	0.52	0.53	0.52	0.45		M4	0.43	0.36	0.23	0.15	0.09	0.05
	M5	0.52	0.56	0.62	0.65	0.67	0.68		M5	0.38	0.28	0.15	0.09	0.07	0.06
	M6	0.55	0.57	0.63	0.66	0.68	0.69		M6	0.38	0.27	0.14	0.09	0.07	0.05
	M7	0.49	0.45	0.40	0.36	0.32	0.19		M7	0.42	0.35	0.25	0.17	0.13	0.04
	M8	0.46	0.45	0.40	0.36	0.32	0.23		M8	0.43	0.36	0.24	0.17	0.12	0.05
	M9	0.48	0.46	0.41	0.36	0.32	0.21		M9	0.43	0.35	0.24	0.17	0.12	0.04
	M10	0.46	0.43	0.37	0.32	0.28	0.18		M10	0.43	0.36	0.26	0.18	0.14	0.05
	M11	0.52	0.53	0.54	0.54	0.54	0.53		M11	0.39	0.30	0.17	0.10	0.07	0.05
	M12	0.47	0.42	0.35	0.28	0.23	0.06		M12	0.45	0.40	0.31	0.24	0.19	0.05
Republicans and Democrats	R1	0.45	0.40	0.32	0.25	0.20	0.06	Republicans and Democrats	R1	0.47	0.46	0.42	0.39	0.35	0.21
	R2	0.46	0.43	0.36	0.30	0.26	0.08		R2	0.50	0.46	0.42	0.39	0.36	0.20
	R3	0.43	0.36	0.25	0.19	0.15	0.12		R3	0.47	0.46	0.42	0.41	0.40	0.40
	R4	0.45	0.41	0.34	0.29	0.27	0.19		R4	0.43	0.37	0.28	0.24	0.22	0.22
	R5	0.44	0.39	0.30	0.23	0.18	0.09		R5	0.49	0.49	0.47	0.45	0.44	0.40
	R6	0.45	0.40	0.31	0.25	0.21	0.12		R6	0.50	0.49	0.48	0.47	0.45	0.43
	R10	0.45	0.41	0.33	0.26	0.21	0.06		R10	0.48	0.45	0.40	0.36	0.32	0.19
Whigs and Democrats	A1	0.45	0.39	0.30	0.24	0.19	0.10	Whigs and Democrats	A1	0.45	0.39	0.30	0.23	0.18	0.09
	A2	0.44	0.40	0.33	0.26	0.22	0.10		A2	0.45	0.40	0.31	0.25	0.20	0.09
	A3	0.48	0.46	0.42	0.41	0.38	0.25		A3	0.39	0.29	0.18	0.15	0.17	0.20
	A4	0.45	0.41	0.32	0.23	0.18	0.11		A4	0.39	0.30	0.19	0.15	0.15	0.16
	A5	0.46	0.42	0.35	0.31	0.28	0.26		A5	0.44	0.39	0.30	0.25	0.22	0.19
	A6	0.46	0.42	0.35	0.30	0.28	0.26		A6	0.44	0.39	0.31	0.26	0.23	0.21

Note: Table reports statistics describing the probability of inversions, conditional on the Electoral College win going to the indicated party. Columns additionally condition on various two-party popular vote share margins. Model estimates in the table are grouped by period: (M)odern, (R)econstruction, (A)ntebellum.

Table A8: Asymmetry: Who Wins from an Inversion?

		Probability of Republican/Whig Win, Conditional on Inversion Occurring					
		Conditional on Margin of Victory					Any
	Model	<0.5pp	<1pp	<2pp	<3pp	<4pp	Margin
Republicans and Democrats	M1	0.68	0.70	0.73	0.75	0.76	0.77
	M2	0.68	0.71	0.74	0.77	0.78	0.79
	M3	0.79	0.84	0.90	0.92	0.92	0.93
	M4	0.69	0.72	0.78	0.82	0.84	0.85
	M5	0.66	0.72	0.80	0.84	0.85	0.85
	M6	0.68	0.73	0.81	0.84	0.85	0.86
	M7	0.71	0.72	0.75	0.77	0.78	0.79
	M8	0.67	0.71	0.74	0.76	0.78	0.79
	M9	0.72	0.73	0.76	0.78	0.79	0.80
	M10	0.62	0.63	0.66	0.68	0.69	0.69
	M11	0.67	0.72	0.78	0.81	0.82	0.82
	M12	0.61	0.60	0.61	0.61	0.62	0.62
Republicans and Democrats	R1	0.36	0.34	0.31	0.30	0.28	0.27
	R2	0.39	0.39	0.38	0.37	0.36	0.33
	R3	0.77	0.75	0.73	0.71	0.70	0.70
	R4	0.85	0.86	0.86	0.86	0.86	0.86
	R5	0.36	0.33	0.30	0.28	0.26	0.24
	R6	0.45	0.43	0.40	0.37	0.36	0.34
	R10	0.50	0.48	0.46	0.45	0.44	0.43
Whigs and Democrats	A1	0.51	0.51	0.51	0.51	0.51	0.51
	A2	0.50	0.51	0.51	0.51	0.51	0.50
	A3	0.48	0.51	0.54	0.51	0.44	0.37
	A4	0.41	0.43	0.43	0.40	0.35	0.28
	A5	0.51	0.51	0.52	0.52	0.51	0.51
	A6	0.49	0.49	0.49	0.48	0.48	0.48

Note: Table reports statistics describing the probability that inversions were won by the index party (Republican/Whig). Columns condition on various two-party popular vote share margins. Statistics for Democrats are one minus the indicated value in the table. Model estimates in the table are grouped by period: (M)odern, (R)econstruction, (A)ntebellum.

Table A9: Inversion Probabilities in Counterfactuals that Allow for Endogenous Response

	Probabiltiy of Inversion		
	If Margin < 1 p.p.	If Margin < 2 p.p.	Unconditional
	(1)	(2)	(3)
	Panel A: Baseline Decomposition from Main Paper		
status quo	0.424	0.350	0.123
without +2 Senators	0.416	0.337	0.114
award ECVs proportionally	0.296	0.176	0.047
without +2 & award ECVs proportionally	0.293	0.170	0.045
	Panel B: Adjust Tightness of Race in New Safe/Battleground States		
status quo	0.424	0.350	0.123
without +2 Senators	0.411	0.340	0.126
award ECVs proportionally	0.375	0.251	0.070
without +2 & award ECVs proportionally	0.355	0.240	0.068
	Panel C: Adjust Variance in New Safe/Battleground States		
status quo	0.424	0.350	0.123
without +2 Senators	0.410	0.327	0.109
award ECVs proportionally	0.322	0.209	0.051
without +2 & award ECVs proportionally	0.303	0.181	0.044
	Panel D: Adjust Turnout in New Safe/Battleground States		
status quo	0.424	0.350	0.123
without +2 Senators	0.416	0.340	0.115
award ECVs proportionally	0.331	0.206	0.054
without +2 & award ECVs proportionally	0.304	0.181	0.047

Note: Table reports summary statistics for the counterfactuals considered in Figure A15. See Figure A15 and Appendix F for additional detail.

Bibliography to the Online Appendix

Bakthavachalam, Vinod, and Jake Fuentes. 2017. "The Impact of Close Races on Electoral College and Popular Vote Conflicts in US Presidential Elections." Princeton University mimeo.

Ball, William J, and David A Leuthold. 1991. "Estimating the likelihood of an unpopular verdict in the electoral college." *Public Choice*, 70(2): 215–224.

Banzhaf III, John F. 1968. "One man, 3.312 votes: a mathematical analysis of the Electoral College." *Vill. L. Rev.*, 13: 304.

Blair, Douglas H. 1979. "Electoral College reform and the distribution of voting power." *Public Choice*, 34(2): 201–215.

Butler, David E. 1951. "Appendix to The British General Election of 1950. Ed. H. G. Nichols."

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2008. "Bootstrap-Based Improvements for Inference with Clustered Errors." *The Review of Economics and Statistics*, 90(3): 414–427.

Cervas, Jonathan R, and Bernard Grofman. 2019. "Are Presidential Inversions Inevitable? Comparing Eight Counterfactual Rules for Electing the US President." *Social Science Quarterly*.

Cullen, Julie Berry, Nicholas Turner, and Ebonya Washington. 2018. "Political Alignment, Attitudes Toward Government, and Tax Evasion." Working Paper.

de Mouzon, Olivier, Thibault Laurent, Michel Le Breton, and Dominique Lepelley. 2018. "The theoretical Shapley–Shubik probability of an election inversion in a toy symmetric version of the US presidential electoral system." *Social Choice and Welfare*, 1–33.

de Mouzon, Olivier, Thibault Laurent, Michel Le Breton, and Dominique Lepelley. 2019. "Exploring the effects of national and regional popular vote Interstate compact on a toy symmetric version of the Electoral College: an electoral engineering perspective." *Public Choice*, 179(1-2): 51–95.

DeWitt, Darin, and Thomas Schwartz. 2016. "A calamitous compact." *PS: Political Science & Politics*, 49(4): 791–796.

Dubey, Pradeep, and Lloyd S Shapley. 1979. "Mathematical properties of the Banzhaf power index." *Mathematics of Operations Research*, 4(2): 99–131.

Enos, Ryan D, and Anthony Fowler. 2018. "Aggregate effects of large-scale campaigns on voter turnout." *Political Science Research and Methods*, 6(4): 733–751.

Epperly, Brad, Christopher Witko, Ryan Strickler, and Paul White. 2019. "Rule by violence, rule by law: Lynching, Jim Crow, and the continuing evolution of voter suppression in the US." *Perspectives on Politics*, 1–14.

Estes, Todd. 2011. "The Connecticut effect: The great compromise of 1787 and the history of small state impact on Electoral College outcomes." *Historian*, 73(2): 255–283.

Gaines, Brian J. 2001. "Popular myths about popular vote-electoral college splits." *PS: Political Science and Politics*, 34(1): 71–75.

Garand, James C, and T Wayne Parent. 1991. "Representation, swing, and bias in US presidential elections, 1872-1988." *American Journal of Political Science*, 1011–1031.

Gelman, Andrew, and Gary King. 1990. "Estimating the electoral consequences of legislative redistricting." *Journal of the American Statistical Association*, 85(410): 274–282.

Gelman, Andrew, and Gary King. 1994. "A Unified Method of Evaluating Electoral Systems and Redistricting Plans." *American Journal of Political Science*, 38(2): 514–554.

Gelman, Andrew, Gary King, and W John Boscardin. 1998. "Estimating the probability of events that have never occurred: when is your vote decisive?" *Journal of the American Statistical Association*, 93(441): 1–9.

Gelman, Andrew, Jonathan N Katz, and Francis Tuerlinckx. 2002. "The mathematics and statistics of voting power." *Statistical Science*, 420–435.

Gelman, Andrew, Nate Silver, and Aaron Edlin. 2012. "What is the probability your vote will make a difference?" *Economic Inquiry*, 50(2): 321–326.

Grofman, Bernard, William Koetzle, and Thomas Brunell. 1997. "An integrated perspective on the three potential sources of partisan bias: Malapportionment, turnout differences, and the geographic distribution of party vote shares." *Electoral Studies*, 16(4): 457–470.

Gudgin, Graham, and Peter J Taylor. 1979. "Seats." *Votes and the Spatial Organisation of Elections (Pion, London)*.

Kallina, Edmund F. 1985. "Was the 1960 Presidential Election Stolen? The Case of Illinois." *Presidential Studies Quarterly*, 113–118.

Katz, Jonathan N, Andrew Gelman, and Gary King. 2004. "Empirically evaluating the electoral college." In *Rethinking the Vote: The Politics and Prospects of American Election Reform.* , ed. Jon Krosnick, JM Miller, MP Tichy, AN Crigler, MR Just and EJ McCaffery. Oxford University Press.

Kikuchi, Kazuya. 2017. "The likelihood of majority inversion in an indirect voting system." *SSRN*.

Koza, John R. 2016. "A not-so-calamitous compact: a response to DeWitt and Schwartz." *PS: Political Science & Politics*, 49(4): 797–804.

Kuziemko, Ilyana, and Ebonya Washington. 2018. "Why did the Democrats lose the south? Bringing new data to an old debate." *American Economic Review*, 108(10): 2830–67.

Leip, Dave. 2018. "David Leip's atlas of U.S. Presidential Elections, Datasets."

Lichtman, Allan J, and M Kazin. 2010. "Elections and electoral eras." *The Princeton Encyclopedia of American Political History*, 281–289.

Merrill, Samuel. 1978. "Empirical estimates for the likelihood of a divided verdict in a presidental election." *Public Choice*, 33(2): 127–133.

Miller, Nicholas R. 2012. "Election inversions by the US Electoral College." In *Electoral Systems*. 93–127. Springer.

Miller, Nicholas R. 2013. "A priori voting power and the US Electoral College." In *Power, Voting, and Voting Power: 30 Years After*. 411–442. Springer.

Owen, Guillermo. 1975. "Evaluation of a presidential election game." *American Political Science Review*, 69(3): 947–953.

Rakove, Jack N. 2004. "Presidential selection: Electoral fallacies." *Political Science Quarterly*, 119(1): 21–37.

Riker, William H, and Peter C Ordeshook. 1968. "A Theory of the Calculus of Voting." *American Political Science Review*, 62(1): 25–42.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek. 2018. "IPUMS USA: Version 8.0 [dataset]."

Silver, Nate. 2016. "A User's Guide To FiveThirtyEight's 2016 General Election Forecast." *FiveThirtyEight.com*. https://fivethirtyeight.com/features/a-users-guide-to-fivethirtyeights-2016-general-election-forecast/. Accessed: 2019-6-13.

Strömberg, David. 2008. "How the Electoral College influences campaigns and policy: the probability of being Florida." *American Economic Review*, 98(3): 769–807.

Thomas, AC, Andrew Gelman, Gary King, and Jonathan N Katz. 2013. "Estimating Partisan Bias of the Electoral College Under Proposed Changes in Elector Apportionment." *Statistics, Politics and Policy*, 4(1): 1–13.

Warf, Barney. 2009. "The US Electoral College and spatial biases in voter power." *Annals of the Association of American Geographers*, 99(1): 184–204.