NBER WORKING PAPER SERIES

THE RISK OF NARROW, DISPUTABLE RESULTS IN THE U.S. ELECTORAL COLLEGE: 1836-2020

Michael Geruso
Dean Spears

Working Paper 27993
http://www.nber.org/papers/w27993

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2020, Revised November 2021

We thank Samuel Arenberg, Marika Cabral, Maya Eden, Andrew Gelman, Brendan Kline, David Molitor, Paolo G. Piacquadio, Itai Sher, Ishaana Talesara, Caroline Thomas and seminar participants at the University of Texas at Austin and the 2021 Welfare Economics and Economic Policy Seminar for helpful comments. We thank Nathan Franz for excellent research assistance. 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 additional relationships of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w27993.ack

The Risk of Narrow, Disputable Results in the U.S. Electoral College: 1836-2020
Michael Geruso and Dean Spears
NBER Working Paper No. 27993
October 2020, Revised November 2021
JEL No. H8,J1,J18,K16

ABSTRACT

Close elections are important for many reasons, including that consequent election disputes can weaken democratic legitimacy and risk political violence. We show that, in theoretical principle, either a popular vote or a two-stage electoral system, such as the US Electoral College (EC), could generate closer outcomes in expectation. We then quantify the probability of close outcomes in US presidential races with novel applications of empirical election models spanning all of US voting history. We show that razor-thin margins are, in fact, very likely under the EC. We then establish that the EC causes this closeness: It would not occur under any plausibly comparable popular vote system. The tendency of the EC to generate close elections is true today, throughout US presidential voting history, and for most likely configurations of future US politics.

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

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

1 Introduction

An election outcome is *close*[1] if changing some small fraction of the votes could reverse the outcome. A discipline-spanning literature in law, history, economics, and political science has been concerned with the consequences of close elections, including unclear results, weakened state capacity, and diminished legitimacy of leaders. Because the contribution of our paper is to comprehensively describe the *probability* and institutional *causes* of close elections in US presidential races, we begin by noting that this prior literature has identified a number of important *consequences*.

The prior literature has argued that: (i) Close elections are more vulnerable to manipulation or fraud, whether by foreign powers or domestic political actors (e.g., Posner, 2001; Hasen, 2005; Hirsch, 2020), and that the expectation of a close outcome may, in fact, incentivize fraud and other bad faith actions (Rapoport and Weinberg, 2000; Grofman and Feld, 2005). (ii) Close elections are more likely to be disputed, undercutting the value of elections in decisively transferring political power (Przeworski, 2018; Hasen, 2020).[2] (iii) Election officials (often in partisan jobs) and judges may replace voters as the pivotal actors in a narrowly decided election (Rakove, 2004; Gelman, Katz and Bafumi, 2004), such as when making determinations about voter intent from partially punched paper ballots or discarding "naked" mail ballots.[3] (iv) Voter perceptions of fairness and election integrity are critical for legitimacy (see, e.g., Weatherford, 1992; Banducci and Karp, 2003), and narrow margins are empirically linked to lower voter confidence in election integrity (Birch, 2008; Sances and Stewart III, 2015). Indeed, many surveyed Americans viewed the results of the close, disputed presidential elections of 2000 and 2020 as illegitimate.[4] These considerations are independent of any normative claim about who should win an election, given a set of voter intentions or voter actions; they are about the effects of closeness *per se*.

In fact, several US presidential elections have been uncannily close. The election in 2000 hinged

[1]We use the terms *close* and *narrow* as synonyms throughout, when referring to realized or potential election outcomes.

[2]For example, Przeworski (2018) argues that, "the greatest value of elections, for me by itself sufficient to cherish them, is that at least under some conditions they allow us to process in relative liberty and civic peace whatever conflicts arise in society, that they prevent violence." (p. 4) A tendency to extremely narrow outcomes, in practice, is likely to be a condition that interrupts such processing.

[3]Officials in each swing state typically discard thousands of mail ballots in US presidential elections due to signatures, enveloping, or postmarks. Such judgements are subjective and frequently contested. For example, in the 2016 presidential election, FL officials rejected 21,973 absentee ballots (U.S. Election Assistance Commission, 2017). In the 2020 election, Justice Alito ordered the quarantine of PA's mail ballots arriving after 8 p.m. on election day (for possible elimination), affecting about 10,000 ballots. Neither action was pivotal *ex post*, though similar judgements were possibly pivotal in the 2000 election.

[4]As of midyear 2001, following the Bush-Gore election, about one quarter of Americans said of Bush they "don't accept him as the legitimate president," and 17% said Bush "stole the election" (Gallup, 2001). Immediately following the Biden-Trump 2020 election, about one third of all surveyed Americans and over three quarters of Trump supporters reported believing that voter fraud was pivotal in Biden's 2020 victory (Monmouth University Poll, 2021). As of late April 2021, 30% of Americans reported "Biden did not legitimately win enough votes to win the presidency" (CNN, 2021).

on a final tally of just 537 votes in a single state, Florida. The 1916 presidential election hinged on 3,773 votes in California. The 1876 election was decided by 889 votes in South Carolina. The 1876 Hayes-Tilden election provides an illustrative example of the type of legitimacy crisis that can attend close election outcomes and why a close election may not produce a clear winner.[5] In that election, there were credible allegations that fraud favoring Hayes was pivotal to his statewide win in South Carolina. Because the presumptive Electoral College vote tally was within a single ballot nationally—185 for Hayes versus 184 for Tilden—a single Hayes elector lost would mean that neither candidate carried an Electoral College majority, which would trigger the US House to conduct its own vote to determine the presidency.[6] The razor-thin margin presented an opportunity to seize power through post-election maneuvers, and so several states sent opposed slates of electors (e.g., one certified by a Republican governor and another certified by a Democratic attorney general from the same state). Democrats in the US Senate threatened to filibuster the Electoral College vote count and certification process. Ultimately, the winner of the presidency was resolved via an unwritten, informal political compromise among party elites just a few days before the scheduled inauguration.[7] In short, the close 1876 election was attended by potentially pivotal fraud, several attempted subversions of statewide vote results by elected leaders, the threat of violence (including a shot fired at Hayes' residence and federal troops mobilized in the capital to brace for insurrection), and a president who was arguably selected via a political deal by party leaders, rather than by voters.[8]

The importance of these consequences motivates the question that our research answers: Were the 1876 election and other close, disputed election outcomes for US president statistical flukes? Or are close elections—and the risks they carry of perceived illegitimacy and Constitutional crisis—especially likely in the US's two-tiered Electoral College system, relative to a popular vote? Although there has been significant study among historians, legal scholars, and political scientists of the potential consequences of close elections in the US and abroad, such work hasn't established the empirical

[5]Facts about the 1876 election in the remainder of this paragraph are sourced from Haworth (1906) and Holt (2009).

[6]If neither candidate had a majority, the election would be decided by a new vote in the House of Representatives, nullifying all votes cast by citizens in the election. In Oregon, where the vote went for Hayes (a Republican), the state leadership was Democratic and officials attempted to send a Democratic elector to the Electoral College in defiance of the vote outcome. In Louisiana and Florida, the outcome was contested, and each state sent two opposed slates of electors. The election result remained disputed in the weeks leading up to the scheduled inauguration in March of 1877.

[7]Under this compromise, the 1876 presidency would go to the northern Republican Hayes in exchange for the end of Reconstruction and ultimately the freedom for southern Democrats to institute Jim Crow.

[8]This understanding of the risks associated with close elections, in which a governing party of a state might take actions that are pivotal in affecting an election outcome (holding fixed voter intentions or even holding fixed votes cast) is importantly at odds with Downs (1957), in which the assumption is "The only limit on government's powers is that the incumbent party cannot in any way restrict the political freedom of opposition parties or of individual citizens." (p 137)

probability of these events.[9] Nor has any prior work settled whether the Electoral College system is especially statistically prone to close outcomes relative to feasible alternatives. In this paper we address these questions, characterizing the probability of close outcomes under the US Electoral College (EC) and under the plausible alternative—a National Popular Vote (NPV). To emphasize: Although the consequences of close elections motivate our analysis, we do not make any new causal claims about such consequences, focusing instead on the less-studied issues of the likelihood and institutional causes of of close elections.

We begin by showing that, in theoretical principle—and contrary to the accounts of many election reform advocates—either a single-stage system (a popular vote) or a two-stage system (such as the EC) could generate closer outcomes in expectation. Our toy model makes clear that which type of electoral system is more likely to produce a close outcome depends on the facts of political alignment and polarization across second-tier voting units (states). Therefore, whether the EC is particularly likely to generate close outcomes is an empirical question.

To answer that question, we make a novel application of the presidential election models from several sources, primarily focusing on three: Gelman, Heidemanns and Morris (2020), Geruso, Spears and Talesara (2020), and Silver (2020). Geruso, Spears and Talesara (2020) estimate probability distributions over US presidential elections from 1836 to 2016 based on historical voting data. Silver (2020) and Gelman, Heidemanns and Morris (2020) are two prominent election forecasting models of the 2020 election, both estimated using polling data. We take the election models (these and others) as basic data in our empirical exercise. We show results from several sources to illustrate that our conclusions are robust to alternative modeling assumptions and estimation approaches.

As our first main result, we establish the fact that the EC is very likely to generate a close outcome. For example, the probability that an EC winner today is decided by 7,500 ballots or fewer (about 0.005% of votes cast) in a single pivotal state is greater than 1-in-40. The probability that the EC winner is decided by 75,000 ballots or fewer in a single state is about 1-in-6. That probability rises to 1-in-4 if we allow for the 75,000 votes to be spread across two or three states that together would be pivotal in reversing the outcome (though as we discuss, the risks of dispute and conflict may be lower if subverting voter intentions requires the coordination of activities across actors from different states).

[9] For the last few presidential elections, predictions of the risk of a close election have been byproducts of election forecasting models, which we incorporate here. No prior work has offered the comprehensive characterization that our research contributes (including how these probabilities have evolved over US history, the stability of these findings to alternative modeling choices, or any counterfactual comparisons to a plausible alternative electoral system).

Even an election decided by the extremely narrow margin of 150 votes in a single state has a non-negligible chance of occurring: a little less than 0.1%. In other words, the probability that a one-one millionth share of turnout is pivotal in deciding the election is about a one-in-one thousand event. These findings indicate that recent electoral experience of close presidential outcomes in 2000, 2016, and 2020 (decided by <0.001% of national turnout in one state, 0.060% in three states, and 0.028% in three states, respectively) were not statistical flukes. They were ex-ante likely.

We further examine whether close election risk is a modern phenomenon, specific to the current configuration of US politics. Using election models that extend back to 1836, we show that the EC has been likely to deliver very close outcomes even under markedly different territorial boundaries, party systems, and political configurations. This includes periods during which the US was made up of different states than today (25 states plus the voting Territory of Michigan at the beginning of our study period), different dominant parties (Whigs, Democrats, Republicans), and different groups of enfranchised voters (black men, women, the poor). The statistical tendency of the EC to produce narrow election outcomes has been present over the entire history of presidential voting in the US, despite such changes.

As a second main contribution, we then address whether the Electoral College *causes* this closeness. A plausible alternative explanation is that closeness is the product of strong positional competition between two parties. Under such a view, close outcomes would be frequent under any feasible electoral system, including a National Popular Vote. To shed light on this, we draw on evidence from popular-vote presidential elections in other countries and from US governor and US Senate elections. These are settings in which parties and voters endogenously respond to the electoral incentives of popular voting. Contrasting the empirical properties of the multi-stage EC with a wide range of single-stage electoral systems, we show that the EC is one or two orders of magnitude more likely to generate a close election than any comparable popular-vote electoral system that exists in practice (for any threshold of *closeness* one chooses).

Of course, claims about the statistical properties of a hypothetical popular vote for US president are necessarily difficult. Perhaps the observed patterns of popular voting for other offices can offer little guidance for this counterfactual. As an alternative approach, we assess the conditions under which an NPV system for US president would generate the same close-election risk as the EC presently does, and then ask whether such conditions are feasible. We show that to exactly match the likelihood that an election is decided by a few thousand votes in the EC, a calibrated model of voting under an NPV system would generate a distribution of voting outcomes that is implausibly tight: In such a model, it would be an extreme statistical fluke ($p < 10^{-10}$) for a presidential candidate to ever win more than 50.5% of the vote. While we cannot rule out the abstract possibility that national politics under an NPV would converge so completely and persistently to a near-50-50 tie, this is deeply implausible. We show that popular voting under any reasonable parameterization simply does not deliver close outcomes as frequently as does the Electoral College in statistical expectation.

Though no prior work has established the catalogue of facts we develop about the EC's close-election risk over history and relative to a potential NPV system,[10] a concern with the *consequences* of close elections is common. This concern motivates calls for electoral reform (see, e.g., Posner, 2001; Miller, 2012; Hasen, 2020; Hirsch, 2020). In *A Short History of Presidential Election Crises* (published prior to the 2020 election), Hirsch (2020) writes about the Bush-Gore race that "In 2000, we saw that a contested presidential election can produce chaos... Our systems remain vulnerable and another disputed election result seems inevitable, and yet, we are woefully unprepared for the crises that will ensue."[11] In the months prior to the 2020 US presidential election, experts frequently warned that the inadequacies of the electoral system meant that the US presidential race could result in violence—*if it turned out to be close*.[12] Our work informs this perennial public debate over whether the likelihood and danger of narrow, disputable election outcomes are sufficient to motivate electoral reform. Our findings establish that the empirical tendency towards frequently generating close electoral outcomes—and therefore also the theorized and documented challenges to a functioning democracy that result from close outcomes—is not a transitory, happenstantial phenomenon arising from present politics

[10]Our empirical results complement contemporaneous theoretical work attempting to understand these risks in stylized, mathematical models, including Merlin and Nagel (2020). Methodologically, our work is most closely related to studies in positive political science, economics, and mathematics that have estimated the probability that a voter is pivotal in determining an election (including Shapley and Shubik, 1954; Banzhaf III, 1964; Chamberlain and Rothschild, 1981; Gelman, King and Boscardin, 1998; Gelman, Katz and Tuerlinckx, 2002; Gelman, Katz and Bafumi, 2004; Gelman, Silver and Edlin, 2012), as well as the response of parties and campaign investment to these probabilities (Strömberg, 2008; Wright, 2009) and the voter turnout response to a close expected outcome (e.g., Cox and Munger, 1989; Vogl, 2014; see Gerber et al., 2020 for an overview). Prior work has focused largely on characterizing or predicting which state-level outcomes are likely to be pivotal in the national races; where political power thus resides geographically; and how optimizing campaigns or voters might respond. Our aims are different, characterizing the risk of a close election and comparing this risk to single-tier systems. Our work also connects to recent econometric studies at the intersection of demographics and voting power over US history, including Cascio and Washington (2014), Kuziemko and Washington (2018), and Cascio and Shenhav (2020).

[11]Hirsch is concerned primarily with the ability of electronic election tallies to be hacked in a close statewide race.

[12]See, for example: "Why the risk of post-election violence in the US is higher than at any time in recent memory" by Alex Ward, *Vox*, November 2, 2020; "'All the red flags': Political analysts warn of U.S. election violence in November" by Elliot Smith, *CNBC*, October 1, 2020; "Election Dispute Increases Risk Of Political Violence, Analysts Warn" by Hanah Allam, *NPR*, November 5, 2020.

and demographics. It is a fundamental property of the US Electoral College system. Narrow outcomes are likely to be an enduring feature of future elections, if held under the present system.[13]

2 Illustrative model

In this section, we present a toy model of two-tier voting that motivates our empirical analysis beginning in the next section. The model illustrates two facts. The first fact is that, in principle, either a two-tier system like the Electoral College (EC) or a single-tier system like a National Popular Vote (NPV) could be more likely to generate close elections. This fact is important to briefly illustrate because it contrasts with a popular intuition among some of the political scientists and election law experts most concerned with examining the *consequences* of narrow, disputable election outcomes. Despite widely varying views on the need for or potential consequences of EC reforms, many in this literature assume that a race decided by the US Electoral College is likelier to end in a closer electoral outcome than a popular vote simply because it is composed of many winner-takes-all state-level contests, and each of these smaller contests creates an additional opportunity for a narrow-margin outcome in some potentially pivotal voting unit (see, e.g., Posner, 2001, Hasen, 2005, and Hirsch, 2020).[14] This conjecture, assumption, or intuition is pervasive in the legal and political literature, despite being unquantified (and indeed, as we establish in this model, false).

The second insight from the model is that the partisan distribution total of citizen votes across the second-tier voting units (states) is a key factor determining whether a close election is more likely under a NPV or EC system. Illustrating these facts in our simple model helps to clarify confusion that has propagated in this literature, in which there has been remarkably little dialogue between the researchers focused on statistical forecasting and researchers concerned with election legitimacy and

[13] Of course, one cannot strictly rule out the possibility that changes to American politics and voter demographics in coming years will be even more extreme than those associated with the US Civil War and emancipation; the annexation of Texas, California, and other large swathes of land; women's suffrage; etc. In Section 5.3 we provide examples of the type of extreme political changes that would be sufficient to reduce the risk of a close outcome to various thresholds.

[14] From Hirsch (2020): "The argument that the Electoral College will produce more squeakers than the national vote... rests on the basic statistical principle that the greater the sample size, the greater the variation. Imagine there were only ten voters in every state. Assuming the candidates are equal, as a sheer matter of probability the fifty states would be almost certain to produce many five-five ties." From Hasen (2005): "On the one hand, the popular vote method opens up the possibility of post-election legal challenge anywhere in the United States in an effort to harvest votes, thereby broadening the scope of potential election problems for creative lawyers to examine. On the other hand, the popular vote method increases the vote margin between candidates, because it aggregates votes from all fifty states plus the District of Columbia." See also the National Popular Vote Interstate Compact (Koza et al., 2013): "The current state-by-state winner-take-all system repeatedly creates artificial crises because every presidential election generates 51 separate opportunities for a dispute because of an outcome-altering statewide margin."

vulnerability.

For the purposes of the stylized model, assume that there are three equal-sized states ($s \in \{1,2,3\}$) with equal EC representation (here, one EC vote each, though this does not matter). Denote the realized Republican vote share in each state as $x_s \in [0,1]$ and the expected Republican vote share—its partisan lean—as μ_s. States follow the "unit rule," so that a simple statewide vote plurality gains all electoral votes in the state. Numbering states according to increasing μ_s and assuming that the probability that state 2 is not pivotal for the EC is small enough to ignore, a candidate wins in an EC-like system if and only if they win the tipping point state—that is, if and only if their realized vote share in state 2 is greater than 0.5. A candidate wins in an NPV system if and only if their national vote share ($\bar{x} \equiv \frac{1}{3}\sum_s x_s$) is greater than 0.5. For simplicity, we are not assuming here that vote shares in each state are endogenous to the voting system, though doing so would not change the point illustrated by the model, which is that there are configurations of partisan leans of states that can produce either very close NPV outcomes or very close EC outcomes. The question of how voting occurs in an equilibrium in which parties and voters respond to the electoral systems in which they are embedded is deferred to the empirical investigation below (beginning in Section 3 for the EC system and Section 5 for an NPV).

To examine probability distributions over outcomes, let the vote shares x_s be normally distributed for each state. Assume a homoskedastic variance ($\sigma_s^2 = \sigma^2$) and that x_s are independent.[15] These assumptions help to focus this toy mathematical model and enable an analytical solution; we do not retain them in the empirical exercise that follows.[16] Because the sum of independent normal random variables is normal, the national vote (expressed as the Republican share) is normal with an expectation of $\frac{1}{3}\sum_s \mu_s$ and a variance of $\frac{1}{3}\sigma^2$.[17]

Let each state contain n voters. Under the EC system, the vote margin that is pivotal in determining

[15]The normal parameterization allows small tail probabilities to fall outside of [0,1]. See Appendix Section A for an alternative model formulation without this feature; it instead parameterizes vote distributions as normal in the log-odds ratios. This alternative also relaxes the assumption that state 2 is pivotal. The alternative is more complex to describe and solve but produces the same predictions.

[16]The *empirical* model of the data-generating process of election outcomes that we use below to establish our main results incorporates correlated national election-year shocks, as well as regional shocks, facts about states that are heterogenous sizes and in apportionment representation, etc.

[17]The variance of the national vote, expressed in terms of vote shares (not vote counts), is the variance of the across-state mean share, $\mathrm{Var}\left(\frac{1}{3}(x_1 + x_2 + x_3)\right)$.

Table 1: Disputed Election Risk in the EC and NPV: An Illustrative Model

map	states' sizes	state expectations μ_1	μ_2	μ_3	NPV, $\mathbb{E}[\bar{x}]$	expected number of votes needed to reverse outcome if EC system, $\mathbb{E}[m^{EC}]$	expected number of votes needed to reverse outcome if NPV system, $\mathbb{E}[m^{NPV}]$
A	1,000	35%	53%	65%	53%	47	73
B	1,000	20%	63%	70%	53%	130	73

Note: Table shows how the number of votes that are pivotal in determining the presidential winner in an EC or NPV system depends on geographic polarization. The table considers two alternative "maps" of where Democratic and Republican voters reside. In both maps, we hold the Republican share of the national popular vote, $\mathbb{E}[\bar{x}]$, fixed at 53%. In both maps, states each contain 1,000 voters. Map B is more geographically polarized than map A. As a result, in map A, it is more likely that a small number of votes could reverse the outcome in an EC system than in an NPV system. In map B, it is more likely that a small number of votes could reverse the outcome in an NPV system than in an EC system. Appendix Section A presents a more complex model that relaxes various assumptions imposed here but produces the same substantive results.

the outcome is merely the margin, m, in the pivotal state:

$$m^{EC} = n\,|0.5 - x_2| \text{ where } x_2 \sim \mathcal{N}\left(\mu_2, \sigma^2\right). \tag{EC}$$

Under the NPV system, the margin that is pivotal in determining the outcome is the national margin:

$$m^{NPV} = 3n\,|0.5 - \bar{x}| \text{ where } \bar{x} \sim \mathcal{N}\left(\frac{1}{3}\sum_s \mu_s, \frac{1}{3}\sigma^2\right). \tag{NPV}$$

Taking the expectations of these expressions for the pivotal margins shows that which voting system is more likely to deliver close election outcomes depends on how far μ_2 is from 0.5 relative to σ^2 and how far $\frac{1}{3}\sum_s \mu_s$ is from 0.5 relative to $\frac{1}{3}\sigma^2$.[18]

Table 1 illustrates this numerically, in an example in which we assume states of 1,000 voters, expected vote shares μ_s as indicated in the table, and a standard deviation of 5% ($\sigma = 50$) for all states. The two rows consider two alternative scenarios, described by electoral "maps," A and B. Both have the same expected Republican share of the vote nationally, so the national popular vote tally is held fixed across them. But the location of the Democratic and Republican voters differs, with map B more geographically polarized.

The electoral system expected to deliver closer outcomes depends on the map. In map A, the expectation for the decisive number of votes in the EC system ($\mathbb{E}[m^{EC}]$) is smaller than the expectation

[18] In particular, the expectations $\mathbb{E}[m^{EC}]$ and $\mathbb{E}[m^{NPV}]$ can be calculated as the means of folded normals: $\sqrt{\frac{2}{\pi}}\sigma \exp\left(\frac{-\mu^2}{2\sigma^2}\right) + \mu\left(1 - 2\Phi(\frac{-\mu}{\sigma})\right)$.

of the decisive number in the NPV system ($\mathbb{E}\left[m^{NPV}\right]$) because the tipping point state is close to 50-50 in expectation. In map B, in contrast, an NPV system is more likely to be decided by a small number of votes. Of the two maps, A is the scenario that better corresponds to the actual geographic alignment in US politics.[19]

Thus it is not the case that there is inherently more risk of a close outcomes in a two tiered system, as has been often assumed or conjectured. The facts of geographic partisan alignment matter. With that in mind, we now turn to empirical distributions of likely presidential election outcomes, estimated from voting and polling data and fully accounting for the empirical patterns of partisan alignment and polarization across states—as well as the significant heterogeneity in state sizes, EC representation per capita, and turnout.

3 Background and Methods

3.1 Background on the EC system

The two-tiered EC system is established in Article II, Section 1 of the Constitution, and affected by various constitutional amendments and state laws. In current practice, the first tier involves citizens voting for electors pledged to a candidate. Today, in all states except Maine and Nebraska, the statewide popular vote winner is awarded all of the state's EC representation, though there is no constitutional requirement to award electors in this way (or to involve citizens in presidential elections at all), and the state-specific arrangements meant to bind electors to vote for their pledged candidate are imperfect. At the second tier, electoral votes are tallied.

States are apportioned electoral votes equal to the size of the congressional delegation (US senators plus US representatives). The number of electors are thus linked to the Apportionment Acts of Congress, which determine the number of US House seats. In earlier periods that we study, fewer states and different apportionment law generated different sizes of the US Electoral College. For example, in 1836 when Whigs competed with Democrats across 25 states plus the Territory of Michigan, there were 294 EC ballots, and a winning majority was 148. Today there are 538 electors: 435

[19] In the modern US electoral map, the expected outcome of the states that are likely to be pivotal in the EC—here represented by μ_2—is close to 0.5, so the EC is more likely to have a very small number of decisive votes. But, as we show in a counterfactual computation in Section 5.3 and Figure 6, if the map changed such that geographic polarization increased dramatically (so that each state's μ_s moved away from 0.5 while preserving the national vote shares expected by each party), then the pivotal state's expectation (μ_2) could be much farther from 0.5 than the national expectation, and a small number of decisive votes would be more likely under the NPV than under the EC.

corresponding to US House members, 100 corresponding to senators, and 3 for Washington DC. The candidate with a simple majority of 270 votes in the electoral college wins the presidency. If neither candidate secures 270 ECVs, the US House of Representatives choses a candidate to become president, under special rules.

3.2 Data: External Election Model Output

To examine the probability that election outcomes in US presidential races will be close at some threshold of interest, we take existing probabilistic models of the US elections from several published sources. Two sources are widely-reported-on forecasting models for the 2020 presidential election by *The Economist* (Gelman, Heidemanns and Morris, 2020) and *FiveThirtyEight* (Silver, 2020), which generated probability distributions over the 2020 election outcome, estimated largely from polling data. We use these models as they were produced in late October 2020 just before the 2020 presidential election. These two models were perhaps the most scrutinized in the history of election forecasting.[20]

The third data source is Geruso, Spears and Talesara (2020), which is built from historical voting data, rather than polls, and back-casts distributions of potential outcomes in past presidential elections. Geruso, Spears and Talesara (2020) (GST for convenience) includes models describing presidential elections in various discrete periods up to 2016 and extending back to 1836.[21] For example, the Antebellum and post-Reconstruction election models in GST cover the periods 1836–1852 and 1872–1888, respectively, and are intended to describe potential outcomes in a typical election over those periods of relatively stable partisan geographies for a generic pair of Democratic, Republican, or Whig candidates competing.[22]

The three data sources each generate many thousands of simulated election outcomes as draws from their estimated statistical processes. Because the EC is a complex statistical object, parameter estimates (such as the conditional covariance of two states' shock process net of regional- and demographics-based covariance) are not in all cases commensurable across the three models, which have different structures and therefore estimate different parameters. The sets of simulated election draws, in contrast, are directly comparable across models and describe each model's implied distribu-

[20]See, e.g., "Nerd Wars: Nate Silver and G. Elliott Morris are trying to make sense of the 2020 election and each other," New York Magazine, October 15, 2020 by Matthew Zeitlin.

[21]In earlier periods, electors were most commonly decided by state legislatures, rather than citizen voters.

[22]These model periods correspond to periods of relatively stable partisan geographies, aligning mostly with political scientists' notion of "Party Systems."

tion over election possibilities. For simplicity, we often refer to the simulation output that characterizes these probability distributions as the *model*.

Each simulation draw represents an observation in our analysis, consisting of a vector of state-level vote shares for each of the two major-party contenders.[23] These state-level outcomes for the presidential vote, combined with data on apportioned electors and turnout by state, completely describes an election possibility in terms of the allocation of a state's ECVs, the Electoral College outcome, and the popular vote.[24] Because the election outcome in 2020 revealed significant polling bias that was reflected in the forecasting models, we subtract a 2 percentage point ex-post model bias favoring Democrats from every state-level simulated outcome in the 2020 *Economist* and *FiveThirtyEight* forecasts, but also report results using uncorrected model output.[25]

Figure 1 describes the data. The *Economist* and *FiveThirtyEight* models each consist of 40,000 simulation draws. GST consists of 100,000 draws per time-period-specific model. Only the most recent period model from GST (1988–2016) is included in Figure 1; summary statistics for models of other historical periods are presented below in Section 4.2. Panel (A) plots the probability distribution over the popular vote across the three models. Panel (B) plots the implied probability distribution over the EC outcome. The distributions in both panels are calculated simply as the frequency of discrete events in each set of simulation draws.

The Gelman, Heidemanns and Morris (2020), Geruso, Spears and Talesara (2020), and Silver (2020) models disagree on many particulars—central tendencies and higher-order moments are significantly different. Indeed, the three models are not even attempting to represent the same phenomena. The forecasts—Gelman, Heidemanns and Morris (2020) and Silver (2020)—are predicting a particular event one week before the outcome was resolved. These were estimated after the candidates were selected via the party primaries and after a global pandemic and a sharp recession occurred, among other facts. Differences between these two forecasting models, although the subject of much attention

[23] None of the three primary data sources in our study report third party vote shares. See GST for detailed discussion of the potential influence of third parties in these models.

[24] To facilitate comparability between the three datasets, we treat the 5 Electoral College ballots at stake at the congressional district level in Maine and Nebraska as if these were determined at large, even though the source for two of three models provides forecasts at the district level for these states.

[25] So a simulated Democratic vote share of 0.56 in a state becomes 0.54. The 2 percentage point vote share adjustment is based on comparing the realization of vote shares nationally and in battleground states to the model predictions. See also Gelman and Morris "Comparing election outcomes to our forecast and to the previous election" (https://statmodeling.stat.columbia.edu/2020/11/06/comparing-election-outcomes-to-our-forecast-and-to-the-previous-election/) for the assessment that the average state prediction in the *Economist* model was off by about 2.5 percentage points. Implementing the correction does not change any conclusion of this paper. See Figures A1 and A2 for uncorrected versions of the data and main results.

and debate, are unimportant for our purposes and are small in comparison to differences between either forecasting model and the non-forecasting model by Geruso, Spears and Talesara (2020). GST is methodologically and conceptually different. It estimates probability distributions that describe likely outcomes for a *generic* Democratic and Republican presidential candidate pair over a specified window (in Figure 1, 1988-2016). The GST back-casting models thus are intended to build-in higher-order uncertainty. This is clearly reflected in the dispersion differences in Figure 1.

We present results using various models/data sources in order to help the reader understand the extent to which our results are robust to different election models, estimated with different techniques and structures, and which are fundamentally different in the facts they are intended to describe. As we explain below when examining robustness, we further permute the model set beyond those in Figure 1. Model uncertainty and parameter uncertainty are important considerations because the set of observations is small—just 25 presidential election results per century.

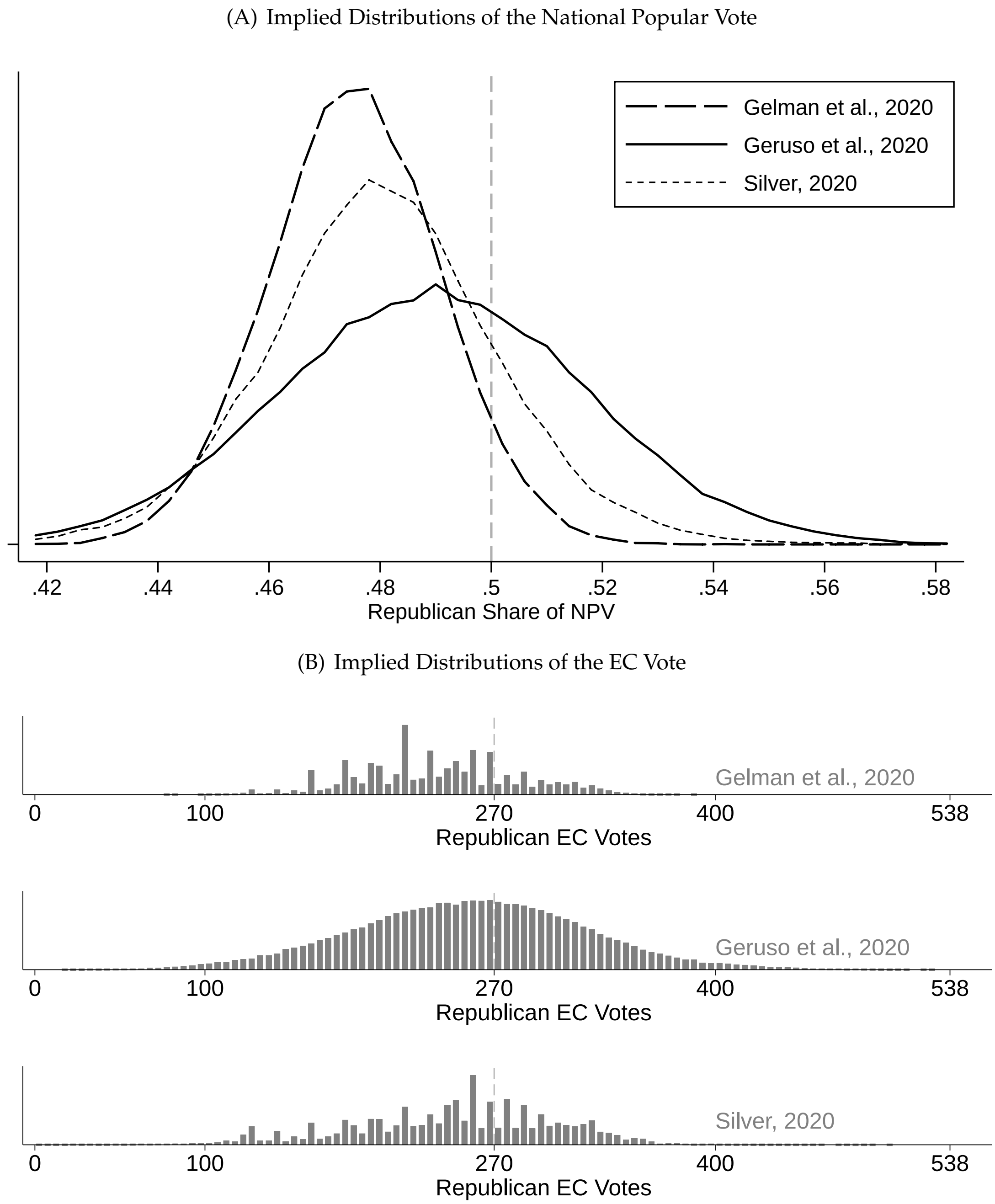
Unsurprisingly, for some outcomes, these models yield very different predictions. Yet, as we show below, our main finding regarding the probability of a close outcome in the Electoral College relative to the probability of a close outcome in a single-tier popular vote is not very sensitive to model choice. The types of fine differences important for forecasters—e.g., does some candidate have a 40%, 50%, or 60% chance of winning?—are unimportant here. That is because our findings are concerned with orders of magnitude: We show that voting in the EC system is about 100 times more likely to result in a race decided by a few hundred or few thousand votes than voting in a single-tier system. Thus, there need not be agreement on the "right" model of the EC to learn that the EC is especially likely to deliver close outcomes.

3.3 Identifying narrow outcomes

We compute the probability of a *close* election by observing how frequently, within the set of simulation draws of each model, discarding a small number of votes from the simulated winner's tally could overturn the election result. In particular, for each simulated election draw $q^j \in Q^j$ for election model j, we find the minimum number of discarded votes needed to change the outcome of the election. Because the EC tends to deliver close elections, even very close outcomes within a few hundred votes are not vanishingly rare events in the data.

Consider discarding some number of votes, n, from the winning candidate in a single state. An

Figure 1: Data and Descriptive Statistics: Distributions of Likely Election Outcomes from Several Sources



Note: Figure shows the probability distributions over electoral outcomes implied by three datasources, constructed from the 40,000 or 100,000 simulation draws in each datasource. Panel (A) plots the distribution of popular vote outcomes. Panel (B) plots the distribution of electoral vote outcomes.

election outcome could be altered by discarding those votes only if it were the case that *n* was greater than the difference between the state-level winner and loser and only if that state were individually pivotal in deciding the election. In our baseline results, we focus on cases in which the electoral outcome hinges on a small number of votes in a *single* state. Historical examples include 537 votes in Florida in 2000, 3,773 votes in California in 1916, and 889 votes in South Carolina in 1876. In other results, we allow the possibility that an election could hinge on discarding a small number of votes split across two or three carefully chosen states—like 2020, in which 43,000 votes split across AZ, GA, and WI decided the outcome.

There are reasons to believe that the scope for disputation and attendant adverse consequences is higher if the disputable margin resides in a single state. In the case of the 2020 presidential election contestation, for example, commentators conjectured that because election officials from multiple states would have had to act in concert to change the presidential election result in favor of Trump, the risk of that possibility was smaller.[26] Depending on the type of risk one has in mind—e.g., popular perceptions of illegitimacy versus a substantive subversion of voter intentions by state elections officials that actually changes the presidential outcome—the single state or multi-state case may be more useful metric. We present both types of results below.

To allow comparisons across time periods in which vastly different numbers of votes were cast—1.5 million in 1836 versus in excess of 150 million in 2020—we report closeness thresholds that are normalized by the national turnout. When we characterize the probability that an election is resolved by some margin *m* or less, we define *m* as the number of votes dropped from the winner's tally divided by the total number of votes cast. We report results for various values of *m* from 0.00001% of turnout to 0.1% of turnout. For context, the 2000, 2016, and 2020 races—the closest in recent history—were decided by 0.0005%, 0.060%, and 0.028% of the national two-party vote, respectively.

[26] E.g., Chris Hayes writes: "One thing becomes clearer and clearer by the day, which is that if it had come down to only one state, they probably would have pulled it off" (January 23, 2021, https://twitter.com/chrislhayes/status/1353158526017875978). See also Richard Hasen: "Trump Is Planning a Much More Respectable Coup Next Time" in *Slate* August 5, 2021 (https://slate.com/news-and-politics/2021/08/trump-2024-coup-federalist-society-doctrine.html).

4 Results

4.1 Main Result: Probabilities of Close Outcomes

Figure 2 presents our main result: the probability of a disputably narrow outcome, in which a small number of votes would change the winner of the presidency in the EC system. Separately for each data source, we plot probabilities that the election is decided by the indicated vote margin (indexed along the horizontal axis) in a single pivotal state. These first results focus on contemporary politics, using the 2020 forecasting models and the GST model covering the 1988-2016 period. In Section 4.2, we examine the same set of outcomes using models that span US history.

Figure 2 and the corresponding tabulations in Table 2 show that the probability an election would be reversed by a small share of votes is large and is not different across the three data sources. Scaling to 2020 turnout, the probability of an election decided by 150,000 votes or fewer (about 0.1 percent of turnout) is at least 18% in all models. At a threshold of 7,500 votes, the corresponding probability is about 3% in all models. Even the small margin of 150 votes has a non-negligible chance of occurring: 0.073% in the GST model (the only model precise enough to estimate that statistic).[27] In other words, the probability that less than a one-one millionth share of turnout is pivotal in deciding the election is nearly a one-in-one thousand event.

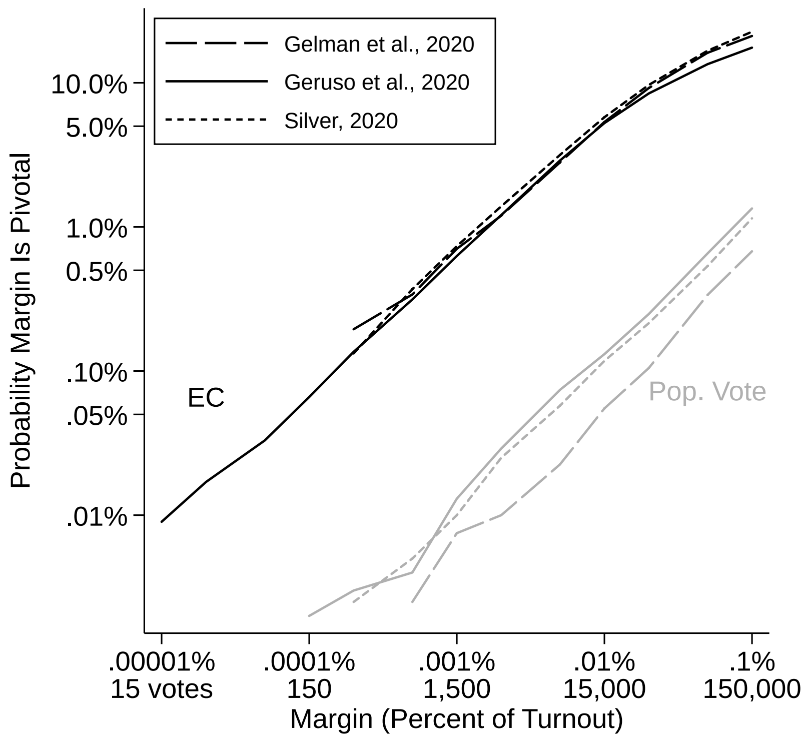
Although such an outcome may seem strikingly unlikely, history supports its plausibility. The 2000 election, in which over 100,000,000 votes were cast, was decided by a 537-vote margin. Of course, the value of such plots is not in producing ex-ante probability distributions over some particular past event like the 2000 race, which has been resolved with certainty, but in understanding the full function describing the risk of possible outcomes.

For reference, we also plot in Figure 2 the closeness of the national popular vote as it occurs in the same election simulation data. The figure shows that, for any threshold of interest, the popular vote is less likely to be close than the EC outcome by nearly two orders of magnitude. Despite close agreement on the probability of a close outcome in the Electoral College, the three models are dispersed in their predictions of the popular vote, highlighting the EC's complex mapping from citizen votes

[27] With 100,000 simulation runs in GST, we observe dozens of draws with elections decided by as little as 0.0001% of the vote ($\sim$ 150 votes by 2020 turnout). We cannot produce estimates at similarly low thresholds for the Gelman et al. and Silver models in part because fewer simulation draws are provided for those models (40,000 each), but primarily because the precision of their published data are truncated at 4 digits: A reported Republican share of 0.5000 for state s in simulation q does not distinguish between a state race won by two votes or two hundred votes.

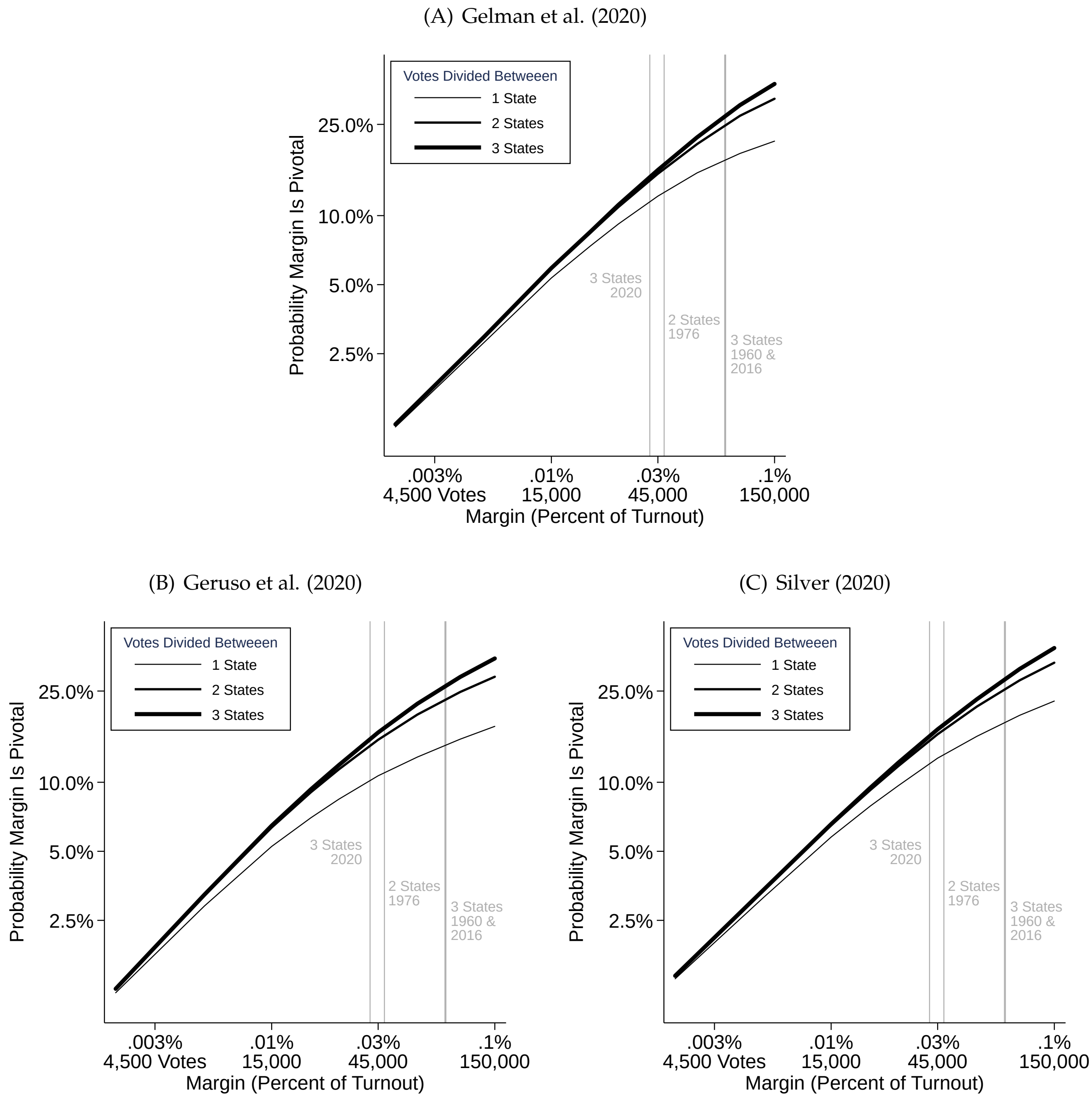
Figure 2: Main Result: Close Election Risk in the Electoral College

(A) Probabilities in Three Models (Single Pivotal State)



Note: Figure plots the probability that a national election outcome would be reversible by changing a small number of votes in a single state. The horizontal axis indicates the closeness margin being evaluated, labelled in the top line as a percent of overall turnout, and labelled in the bottom line as vote counts, assuming turnout of 150 million voters (roughly 2020 turnout). The vertical axis indicates the probability that the indicated margin is pivotal. Underlying data are derived from the simulation models of Gelman, Heidemanns and Morris (2020), Geruso, Spears and Talesara (2020) (modern period, 1988–2016), and Silver (2020). For the same set of simulation models, the lighter gray lines indicate the probabilities that the corresponding popular vote generated by the model is within the indicated closeness margin.

Figure 3: Close Election Risk, Allowing that Multiple States Could Be Pivotal if Combined



Note: Figure plots the probability that a national election outcome would be reversible by changing a small number of votes. The construction follows Figure 2 but allows for the indicated vote margin to be split across 2 or 3 states that together could be pivotal in determining the EC outcome. The horizontal axis indicates the closeness margin being evaluated, labelled in the top line as a percent of overall turnout, and labelled in the bottom line as vote counts, assuming turnout of 150 million voters (roughly 2020 turnout). The vertical axis indicates the probability that the indicated margin is pivotal. Panel (A) plots the probability of a close EC outcome in the models from Gelman, Heidemanns and Morris (2020); Panel (B) uses Geruso, Spears and Talesara (2020) (modern period, 1988–2016); and Panel (C) uses Silver (2020). The 1960, 1976, 2016, 2020 outcomes are plotted for reference, and set along the horizontal axis according to the margin in those races expressed as a fraction of the turnout in those election years.

to the electoral votes that determine the presidency. The popular vote statistics are, of course, not a projection of what *would* happen under a National Popular Vote system. Instead, they describe the likely outcomes for the popular vote as it is generated under the EC system. Below in Section 5, we characterize the expected closeness of single-tier—i.e., popular vote—systems using data generated in elections in which the popular vote decides the electoral contest.

In Figure 3, we calculate the probabilities that an m share of the national votes is pivotal in determining the outcome, if that share is split across two or three carefully chosen states. For very small margins—a few hundred to a few thousand votes—the probabilities are essentially identical to the single-state case. This is because it is unlikely that two states' elections, which together would be pivotal for the EC outcome, would coincide in being determined so narrowly. We thus focus on larger margins compared to Figure 2, though see Figure A3 for the full m range along the horizontal axis.

In Figure 3, allowing for the pivotal votes to be split across two or three states raises the odds of a close election. For a 0.1% margin (less than 150,000 votes by 2020 turnout), the probability rises from at least 18% (for a single pivotal state) to at least 35% (for three states) across all election models. To put the multi-state results in context, the figure marks the outcomes of the 1960, 1976, 2016 and 2020 election results, which were each decided by a small number of votes split across two or three states that were jointly pivotal. For 2016, this was about 78,000 votes split across PA, MI, WI. For 2020, this was about 43,000 votes split across AZ, GA, WI. Figure 3 makes clear that these events were not statistical flukes, but were ex ante likely outcomes. In fact, Panel (B), which uses a model trained on data that stops at 2016 (and for which the model output was fixed prior to the 2020 presidential primaries), predicts that observing a margin in 2020 that was as small as the realized 2020 margin was a 1-in-6 event, ex ante.

Is this risk conveyed in Figures 2 and 3 large? Over a voting lifetime (15 presidential elections in 60 years), the results imply greater than a 50% chance that a voter will experience a presidential election decided by 0.01% of turnout (today about 15,000 votes or fewer) in a single pivotal state. Considering a slightly larger threshold and allowing that closeness could consist of several states each with small margins combining to be pivotal, the results imply greater than a 98% chance that 0.05% of turnout (75,000 votes) spread across three or fewer states decides the presidency. Of course, the ex post probabilities of these events is 1.0 for voters who experienced the 2000 election (and the disputation constitutional crises that resulted) or the 2020 election (and the insurrection at the US

Table 2: Main Results (Tabulations from Figures 2 and 3)

Margin (Winner's Tally - Loser's Tally)		Probability		
Margin in Points (Shares*100)	Margin in Votes, Assuming 150M Two-Party Turnout	Gelman et al. (2020)	Geruso et al. (2020)	Silver (2020)
(1)	(2)	(3)	(4)	(5)
Panel A: Race Decided within Margin in a Single Pivotal State				
0.00001	15	-	0.01%	-
0.00002	30	-	0.02%	-
0.00005	75	-	0.03%	-
0.0001	150	-	0.07%	-
0.0002	300	0.20%	0.14%	0.13%
0.0005	750	0.34%	0.31%	0.37%
0.001	1,500	0.71%	0.63%	0.74%
0.002	3,000	1%	1%	1%
0.005	7,500	3%	3%	3%
0.01	15,000	5%	5%	6%
0.02	30,000	9%	8%	10%
0.05	75,000	16%	13%	17%
0.1	150,000	21%	18%	23%
Panel B: Race Decided within Margin Split Across Two States or Fewer				
0.00001	15	-	0.01%	-
0.00002	30	-	0.02%	-
0.00005	75	-	0.03%	-
0.0001	150	-	0.07%	-
0.0002	300	0.20%	0.14%	0.13%
0.0005	750	0.34%	0.32%	0.38%
0.001	1,500	0.71%	0.64%	0.74%
0.002	3,000	1%	1%	1%
0.005	7,500	3%	3%	3%
0.01	15,000	6%	6%	6%
0.02	30,000	11%	11%	12%
0.05	75,000	22%	21%	23%
0.1	150,000	32%	29%	33%
Panel C: Race Decided within Margin Split Across Three States or Fewer				
0.00001	15	-	0.01%	-
0.00002	30	-	0.02%	-
0.00005	75	-	0.03%	-
0.0001	150	-	0.07%	-
0.0002	300	0.20%	0.14%	0.13%
0.0005	750	0.34%	0.32%	0.38%
0.001	1,500	0.71%	0.64%	0.74%
0.002	3,000	1%	1%	1%
0.005	7,500	3%	3%	3%
0.01	15,000	6%	6%	7%
0.02	30,000	11%	12%	12%
0.05	75,000	24%	24%	25%
0.1	150,000	38%	35%	39%

Note: Table lists the probabilities that a national election outcome would be reversible by changing a small number of votes. Tabulations correspond to the results in Figures 2 and 3. See corresponding figure notes for additional detail. We cannot produce estimates at very low *closeness* thresholds for the Gelman, Heidemanns and Morris (2020) and Silver (2020) models because the precision of their published data is truncated at 4 digits: A reported Republican share of 0.5000 for some large state s in simulation q does not distinguish between a state race won by two votes or won by two hundred votes.

Capitol that followed it), but it is important to establish whether such outcomes were flukes or the results we should routinely expect of the EC system. Figures 2 and 3 and Table 2 reveal that a high risk of close elections (and their consequences) is built into the US electoral system.

These findings are remarkably similar across the several models of the EC we examine. But the close correspondence in Figure 2 may give an impression of more precision in these results than is warranted. It is possible to construct election models where the risk at each threshold m is very high, but somewhat lower than in Figure 2. In the Appendix we repeat the exercise of Figure 2 for a catalogue of about 100 alternative election models for the modern period (1988–2016) produced by GST, along with uncorrected versions of the Gelman, Heidemanns and Morris (2020) and Silver (2020) models and one of the earliest presidential simulation models for which simulation draws were obtainable (Gelman, Silver and Edlin, 2012). (See Figures A1 and A5.) These models predict or describe different events (e.g., different election years) and make different structural assumptions, such as about the correlation structure that links cross-state shocks in an election year. The patterns are more differentiated than in Figure 2, but the risk of a close outcome remains high across all variants.

4.2 Closeness Risk Over Presidential Voting History

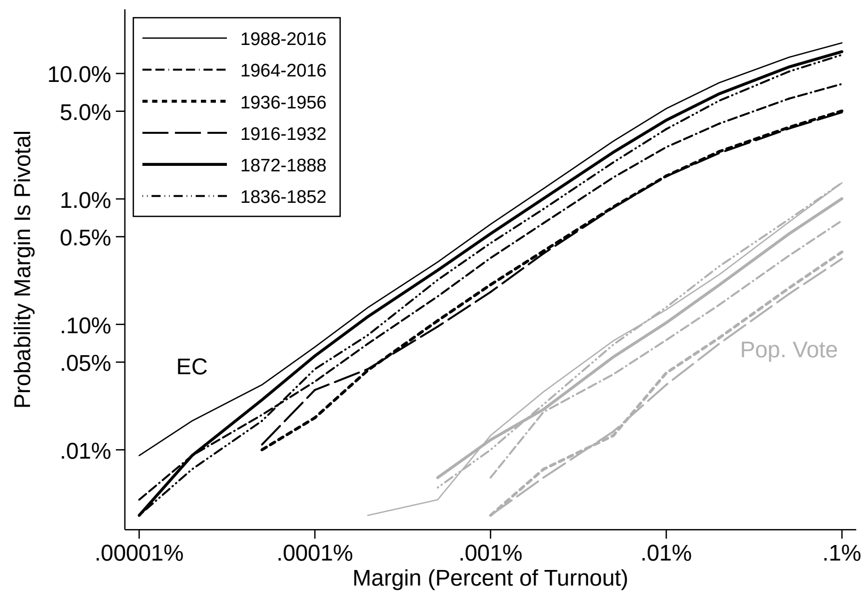
In Figure 4 we repeat the analysis of Figure 2 for earlier time periods, adding in GST's election models for the Antebellum (1836–1852), Post-Reconstruction (1872–1888), and twentieth century and later periods (1916–1932, 1936–1956, 1964–2016). Each period aligns with a *party system*, characterized by fixed parties and fairly stable political geographies.

Figure 4, Panel (A) shows that the risk of a close, disputable election has been high throughout US history, though not uniformly so.[28] In the Antebellum and post-Reconstruction periods, the probability of a narrow outcome that could be reversed by changing a small number of votes was nearly as large as today. And like today, close elections were hotly contested and a source of crisis. The early and mid-twentieth century were about half as likely to result in a close outcome as the late twentieth century. In both 1916–1932 and 1936–1956, the chances that a presidential race was decided by 0.01% of the vote was about 1.5%, compared to over 5% today. That is unsurprising, as these periods were characterized by many landslide elections, and the models fit to these periods reflect that. For example,

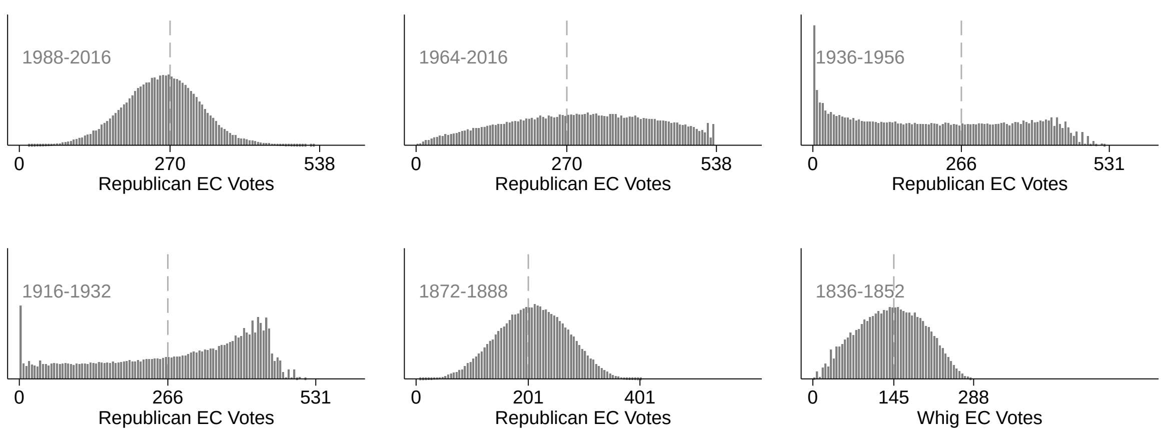
[28] We remove the secondary labels on the horizontal axis that indicate margins in terms of votes, because closeness as a fraction of turnout corresponds to very different numbers of citizen votes across periods. In the 2020 election, 0.01% of the two-party vote corresponded to about 15,000 ballots. In 1836, 0.01% was just 150 ballots.

Figure 4: Close Election Risk Has Been High over the History of Presidential Voting

(A) Probabilities of Election Closeness over US History



(B) Electoral Vote Probability Distributions over US History



Note: Figure plots the probability that a national election outcome would be reversible by changing a small number of votes, as in Figure 2. In Panel (A), black lines indicate the probability of a *close* EC outcome, and gray lines plot the probabilities of popular vote margins from the same models, for reference. Each line represents a model from Geruso, Spears and Talesara (2020) describing a different time period, as indicated. For the same set of models, panel (B) plots the distribution of the expected Republican (or, for the earliest period, Whig) electoral vote total. For the 1836-1852 panel, the GST dataset includes a total of 288 EC votes at stake. This excludes South Carolina's 8 ballots, because during this period there was no citizen vote for president in South Carolina.

the four races won by FDR in the early twentieth century were decided by an average margin of 15 percentage points of the popular vote, including a margin of 24 points in 1936, the largest over the past century. Even though these races also involved large electoral vote margins—in 1936, FDR won 98% of the electoral votes—Figure 4 shows that the expected popular votes margins in every period were orders of magnitude larger in expectation than the number of votes needed to reverse the EC outcome.

For reference, Panel (B) of Figure 4 plots the distribution of likely electoral vote outcomes in each period, which are quite different from each other and the modern period in terms of both means and variances.[29] For example, the standard deviation of Republican/Whig share of the popular vote is less than 3.0 percentage points in the 1988-2016 and 1836–1852 models, but more than three times that size in the 1916–1932 and 1936–1956 models. Examining periods of more diffuse election possibilities and during which different political parties were favored—e.g., Republicans in the late twentieth century, Whigs in the mid-1800s—is useful in disentangling how much of the recent experience with close elections is a consequence of the facts of contemporary politics rather than the fundamental properties of the EC's aggregation mechanics.

The spread in the probabilities of narrow election margins across the different periods in Figure 4 is non-negligible, but ultimately not large when compared to the potential model and parameter uncertainty inherent in the models that generate these simulations. For example, the differences between the mid-twentieth century and today in terms of the close-election probabilities in Figure 4 are similar in magnitude to the difference in probabilities between the polling-bias-corrected versions of the Silver (2020) simulation model of the 2020 election and its uncorrected variant. (See Figure A1.) Most importantly, this spread is small relative to the difference in magnitude of these risks when comparing any model of the EC to a single-tier voting system, which we take up next in Section 5.

Figure 4 establishes an important new result: The tendency of the EC to produce narrow elections is an enduring empirical feature, true for every political geography that the US has to date produced. The risk of a close, disputable election has been high for as long as citizens have participated in US presidential elections. This statistical property has persisted in periods characterized by radically different geographic patterns of partisan politics—including the addition of new states and voting territories, new enfranchised voters (non-whites, the poor, women), and the birth and death of US political parties.

[29] The popular vote distributions corresponding to these electoral vote distributions are also quite varied. See Figure A4.

5 Is the EC Uniquely Vulnerable to Close Elections?

5.1 Closeness in Single-Tier Systems

Would a National Popular Vote system produce US presidential elections just as close in expectation as the EC does? The empirical models considered above, which describe statistical properties of presidential elections as politics is organized under the EC system, do not necessarily answer this. But there is considerable evidence on the closeness of outcomes in other single-tier electoral systems in the US and abroad that can materially inform the comparison. In this section, we quantify the risk of close elections in US Senate races, US gubernatorial races, and single-tier presidential elections in the 15 OECD countries that have them. We show that none of these 115 single-tier electoral systems produces probabilities of a close outcome within an order of magnitude of the corresponding estimate for the EC.

Single-tier (popular vote) elections are much simpler to model than a two-tier system like the EC. In a single-tier election, only the vote tallies matter, not the geographic distribution of votes across the tiered voting units (e.g., states) where they are cast, so the vector of information that fully characterizes an election is much smaller. Consider, for illustration, maps A and B in the simple toy model of Table 1; they have different properties in a two-tier system, but the exact same distribution of possible outcomes in a one-tier electoral system. The uncertainty in a single-tier election can be well summarized in two parameters: the mean and variance of possible vote outcomes. Throughout this section, we focus only on the variance: We artificially replace the mean vote shares with 0.5 when examining popular-vote elections, in order to bias our computations in favor of finding that these elections are likely to be close. Thus we ask whether the empirical dispersion of likely outcomes in any observed popular-vote election is narrow enough to produce outcomes as close as the EC in expectation, even when we assume these elections end in a tie in expectation?

The answer is no. Table 3 reports summary statistics that describe two important single-tier electoral contests in the US: governorships and US Senate seats. For these, we use the Leip Election Atlas (Leip, 2019) and include all available data on races since 1960. A useful summary statistic for understanding the likelihood of a close outcome is the standard deviation of the historical vote shares. The table lists, for governor and Senate races in each state (combining Senate seats in a state), the standard deviation of the Republican share of the two-party vote, expressed in percentage points.

Even prior to fitting a model to these data, it is apparent that the spread of historical outcomes in these races is far more dispersed than the EC's distributions of pivotal vote shares. From Table 2 (Panel C), there is a greater than 1-in-3 chance that the EC is decided by less than 0.1 percentage point.[30] In contrast, from Table 3, the standard deviation of voting outcomes for a US governor or US Senate race is typically greater than 10 percentage points.

To more precisely compare the probabilities of close outcomes in US single-tier elections against the EC, we convert the empirical standard deviations reported in Table 3 to probabilities that each state race is decided by various margins (0.1 point, 0.01 point, or 0.001 point) by making two conservative assumptions. First, we demean the vote share outcomes for every elected seat and recenter these at 50-50, so that the expectation of these (mean-shifted) elections is a tie. Ignoring the empirical partisan lean for each seat increases the probability density around a close outcome and allows us to sidestep any objection that the US presidency, to which we will compare, is more likely to center national party politics around a 50-50 Republican-Democrat split than is true of state races.

Second, we fit the observed Republican vote shares to a normal distribution of vote possibilities,[31] rather than the fat-tailed distributions preferred by election forecasters (e.g., Silver, 2016). These assumptions are conservative in the sense that we claim that the EC generates closer outcomes in expectation, and these assumptions overstate the likely closeness of the popular-vote systems they are meant to describe.

Of course, the simple statistical "models" of Senate and governor elections that result from these two assumptions may be poor predictors or descriptors of actual US politics—delivering, for example, the expectation that neither Republicans nor Democrats have an advantage in Texas governor races, and implying that the TX governor is a much closer race than it is, in fact. We use these models only to form conservative, upper bound estimates of the probability of a close outcome in popular voting. Such extremely crude bounds are nonetheless sufficient for our purposes because they turn out to be informative for making comparisons to the more carefully modeled Electoral College, for which the probability of a close result is nonetheless orders of magnitude larger.

Under these assumptions, Table 3 lists the implied probabilities of close outcomes in popular vote state races. The implied probabilities are low. For both the US Senate and US governorships, the upper

[30] That is, 0.1 percent of the two-party vote total.

[31] We also show robustness to fitting the log odds ratio of votes shares, rather than the shares themselves, to a normal, which has the theoretically desirable property that it places no probability weight (however small) on vote shares outside of [0%,100%]. This difference is unimportant. See Table A3.

bound on the probability that a race is decided by 0.1% of the vote or less averages about 0.4% across the states. The EC system, in contrast, is expected to deliver an outcome decided by 0.1% of the vote with a greater than 18% chance when restricting attention to cases in which the EC outcome could be reversed by a single state's outcome (Panel A), and with a greater than 34% chance if one allows the votes to be split across two or three potentially pivotal states (Panel C). This is true whether one considers the *Economist*, GST, or *FiveThirtyEight* models of the Electoral College system. Comparing the probability of closeness at other margins (0.01 or 0.001 percent of turnout) yields a similar contrast: One or two orders of magnitude separate the probabilities of a close race in popular voting versus in the EC.

These findings are consistent with the small sample of presidential elections in recent history. Four of the six presidential races since 2000 have been decided by less than 0.1% of the vote. One of these, the 2000 race was decided by less than 0.001% of the vote. The recent history of voting Senate and governor races indicates that such tight margins are deeply unlikely in those statewide contests.

What about presidential elections? Perhaps there is something different about *national* elections that is important for generating close outcomes. In Table 4, we consider the typical closeness of popular-vote presidential races in other OECD countries. Because OECD countries do not all exhibit stable two party systems, rather than reporting the standard deviation of a fixed party's vote share, we report the standard deviation of the winning party's vote share (among the top two parties).[32] Using the winning party's vote share in these calculations, rather than a fixed party, adds another layer of conservatism to our claim that popular votes are less likely to be close.[33]

In Table 4, we report these statistics for every OECD nation that elects its chief of state ("president") via popular vote. We include the outcome of every popular, winner-take-all election for a chief of state office that has taken place in a current OECD member country from 1988 to 2000. The table also lists the implied probabilities of close outcomes at various thresholds in these popular vote races, using the same parametric assumptions as in Table 3 for US races (50-50 re-centering and fitting a normal distribution with the empirical standard deviation).

Table 4 shows that the presidential elections of OECD countries in the past 30 years have shown

[32] In some cases, a country uses a two-party runoff, which we use as the two-party election.

[33] Calculating the standard deviation in terms of the winner's vote share, rather than over a fixed party (that could win or lose) tends to understate the variance of the potential outcome—analagous to the reduction in variance that would result from folding a mean-zero normal. Like the other assumptions, this tends to generate an overestimate of the probability of a close race in a popular vote system.

Table 3: EC *Closeness* Compared to US Statewide Single-Tier Elections

	Republican Share Statistics: Empirical Mean in points (ignored)	Republican Share Statistics: Empirical Standard Deviation in points	Probability of a Race Decided by Less Than: 0.1 points	Probability of a Race Decided by Less Than: 0.01 points	Probability of a Race Decided by Less Than: 0.001 points
President: Electoral College					
Gelman et al. (2020)			**37.5%**	**5.9%**	**0.7%**
Geruso et al. (2020)			**34.7%**	**6.5%**	**0.6%**
Silver (2020)			**38.5%**	**6.6%**	**0.7%**
US Senate					
Senate Average	**49.5**	**12.3**	**0.38%**	**0.04%**	**0.00%**
Alabama	49.3	20.9	0.19%	0.02%	0.00%
Alaska	59.6	17.4	0.23%	0.02%	0.00%
Arizona	57.3	13.5	0.30%	0.03%	0.00%
Arkansas	45.5	25.7	0.16%	0.02%	0.00%
California	41.3	14.9	0.27%	0.03%	0.00%
Colorado	52.1	6.4	0.62%	0.06%	0.01%
Connecticut	41.8	8.1	0.49%	0.05%	0.00%
Delaware	45.4	9.7	0.41%	0.04%	0.00%
Florida	47.3	11.2	0.36%	0.04%	0.00%
Georgia	39.6	20.5	0.19%	0.02%	0.00%
Hawaii	27.3	8.3	0.48%	0.05%	0.00%
Idaho	62.4	14.3	0.28%	0.03%	0.00%
Illinois	44.4	8.7	0.46%	0.05%	0.00%
Indiana	56.3	13.6	0.29%	0.03%	0.00%
Iowa	54.7	10.0	0.40%	0.04%	0.00%
Kansas	69.0	13.2	0.30%	0.03%	0.00%
Kentucky	49.4	10.3	0.39%	0.04%	0.00%
Louisiana	45.0	19.5	0.20%	0.02%	0.00%
Maine	54.7	15.2	0.26%	0.03%	0.00%
Maryland	40.6	9.8	0.41%	0.04%	0.00%
Massachusetts	38.4	14.2	0.28%	0.03%	0.00%
Michigan	44.2	6.5	0.62%	0.06%	0.01%
Minnesota	45.8	8.0	0.50%	0.05%	0.00%
Mississippi	62.4	20.4	0.20%	0.02%	0.00%
Missouri	51.0	7.6	0.53%	0.05%	0.01%
Montana	45.2	9.1	0.44%	0.04%	0.00%
Nebraska	52.4	14.1	0.28%	0.03%	0.00%
Nevada	47.3	7.2	0.55%	0.06%	0.01%
New Hampshire	55.0	8.3	0.48%	0.05%	0.00%
New Jersey	44.7	5.0	0.81%	0.08%	0.01%
New Mexico	49.6	12.2	0.33%	0.03%	0.00%
New York	39.4	11.2	0.36%	0.04%	0.00%
North Carolina	49.5	5.3	0.75%	0.07%	0.01%
North Dakota	47.9	15.3	0.26%	0.03%	0.00%
Ohio	48.7	10.1	0.40%	0.04%	0.00%
Oklahoma	56.2	16.1	0.25%	0.02%	0.00%
Oregon	48.2	9.9	0.40%	0.04%	0.00%
Pennsylvania	52.8	6.2	0.65%	0.06%	0.01%
Rhode Island	38.2	13.1	0.30%	0.03%	0.00%
South Carolina	53.2	12.2	0.33%	0.03%	0.00%
South Dakota	54.8	15.6	0.26%	0.03%	0.00%
Tennessee	56.9	12.0	0.33%	0.03%	0.00%
Texas	56.4	7.4	0.54%	0.05%	0.01%
Utah	62.7	9.7	0.41%	0.04%	0.00%
Vermont	51.4	19.7	0.20%	0.02%	0.00%
Virginia	50.5	22.9	0.17%	0.02%	0.00%
Washington	41.6	10.6	0.37%	0.04%	0.00%
West Virginia	37.2	15.0	0.27%	0.03%	0.00%
Wisconsin	42.5	8.1	0.49%	0.05%	0.00%
Wyoming	69.2	9.1	0.44%	0.04%	0.00%

	Republican Share Statistics: Empirical Mean in points (ignored)	Republican Share Statistics: Empirical Standard Deviation in points	Probability of a Race Decided by Less Than: 0.1 points	Probability of a Race Decided by Less Than: 0.01 points	Probability of a Race Decided by Less Than: 0.001 points
President: Electoral College					
Gelman et al. (2020)			**37.5%**	**5.9%**	**0.7%**
Geruso et al. (2020)			**34.7%**	**6.5%**	**0.6%**
Silver (2020)			**38.5%**	**6.6%**	**0.7%**
US Governor					
Governor Average	**49.6**	**10.4**	**0.42%**	**0.04%**	**0.00%**
Alabama	40.6	20.7	0.19%	0.02%	0.00%
Alaska	49.6	9.0	0.44%	0.04%	0.00%
Arizona	51.6	7.4	0.54%	0.05%	0.01%
Arkansas	44.4	12.5	0.32%	0.03%	0.00%
California	49.8	8.1	0.49%	0.05%	0.00%
Colorado	44.7	11.4	0.35%	0.04%	0.00%
Connecticut	49.8	7.6	0.52%	0.05%	0.01%
Delaware	46.4	12.9	0.31%	0.03%	0.00%
Florida	47.7	6.7	0.60%	0.06%	0.01%
Georgia	41.4	15.9	0.25%	0.03%	0.00%
Hawaii	45.0	9.0	0.44%	0.04%	0.00%
Idaho	55.0	10.3	0.39%	0.04%	0.00%
Illinois	53.6	11.6	0.35%	0.03%	0.00%
Indiana	52.4	7.0	0.57%	0.06%	0.01%
Iowa	52.0	7.3	0.54%	0.05%	0.01%
Kansas	51.4	10.6	0.38%	0.04%	0.00%
Kentucky	42.3	8.5	0.47%	0.05%	0.00%
Louisiana	54.5	15.5	0.26%	0.03%	0.00%
Maine	49.5	7.2	0.56%	0.06%	0.01%
Maryland	42.5	10.5	0.38%	0.04%	0.00%
Massachusetts	50.8	10.1	0.39%	0.04%	0.00%
Michigan	51.6	7.8	0.51%	0.05%	0.01%
Minnesota	48.8	7.5	0.54%	0.05%	0.01%
Mississippi	51.2	8.9	0.45%	0.04%	0.00%
Missouri	48.6	8.0	0.50%	0.05%	0.00%
Montana	48.5	11.4	0.35%	0.03%	0.00%
Nebraska	53.3	13.0	0.31%	0.03%	0.00%
Nevada	47.9	15.9	0.25%	0.03%	0.00%
New Hampshire	50.0	11.1	0.36%	0.04%	0.00%
New Jersey	49.0	10.1	0.40%	0.04%	0.00%
New Mexico	48.0	6.6	0.61%	0.06%	0.01%
New York	46.5	12.3	0.32%	0.03%	0.00%
North Carolina	46.8	5.9	0.68%	0.07%	0.01%
North Dakota	56.8	12.7	0.31%	0.03%	0.00%
Ohio	53.0	10.3	0.39%	0.04%	0.00%
Oklahoma	50.5	9.8	0.41%	0.04%	0.00%
Oregon	47.1	7.1	0.56%	0.06%	0.01%
Pennsylvania	48.5	7.9	0.50%	0.05%	0.01%
Rhode Island	45.8	12.8	0.31%	0.03%	0.00%
South Carolina	50.7	9.7	0.41%	0.04%	0.00%
South Dakota	58.8	7.7	0.52%	0.05%	0.01%
Tennessee	54.1	13.4	0.30%	0.03%	0.00%
Texas	53.7	11.2	0.36%	0.04%	0.00%
Utah	56.9	14.6	0.27%	0.03%	0.00%
Vermont	50.2	12.6	0.32%	0.03%	0.00%
Virginia	52.9	14.5	0.28%	0.03%	0.00%
Washington	47.9	5.6	0.71%	0.07%	0.01%
West Virginia	46.0	9.3	0.43%	0.04%	0.00%
Wisconsin	51.9	6.1	0.65%	0.07%	0.01%
Wyoming	51.7	14.6	0.27%	0.03%	0.00%

Note: Table reports summary statistics that describe US governorships and US Senate seats, along with the implied probabilities of a close race, by election type and state. Within a state, data on the two Senate seat races are combined to estimate a single statistic. The probability of a close race is calculated by fitting the empirical standard deviation to a normal distribution centered on a 50-50 vote, as described in the text. Statistics related to the US Electoral College are repeated from Table 2, Panel C for reference. The rows labelled Senate Average and Governor Average report the unweighted means of the indicated statistics across all states. Data for this table come from the Leip Election Atlas (Leip, 2019) and include all available data on races since 1960.

Table 4: EC *Closeness* Compared to OECD Single-Tier Presidential Elections

	Winning Party Share Statistics		Probability of a Race Decided by Less Than:		
	Empirical Mean (ignored)	Empirical Standard Deviation	0.1 points	0.01 points	0.001 points
President: Electoral College					
Gelman et al. (2020)			**37.5%**	**5.9%**	**0.7%**
Geruso et al. (2020)			**34.7%**	**6.5%**	**0.6%**
Silver (2020)			**38.5%**	**6.6%**	**0.7%**
OECD Countries					
OECD Average	**61.3**	**9.3**	**0.64%**	**0.06%**	**0.01%**
Austria	65.9	15.8	0.25%	0.03%	0.00%
Chile	58.4	7.5	0.53%	0.05%	0.01%
Czech Republic	53.1	2.4	1.64%	0.16%	0.02%
Finland	60.7	13.5	0.30%	0.03%	0.00%
France	60.0	12.1	0.33%	0.03%	0.00%
Iceland	75.3	17.7	0.23%	0.02%	0.00%
Ireland	60.9	7.4	0.54%	0.05%	0.01%
Lithuania	61.4	11.9	0.34%	0.03%	0.00%
Mexico	57.4	8.9	0.45%	0.04%	0.00%
Poland	58.8	11.1	0.36%	0.04%	0.00%
Portugal	68.6	10.0	0.40%	0.04%	0.00%
Slovakia	58.1	1.8	2.26%	0.23%	0.02%
Slovenia	64.8	8.9	0.45%	0.04%	0.00%
South Korea	56.2	6.3	0.63%	0.06%	0.01%
Turkey	60.3	4.1	0.97%	0.10%	0.01%

Note: Table reports summary statistics that describe chief of state elections by popular vote in OECD countries, along with the implied probabilities of a close race for each country. The probability of a close race is calculated by fitting the empirical standard deviation of the winning party's vote share to a normal distribution centered on a 50-50 vote, as described in the text. Statistics related to the US Electoral College are repeated from Table 2, Panel C for reference. OECD Average reports the unweighted means of the indicated statistics across the countries in the table.

spreads of vote outcomes fairly similar to the spreads of outcomes in US Senate and governor elections. Upper bounds on the probability of a race decided by less than 0.1%, 0.01%, and 0.001% of the vote, respectively, are 0.6%, 0.06%, and 0.01%. Comparing to the Electoral College system, the EC is expected to deliver close results at much higher rates: >34%, >5%, and >0.6%, respectively, for the same thresholds (Table 2, Panel C).

In sum, although the US has never elected its president by popular vote, there is significant experience with popular voting in the US and abroad. Popular voting—as it exists in practice in political systems in which parties endogenously respond to the electoral incentives of popular voting—is much less likely to generate close outcomes than is the EC system.

5.2 What Hypothetical NPV Voting Process *Would* Equal the EC in Closeness?

Does the EC *cause* its razor-thin margins, or would US presidential politics generate very narrow wins and losses even under a National Popular Vote? Section 5.1 provided indirect evidence on this question by comparing patterns of popular voting in other types of elections. In this section, we characterize the counterfactual properties of a popular election for US president, such that the NPV system would match the EC in terms of close election risk. We then ask whether such properties are plausible.

Consider a counterfactual NPV system for US presidential elections that, as above, is conservative in the sense of following a normal distribution with an expectation centered at a 50-50 (Republican-Democrat) vote share. Expressing vote shares in percentage points, the distribution is $F \sim N(0, \sigma^2)$, where the mean has been normalized to zero by subtracting 50 points. Given this distribution, it is straightforward to calculate the standard deviation σ that generates the same probability as does the EC of arriving at a voting outcome within margin m. Formally, we find the σ that solves $\int_{m/2\sigma}^{m/2\sigma} \phi(x)dx = Pr.^{EC}(m)$, where ϕ is the standard normal density and $Pr.^{EC}(m)$ is the probability of a within-margin-m result in the EC system, as calculated for Figure 2 and Table 2.[34]

For example, we found above that $Pr.^{EC}(.01\text{pp}) \geq 5\%$ (Table 2, Panel A). That is, the EC generated about a 1-in-20 chance of an outcome decided by 15,000 votes or fewer by 2020 turnout. The calibrated NPV distribution that would generate a matching probability of races within this margin has a σ equal

[34] In the integral limits, the margin m is expressed by the range $-m/2$ to $+m/2$. A 0.01% margin is generated by Republican vote shares in the range 49.995 to 50.005 points.

to 0.076 percent of turnout.[35] To emphasize: the calibrated σ is equal to a little over 100,000 votes; we do not mean 7.6 percent.[36]

This implied spread of potential vote outcomes under such a hypothetical NPV system is very small. To put the number in context, Figure 5 plots this popular vote counterfactual ($F \sim N(0, 0.076^2)$) against estimates of the empirical popular vote distributions for US president today (repeating from Figure 1) and against the typical spread of votes in other popular-vote elections in the US and abroad (from Tables 3 and 4).

The figure illustrates that the distribution of vote possibilities in an NPV system would have to be exceedingly narrow—far outside of historical experience with popular voting—in order to match the closeness of the EC system. If the popular vote counterfactual on display in Figure 5 were a correct statistical model summarizing campaign and voter behavior under a National Popular Vote for US president, then experiencing a presidential race in which the winner gained more than 50.5% of the two-party popular vote would be an extreme statistical fluke—about seven standard deviations above the mean, occurring with probability $< 10^{-10}$. In contrast to this prediction, popular-vote elections for US senators, US governors, and OECD presidents routinely have victory margins larger than 50.5%-49.5%.

Although we cannot rule out the abstract possibility that an NPV system for the US presidency would always generate such tight outcomes, we can conclude that it would be inconsistent with all historical experience—and on that basis deeply implausible. National popular voting simply cannot match the EC's actual tendency toward narrow outcomes under plausible assumptions. In this sense, the EC causes close election outcomes, relative to the feasible alternative system.

5.3 Counterfactual: EC Closeness as Cross-State Variance increases

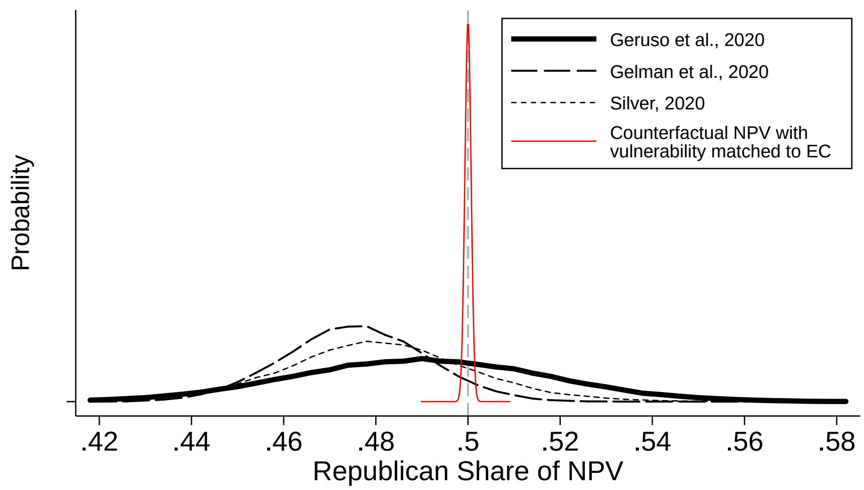
We conclude our comparison between the EC and a popular vote system by returning to an insight from the toy model in Section 2: under certain patterns of geographic polarization, the EC system could in principle deliver results that would not be especially vulnerable to producing close election outcomes. Here, we examine what changes to the partisan alignment of US states could generate elections in which the EC outcome is not particularly close in expectation. Whereas Section 5.2 asked

[35] In particular, we match the estimate for a .01 percentage point margin in the GST column in Panel A of Table 2, which is 5.3%. Of the three models, GST generates the most conservative (lowest) estimates of close-election risk in the EC.

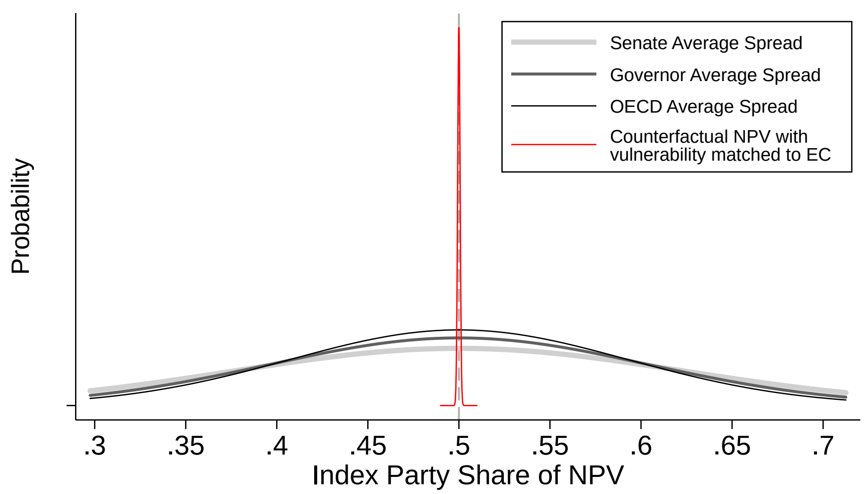
[36] Matching other EC statistics, such as the probability of a race decided by less than .0001%, or .001% of turnout, generates similarly small calibration values for σ. We tabulate these in Appendix Table A5.

Figure 5: A Counterfactual National Popular Vote Distribution that Would Rival EC in *Closeness* Is Implausibly Tight

(A) NPV Counterfactual Compared to the Actual Popular Vote Distribution Generated by the EC System

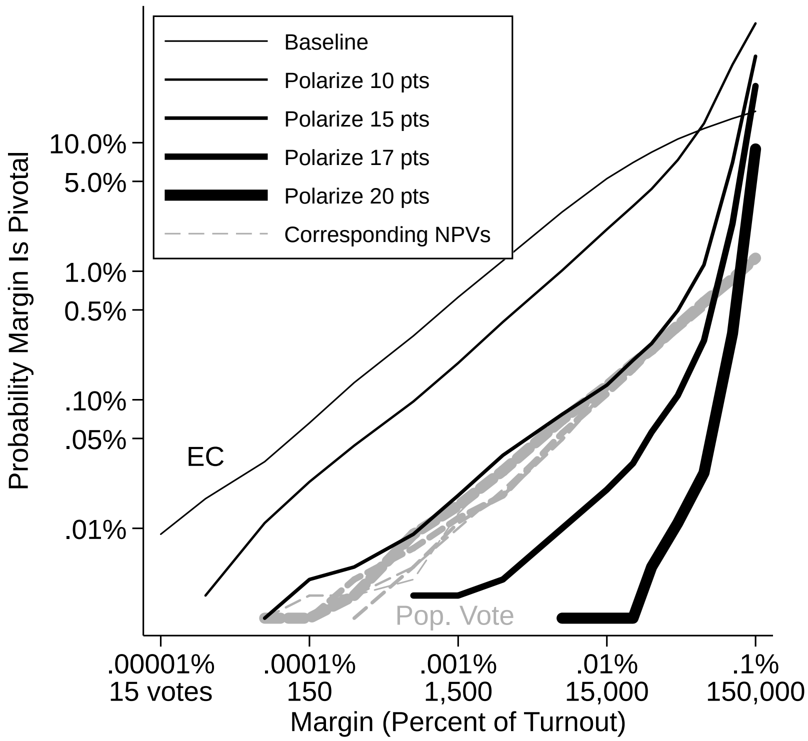


(B) NPV Counterfactual Compared to the Spreads in Actual Popular Vote Outcomes in Popular Vote Systems



Note: Figure compares several empirical distributions of popular vote outcomes with the distribution generated by a hypothetical NPV for US president that would deliver close outcomes with similar expected frequency to the present EC system. See text for full detail. The height of the counterfactual NPV spike is not drawn to scale with other distributions; the spike height is lower relative to truth in order to preserve some visual detail in the other distributions plotted. In Panel (B), the index party is Republican in US elections and is the winning party in OECD elections.

Figure 6: Extreme Counterfactuals for US Politics Could Reduce the Probability of a Close EC Outcome to Levels Similar to a Popular Vote



Note: Figure plots the probability that a national election outcome would be reversible by changing a small number of votes in a single state, as in Figure 2, but under model adjustments that intensify geographic polarization. Starting from the baseline of the Geruso, Spears and Talesara (2020) modern period model (1988-2016), the exercise shifts each state's partisan lean further away from 50-50 to generate alternative simulated election results. For example, under "polarize by 10 points" scenarios, we adjust the state expectation of the Republican vote share upward by 10 percentage points in states with an expected Republican share greater than 0.5 in the baseline model and downward by 10 percentage points elsewhere. Solid black lines indicate the closeness of the EC outcome. Dashed gray lines plot the closeness of the corresponding popular vote result, for reference.

how national politics would have to change to deliver close outcomes in a hypothetical NPV system with high probability, this section asks how state politics would have to change to deliver close outcomes in the present EC system with low probability.

We construct counterfactuals that take as a starting point simulation draws from GST. Given these simulations, we shift each state's partisan lean further away from 50-50. For example, under "polarize by 10 points" scenarios, we adjust the state expectation of the Republican vote share upward by 10 percentage points in states with an expected Republican share greater than 0.5 and downward by 10 percentage points elsewhere, bounding the resulting share expectations at .01 and .99 while retaining all other features of the baseline model. The exercise is analogous to the toy model comparison between maps A and B in Table 1. Although we do not impose it, this process has almost no impact on the mean and variance of the national popular vote.[37]

Figure 6 plots the original simulation results alongside the new results based on these polarization adjustments. The figure shows that counterfactual increases to polarization substantially decreases the risk of a close Electoral College outcome for most margins of closeness, without changing the (irrelevant) popular vote underlying it. For a 15-percentage-point simulated polarization increase, the probability of close election outcome drops by about two orders of magnitude, and the EC's *closeness* is indistinguishable from that of the national popular vote generated under the EC, at least for margins up to about 0.01% of turnout.

It is important to understand that despite contemporary commentators labelling the current political environment as "highly polarized" (Klein, 2020), the type and size of the increase of polarization considered in Figure 6 is considerably more extreme than what is typically meant by "highly polarized" today. Indeed, the 15-point scenario definitionally adds 30-point gap in partisan lean between a pair of states with essentially no actual partisan lean—that is, states that fall just slightly on either side of a 50-50 expectation.

Of course, there are many alternative ways to model increasing polarization or other changes to the geographic distribution of political party support. The aim of Figure 6 is merely to reinforce the observation that it is not *logically necessary* that an Electoral College produces more narrow-election risk than a National Popular Vote. The toy model of Section 2 made the argument from first principles; Figure 6 demonstrates the idea in an estimated model of the EC that incorporates the messy empirical

[37] See Table A6 for details on the mean and variance of the national vote for each counterfactual in Figure 6.

realities of 51 voting units that vary in size, alignment, and apportionment asymmetries.

6 Partisan Asymmetry and Correlated Risks

We conclude with a brief discussion of partisan asymmetry in close election risk. It is well-understood that US political parties perform asymmetrically in the US Electoral College system. It is therefore possible that partisan asymmetry could be relevant in characterizing close election outcomes. In this section, we investigate the asymmetry of close elections.

Asymmetry in the US Electoral College is typically characterized by some measure of the distance between the electoral vote and the popular vote (see, e.g., Miller, 2012). Because the presidency is decided by electoral votes, and because, definitionally, an election is close if changing a small number of citizen votes in one or several states would change the electoral vote winner, close elections tend to be symmetric in electoral votes. For the GST election model of the modern period, for example, the mean and median Republican electoral vote tally in the set of close outcomes is 269 or 270 electoral votes.

But what about the popular vote majority? Geruso, Spears and Talesara (2020) studied mismatches between the popular vote and electoral outcome—*inversions*. They showed that a nearly even split of the popular vote often yields a far-from-even split in the parties' chances of winning the presidency. Given that finding, a close election in the EC system is unlikely to coincide with a popular vote outcome around 50-50. This is visible in Figure 7, which traces, as a function of the closeness of the election outcome, both the probability of a Republican Electoral College victory and the probability of a Republican popular vote majority. For any margin m in the figure, the probability of an EC win is about 50-50 (Republican-Democrat), because the sample is restricted to such close EC outcomes. The probability of a Republican popular vote majority in these same races ranges from about 10% for very narrow EC outcomes decided by a handful of votes to about 30% for the largest margins considered.[38]

Consider an EC margin of $<$5,000 votes, which is less than the number of mail ballots typically disqualified in a large swing state in recent presidential elections (U.S. Election Assistance Commission, 2017). Conditional on this circumstance, there is a 14% chance of a Republican EC victory combined with a Republican popular vote majority and a 35% chance of a Republican EC victory combined with

[38] This statistical relationship between closeness and inversions is an implication of the main results in Figure 2, which showed that the minimum number of votes necessary to reverse an EC winner will usually be smaller than the minimum number of votes necessary to reverse the candidate receiving the popular vote majority.

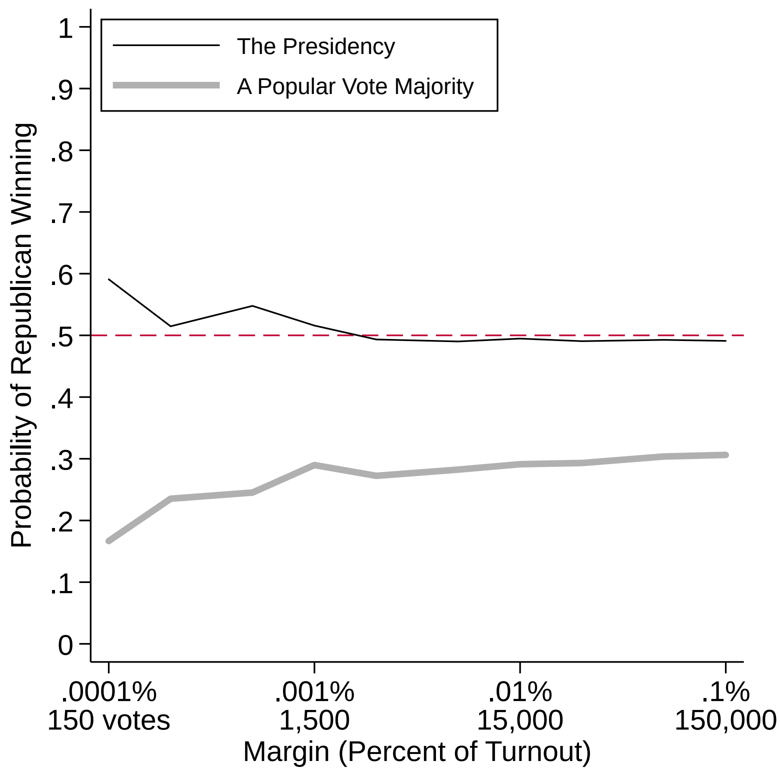
a Republican popular vote minority.[39] In other words, conditional on the outcome being so close, the Republican candidate is more than twice as likely to win in an inversion as they are to win both the electoral and popular vote. This pattern holds in the forecasting election models from Gelman, Heidemanns and Morris (2020) and Silver (2020) as well. (See Figure A6).

Because the patterns of partisan asymmetry in the Electoral College vary significantly across historical periods and even election-to-election (Erikson, Sigman and Yao, 2020), the party likely to win via inversions varies. And so, too, does the asymmetry in close election risk. We show this in Figure A7, which reproduces Figure 7 for periods back to 1836. The figure shows that in the late nineteenth and early twentieth centuries, narrow elections were disproportionately likely to correspond to Democrats winning the presidency and losing the popular vote.

In sum, a close EC outcome is either an inversion or an election that was nearly an inversion—because changing a small number of votes reverses the EC outcome, but not the popular vote majority winner. Thus, the risks of illegitimacy and crisis that result from close elections and the risks of illegitimacy and crisis that result from electoral inversions (see, e.g., Carey et al., 2020) are not independent. These events are disproportionately likely to coincide—as they did in 1876 and in 2000, and as they nearly did in 2016 and 2020.

[39] Analogously: Conditional on an EC margin of $<$5,000 votes, there is a 38% chance of a Democratic EC victory combined with a Democratic popular vote majority and a 13% chance of a Democratic EC victory combined with a Democratic popular vote minority.

Figure 7: Partisan Asymmetry in Narrow Election Outcomes Generates Correlated Risks



Note: Figure plots the probabilities that the Republican candidate wins the presidency and that the Republican candidate secures a popular vote majority. Statistics are conditional on a race decided in the EC system by a margin indicated along the horizontal axis. Data are from the GST model (1988-2016). The figure shows that close presidential outcomes in the modern period are especially likely to be inversions.

7 Conclusion

Extremely narrow election outcomes—such as would be reversed by a few thousand votes—have been claimed in the prior literature to have normatively disadvantageous properties. They may incentivize litigation or recounts; they may be resolved by courts or administrators rather than voters; and they may reduce the legitimacy, peacefulness, or predictability of the transfer of power. Whereas prior work in political science, law, and history has focused on the consequences of close outcomes *if they occur*, we quantify the frequency with which these events are likely to occur. And we further determine the extent to which the voting system itself causes the tendency toward close election outcomes.

We show that, in principle, narrow outcomes could be more likely under the Electoral College or under a National Popular Vote. We show that, in fact, such outcomes *are* more likely in US presidential elections under the Electoral College than would be the case with a National Popular Vote. We demonstrate this with a new application of statistical and forecasting models from several sources. This finding holds over diverse periods of US history, which span different sets of states, parties, and electoral maps. The consistency of this fact over time suggests that a National Popular Vote could significantly reduce the risks of extremely narrow, disputable election outcomes in the future.

Of course, the facts we establish about close election risk exist alongside other considerations of the desirable properties of voting systems, such as whose interests an electoral system should represent. This is a contentious issue in the context of the US Electoral College system. We provide new insights on one important consideration (closeness), without denying the importance of other considerations in the perennial debate over how the US presidential voting should be organized.

Our findings on the statistical tendency of the EC to deliver close election results may serve to clarify thinking on the risk of perceived illegitimacy and litigation following a presidential race. Writing about the 2000 presidential election, Lempert (2016) suggested that the tractability of quarantining recounts to specific states is an advantage of the Electoral College: "Whoever won, Bush or Gore, it was going to be by a hairsbreadth. Because of the Electoral College, we did not have to recount the whole nation. Instead we could focus on a more manageable task—recounting the state of Florida." This argument is built on an empirically mistaken assumption. The error here is that although hypothetical national recount might have been difficult, it would not have been needed and would not have occurred. In statistical expectation, only the Electoral College generates hairsbreadth outcomes.

References

Banducci, Susan A, and Jeffrey A Karp. 2003. "How elections change the way citizens view the political system: Campaigns, media effects and electoral outcomes in comparative perspective." *British Journal of Political Science*, 33(3): 443–467.

Banzhaf III, John F. 1964. "Weighted voting doesn't work: A mathematical analysis." *Rutgers L. Rev.*, 19: 317.

Birch, Sarah. 2008. "Electoral institutions and popular confidence in electoral processes: A cross-national analysis." *Electoral Studies*, 27(2): 305–320.

Carey, John, Gretchen Helmke, Brendan Nyhan, Mitchell Sanders, Susan Stokes, and Shun Yamaya. 2020. "The Effect of Electoral Inversions on Democratic Legitimacy: Evidence from the United States."

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 Na'ama Shenhav. 2020. "A Century of the American Woman Voter: Sex Gaps in Political Participation, Preferences, and Partisanship since Women's Enfranchisement." *Journal of Economic Perspectives*, 34(2): 24–48.

Chamberlain, Gary, and Michael Rothschild. 1981. "A note on the probability of casting a decisive vote." *Journal of Economic Theory*, 25(1): 152–162.

CNN. 2021. "CNN Presidential Election Poll." *SSRS Research*.

Cox, Gary W, and Michael C Munger. 1989. "Closeness, expenditures, and turnout in the 1982 US House elections." *American Political Science Review*, 83(1): 217–231.

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

Erikson, Robert S., Karl Sigman, and Linan Yao. 2020. "Electoral College bias and the 2020 presidential election." *Proceedings of the National Academy of Sciences*, 117(45): 27940–27944.

Gallup. 2001. "Seven out of 10 Americans Accept Bush as Legitimate President."

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, Jonathan N Katz, and Joseph Bafumi. 2004. "Standard voting power indexes do not work: an empirical analysis." *British Journal of Political Science*, 34(4): 657–674.

Gelman, Andrew, Merlin Heidemanns, and G. Elliott Morris. 2020. "State and national presidential election forecasting model." https://github.com/TheEconomist/us-potus-model.

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

Gerber, Alan, Mitchell Hoffman, John Morgan, and Collin Raymond. 2020. "One in a million: Field experiments on perceived closeness of the election and voter turnout." *American Economic Journal: Applied Economics*, 12(3): 287–325.

Geruso, Michael, Dean Spears, and Ishaana Talesara. 2020. "Inversions in US Presidential Elections: 1836-2016." *Forthcoming: American Economic Journal: Applied Economics*.

Grofman, Bernard, and Scott L Feld. 2005. "Thinking about the political impacts of the Electoral College." *Public Choice*, 123(1-2): 1–18.

Hasen, Richard L. 2005. "Beyond the margin of litigation: reforming US election administration to avoid electoral meltdown." *Wash. & Lee L. Rev.*, 62: 937.

Hasen, Richard L. 2020. *Election Meltdown.* Yale University Press.

Haworth, Paul Leland. 1906. *The Hayes-Tilden Disputed Presidential Election of 1876.* Burrows Brothers.

Hirsch, Alan. 2020. *A Short History of Presidential Election Crises:(and how to Prevent the Next One).* City Lights Books.

Holt, Michael. 2009. *By One Vote: The Disputed Presidential Election of 1876.* Taylor & Francis.

Klein, Ezra. 2020. *Why we're polarized.* Simon and Schuster.

Koza, John R, Barry Fadem, Mark Grueskin, Michael S Mandell, Robert Richie, Joseph F Zimmerman, John B Anderson, Birch Bayh, John Buchanan, and Tom Campbell. 2013. *Every Vote Equal: A State-Based Plan for Electing the President by National Popular Vote.* . 4th ed., National Popular Vote Press.

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. 2019. "David Leip's Atlas of U.S. Elections."

Lempert, Richard. 2016. "Two cheers for the Electoral College: Reasons not to abolish it." *Brookings*.

Merlin, Vincent, and Jack Nagel. 2020. "The U.S. Electoral College and the Probability of Disputable Outcomes under Direct and Indirect Elections."

Miller, Nicholas R. 2012. "Why the Electoral College is good for political science (and public choice)." *Public Choice*, 150(1-2): 1–25.

Monmouth University Poll. 2021. "National: More Americans Happy About Trump Loss than Biden Win."

Posner, Richard A. 2001. *Breaking the Deadlock.* Princeton University Press.

Przeworski, Adam. 2018. *Why bother with elections?* John Wiley & Sons.

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

Rapoport, David C, and Leonard Weinberg. 2000. "Elections and violence." *Terrorism and Political Violence*, 12(3-4): 15–50.

Sances, Michael W, and Charles Stewart III. 2015. "Partisanship and confidence in the vote count: Evidence from US national elections since 2000." *Electoral Studies*, 40: 176–188.

Shapley, Lloyd S, and Martin Shubik. 1954. "A method for evaluating the distribution of power in a committee system." *American political science review*, 48(3): 787–792.

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.

Silver, Nate. 2020. "FiveThirtyEight 2020 General Election Forecast." https://projects.fivethirtyeight.com/2020-election-forecast/.

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

U.S. Election Assistance Commission. 2017. "The Election Administration and Voting Survey: 2016 Comprehensive Report." Report to the 115th Congress.

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

Weatherford, M Stephen. 1992. "Measuring political legitimacy." *American Political Science Review*, 86(1): 149–166.

Wright, John R. 2009. "Pivotal states in the Electoral College, 1880 to 2004." *Public Choice*, 139(1-2): 21–37.

ONLINE APPENDIX for "The Risk of Narrow, Disputable Results in the U.S. Electoral College: 1836-2020"

by Geruso and Spears

A Computational robustness check of illustrative model without simplifying assumptions

Section 2 presented an illustrative model of the fact that, in principle, either a two-stage or a one-stage election system could be *closer*, on average and in the sense of this paper, depending on the partisan polarization of states, modeled here as the state-specific expectations for the two-party vote share. In order to present simple analytical expressions, Section 2 made two simplifying assumptions: that the state with the middle mean was always the tipping point state in the second-stage and that vote shares within states were normally distributed, which permitted the possibility that vote shares could be outside of $[0,1]$.

Here we present Monte Carlo simulations of a version of the model without those assumptions and reach the same conclusion. We let the log-odds of each state vote share be normally distributed around a mean μ_s^ℓ such that $\mu_s^\ell = \ln\left(\frac{\mu_s}{1-\mu_s}\right)$, where μ_s, for each state and map, is just as in Section 2. Each of these normal distributions has a standard deviation of 0.202 logit points, which would correspond to the 5 percentage points used in the analytical model for even odds.

We compute 1 million Monte Carlo simulations, which include cases where the order of the states changes. Table A1 presents the results, which are quantitatively similar and substantially identical to those in Table 1 for the simpler model.

Table A1: Disputed Election Risk in the EC and NPV: Monte Carlo Robustness Checks

map	states' sizes	state mean outcomes $\bar{x}_1$	$\bar{x}_2$	$\bar{x}_3$	NPV, $\mathbb{E}[\bar{x}]$	mean number of votes needed to reverse outcome if EC system, $\mathbb{E}[m^{EC}]$	mean number of votes needed to reverse outcome if NPV system, $\mathbb{E}[m^{NPV}]$
A	1,000	35.1%	53.0%	64.9%	51.0%	<u>46.9</u>	69.5
B	1,000	20.2%	62.9%	69.8%	51.0%	129.0	<u>61.5</u>

Table A2: Probabilities of Close Outcomes Over History (Tabulations from Figure 4)

	1964-2016		1936-1956		1916-1932		1872-1888		1836-1852	
Margin in Points Shares*100)	Margin in Votes, Assuming 150M Two-Party Turnout	Probability	Margin in Votes, Assuming 60M Two-Party Turnout	Probability	Margin in Votes, Assuming 40M Two-Party Turnout	Probability	Margin in Votes, Assuming 10M Two-Party Turnout	Probability	Margin in Votes, Assuming 3M Two-Party Turnout	Probability
0.00001	15	0.00%	6	0.00%	4	0.00%	1	0.00%	0	0.00%
0.00002	30	0.01%	12	0.00%	8	0.00%	2	0.01%	1	0.01%
0.00005	75	0.02%	30	0.01%	20	0.01%	5	0.03%	2	0.02%
0.0001	150	0.04%	60	0.02%	40	0.03%	10	0.06%	3	0.04%
0.0002	300	0.07%	120	0.04%	80	0.04%	20	0.12%	6	0.08%
0.0005	750	0.17%	300	0.11%	200	0.10%	50	0.27%	15	0.23%
0.001	1,500	0.34%	600	0.21%	400	0.18%	100	0.53%	30	0.44%
0.002	3,000	1%	1,200	0%	800	0%	200	1%	60	1%
0.005	7,500	1%	3,000	1%	2,000	1%	500	2%	150	2%
0.01	15,000	3%	6,000	2%	4,000	2%	1,000	4%	300	4%
0.02	30,000	4%	12,000	2%	8,000	2%	2,000	7%	600	6%
0.05	75,000	6%	30,000	4%	20,000	4%	5,000	11%	1,500	10%
0.1	150,000	8%	60,000	5%	40,000	5%	10,000	15%	3,000	14%

Note: Table lists probabilities that election outcomes could be reversible by a small number of votes over various periods in US history. Tabulations correspond to the results in Figure 4. See corresponding figure notes for additional detail.

Table A3: Comparative Closeness in US Single-Tier Elections: Log Odds Analog to Table 3

	Republican Share Statistics		Probability of a Race Decided by Less Than:		
	Empirical Mean in log odds (ignored)	Empirical Standard Deviation in log odds	0.1 points	0.01 points	0.001 points
President: Electoral College					
Gelman et al. (2020)			**37.5%**	**5.9%**	**0.7%**
Geruso et al. (2020)			**34.7%**	**6.5%**	**0.6%**
Silver (2020)			**38.5%**	**6.6%**	**0.7%**
US Senate					
Senate Average	**-0.02**	**0.47**	**0.39%**	**0.04%**	**0.00%**
Alabama	-0.04	0.58	0.27%	0.03%	0.00%
Alaska	0.46	0.83	0.19%	0.02%	0.00%
Arizona	0.22	0.41	0.39%	0.04%	0.00%
Arkansas	-0.03	0.35	0.45%	0.05%	0.00%
California	-0.18	0.25	0.63%	0.06%	0.01%
Colorado	0.08	0.26	0.61%	0.06%	0.01%
Connecticut	-0.34	0.34	0.47%	0.05%	0.00%
Delaware	-0.19	0.40	0.40%	0.04%	0.00%
Florida	-0.11	0.47	0.34%	0.03%	0.00%
Georgia	-0.19	0.60	0.27%	0.03%	0.00%
Hawaii	-1.02	0.42	0.38%	0.04%	0.00%
Idaho	0.70	1.21	0.13%	0.01%	0.00%
Illinois	-0.23	0.36	0.44%	0.04%	0.00%
Indiana	0.17	0.38	0.42%	0.04%	0.00%
Iowa	0.20	0.42	0.38%	0.04%	0.00%
Kansas	0.65	0.35	0.46%	0.05%	0.00%
Kentucky	-0.03	0.43	0.37%	0.04%	0.00%
Louisiana	-0.29	0.98	0.16%	0.02%	0.00%
Maine	0.19	0.66	0.24%	0.02%	0.00%
Maryland	-0.39	0.41	0.39%	0.04%	0.00%
Massachusetts	-0.42	0.52	0.31%	0.03%	0.00%
Michigan	-0.24	0.27	0.60%	0.06%	0.01%
Minnesota	-0.18	0.33	0.48%	0.05%	0.00%
Mississippi	0.24	0.80	0.20%	0.02%	0.00%
Missouri	0.04	0.31	0.51%	0.05%	0.01%
Montana	-0.20	0.38	0.42%	0.04%	0.00%
Nebraska	0.11	0.63	0.25%	0.03%	0.00%
Nevada	-0.11	0.30	0.54%	0.05%	0.01%
New Hampshire	0.21	0.35	0.46%	0.05%	0.00%
New Jersey	-0.21	0.20	0.79%	0.08%	0.01%
New Mexico	-0.01	0.51	0.31%	0.03%	0.00%
New York	-0.45	0.47	0.34%	0.03%	0.00%
North Carolina	-0.02	0.22	0.74%	0.07%	0.01%
North Dakota	-0.07	0.67	0.24%	0.02%	0.00%
Ohio	-0.06	0.42	0.38%	0.04%	0.00%
Oklahoma	0.24	0.72	0.22%	0.02%	0.00%
Oregon	-0.08	0.40	0.39%	0.04%	0.00%
Pennsylvania	0.11	0.25	0.63%	0.06%	0.01%
Rhode Island	-0.51	0.57	0.28%	0.03%	0.00%
South Carolina	0.13	0.51	0.31%	0.03%	0.00%
South Dakota	0.11	0.49	0.32%	0.03%	0.00%
Tennessee	0.28	0.50	0.32%	0.03%	0.00%
Texas	0.26	0.30	0.52%	0.05%	0.01%
Utah	0.54	0.41	0.39%	0.04%	0.00%
Vermont	-0.08	0.65	0.24%	0.02%	0.00%
Virginia	-0.11	0.41	0.39%	0.04%	0.00%
Washington	-0.36	0.48	0.33%	0.03%	0.00%
West Virginia	-0.47	0.56	0.29%	0.03%	0.00%
Wisconsin	-0.31	0.34	0.47%	0.05%	0.00%
Wyoming	0.84	0.42	0.38%	0.04%	0.00%

	Republican Share Statistics		Probability of a Race Decided by Less Than:		
	Empirical Mean in log odds (ignored)	Empirical Standard Deviation in log odds	0.1 points	0.01 points	0.001 points
President: Electoral College					
Gelman et al. (2020)			**37.5%**	**5.9%**	**0.7%**
Geruso et al. (2020)			**34.7%**	**6.5%**	**0.6%**
Silver (2020)			**38.5%**	**6.6%**	**0.7%**
US Governor					
Governor Average	**-0.02**	**0.43**	**0.41%**	**0.04%**	**0.00%**
Alabama	-0.39	1.04	0.15%	0.02%	0.00%
Alaska	-0.02	0.39	0.41%	0.04%	0.00%
Arizona	0.06	0.30	0.53%	0.05%	0.01%
Arkansas	-0.24	0.55	0.29%	0.03%	0.00%
California	-0.01	0.33	0.48%	0.05%	0.00%
Colorado	-0.23	0.50	0.32%	0.03%	0.00%
Connecticut	-0.01	0.31	0.51%	0.05%	0.01%
Delaware	-0.15	0.54	0.29%	0.03%	0.00%
Florida	-0.09	0.27	0.59%	0.06%	0.01%
Georgia	-0.25	0.51	0.32%	0.03%	0.00%
Hawaii	-0.20	0.37	0.43%	0.04%	0.00%
Idaho	0.21	0.43	0.37%	0.04%	0.00%
Illinois	0.18	0.57	0.28%	0.03%	0.00%
Indiana	0.10	0.28	0.56%	0.06%	0.01%
Iowa	0.08	0.30	0.53%	0.05%	0.01%
Kansas	0.06	0.45	0.35%	0.04%	0.00%
Kentucky	-0.32	0.36	0.44%	0.04%	0.00%
Louisiana	0.21	0.67	0.24%	0.02%	0.00%
Maine	-0.02	0.29	0.54%	0.05%	0.01%
Maryland	-0.32	0.47	0.34%	0.03%	0.00%
Massachusetts	0.04	0.42	0.38%	0.04%	0.00%
Michigan	0.06	0.32	0.50%	0.05%	0.01%
Minnesota	-0.05	0.31	0.52%	0.05%	0.01%
Mississippi	0.05	0.36	0.44%	0.04%	0.00%
Missouri	-0.06	0.33	0.49%	0.05%	0.00%
Montana	-0.06	0.50	0.32%	0.03%	0.00%
Nebraska	0.14	0.56	0.29%	0.03%	0.00%
Nevada	-0.09	0.70	0.23%	0.02%	0.00%
New Hampshire	0.00	0.47	0.34%	0.03%	0.00%
New Jersey	-0.04	0.42	0.38%	0.04%	0.00%
New Mexico	-0.08	0.27	0.59%	0.06%	0.01%
New York	-0.15	0.51	0.31%	0.03%	0.00%
North Carolina	-0.13	0.24	0.66%	0.07%	0.01%
North Dakota	0.30	0.55	0.29%	0.03%	0.00%
Ohio	0.13	0.43	0.37%	0.04%	0.00%
Oklahoma	0.02	0.40	0.40%	0.04%	0.00%
Oregon	-0.12	0.29	0.54%	0.05%	0.01%
Pennsylvania	-0.06	0.33	0.49%	0.05%	0.00%
Rhode Island	-0.19	0.55	0.29%	0.03%	0.00%
South Carolina	0.03	0.41	0.39%	0.04%	0.00%
South Dakota	0.36	0.33	0.48%	0.05%	0.00%
Tennessee	0.17	0.57	0.28%	0.03%	0.00%
Texas	0.14	0.48	0.33%	0.03%	0.00%
Utah	0.30	0.63	0.25%	0.03%	0.00%
Vermont	0.00	0.54	0.29%	0.03%	0.00%
Virginia	-0.02	0.26	0.60%	0.06%	0.01%
Washington	-0.09	0.23	0.71%	0.07%	0.01%
West Virginia	-0.17	0.39	0.41%	0.04%	0.00%
Wisconsin	0.08	0.25	0.63%	0.06%	0.01%
Wyoming	0.08	0.62	0.26%	0.03%	0.00%

Note: Table replicates Table 3, but fits the log odds ratio of votes shares, rather than the shares themselves, to a normal. This has the theoretically desirable property that it places no probability weight (however small) on vote shares outside of [0%,100%]. See Table 3 notes for additional information.

Table A4: Comparative Closeness in OECD Single-Tier Elections: Log Odds Analog to Table 4

	Winning Party Share Statistics		Probability of a Race Decided by Less Than:		
	Empirical Mean (ignored)	Empirical Standard Deviation	0.1 points	0.01 points	0.001 points
President: Electoral College					
Gelman et al. (2020)			**37.5%**	**5.9%**	**0.7%**
Geruso et al. (2020)			**34.7%**	**6.5%**	**0.6%**
Silver (2020)			**38.5%**	**6.6%**	**0.7%**
OECD Countries					
OECD Average	**0.51**	**0.45**	**0.60%**	**0.06%**	**0.01%**
Austria	0.74	0.78	0.20%	0.02%	0.00%
Chile	0.35	0.32	0.50%	0.05%	0.01%
Czech Republic	0.12	0.10	1.64%	0.16%	0.02%
Finland	0.49	0.66	0.24%	0.02%	0.00%
France	0.44	0.58	0.28%	0.03%	0.00%
Iceland	1.40	1.16	0.14%	0.01%	0.00%
Ireland	0.45	0.32	0.50%	0.05%	0.00%
Lithuania	0.51	0.60	0.27%	0.03%	0.00%
Mexico	0.31	0.38	0.42%	0.04%	0.00%
Poland	0.38	0.49	0.32%	0.03%	0.00%
Portugal	0.82	0.49	0.32%	0.03%	0.00%
Slovakia	0.33	0.07	2.20%	0.22%	0.02%
Slovenia	0.63	0.40	0.40%	0.04%	0.00%
South Korea	0.26	0.26	0.61%	0.06%	0.01%
Turkey	0.42	0.17	0.93%	0.09%	0.01%

Note: Table replicates Table 4, but fits the log odds ratio of votes shares, rather than the shares themselves, to a normal. This has the theoretically desirable property that it places no probability weight (however small) on vote shares outside of [0%,100%]. See Table 4 notes for additional information.

Table A5: Counterfactual National Popular Vote Probability Distributions Calibrated to Match the Expected *Closeness* of the EC System

Margin in Points (Shares*100)	Margin in Votes, Assuming 150M Two-Party Turnout	Probability of Outcome within Margin	Implied Standard Deviation of Normal NPV with Matching Close-Election Probability (in Points)
0.00001	15	0.01%	0.044
0.00002	30	0.02%	0.047
0.00005	75	0.03%	0.060
0.0001	150	0.07%	0.060
0.0002	300	0.14%	0.059
0.0005	750	0.31%	0.064
0.001	1,500	0.63%	0.064
0.002	3,000	1.2%	0.066
0.005	7,500	2.9%	0.069
0.01	15,000	5.3%	0.076
0.02	30,000	8.4%	0.094
0.05	75,000	13.5%	0.147
0.1	150,000	17.5%	0.226

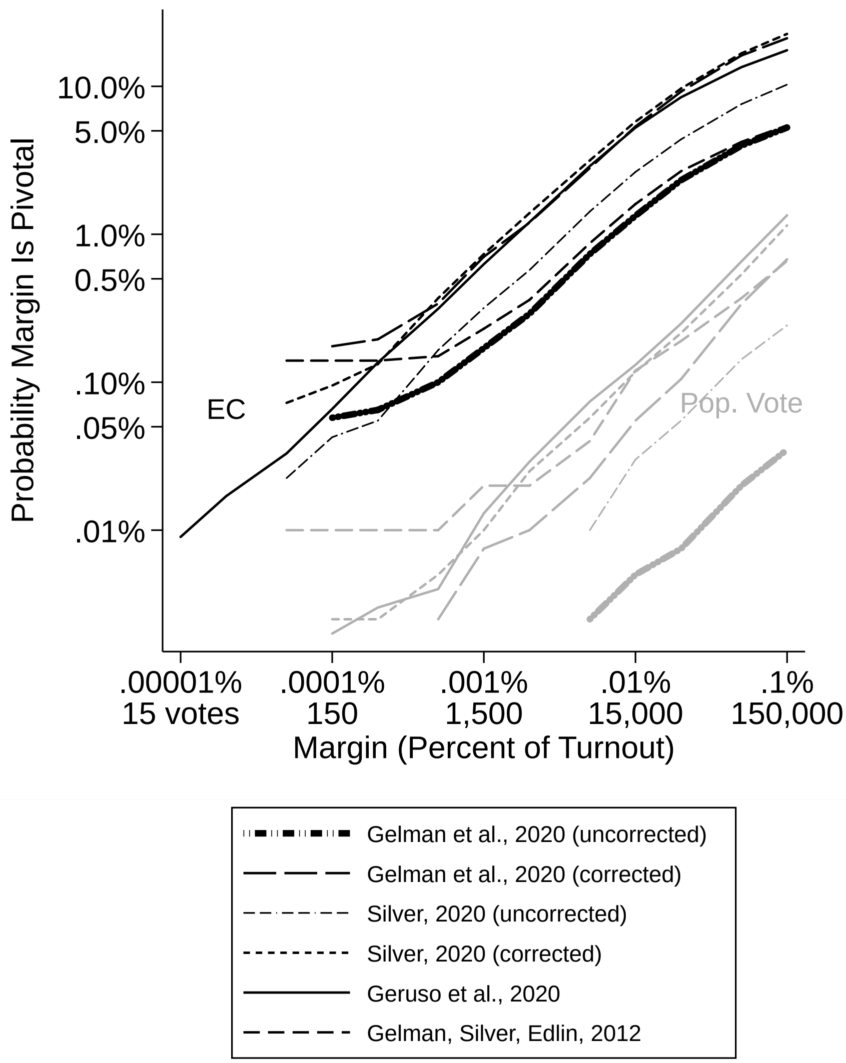
Note: Table lists calibration values for a National Popular Vote probability distribution that would match the expected close-election risk of the EC system. In particular, we calculate the standard deviation σ that generates the same probability as does the EC of arriving at some close outcome within margin m. Formally, we find the σ that solves $\int_{m/2\sigma}^{m/2\sigma} \phi(x)dx = Pr.^{EC}(m)$, where ϕ is the standard normal density and $Pr.^{EC}(m)$ is the probability of a within-margin-m result in the EC system. Results are calibrated to the GST model from Table 2, Panel A.

Table A6: Summary Statistics for the Counterfactuals in Figure 6

Assumed Polarization	National Republican Vote Share	
	Mean	SD
0 (baseline model)	0.490	0.0277
+5 Polarization	0.489	0.0277
+10 Polarization	0.488	0.0277
+15 Polarization	0.487	0.0277
+17 Polarization	0.486	0.0277
+20 Polarization	0.486	0.0277

Note: Table provides additional statistics for the Figure 6 polarization exercise. Under "polarize by 10 points" counterfactuals, we adjust the state expectation of the Republican vote share upward by 10 percentage points in states with an expected Republican share greater than 0.5 in the baseline model and downward by 10 percentage points elsewhere. Means and standard deviations of the resulting probability distributions over the national popular vote are presented in the columns for each counterfactual. See Figure 6 notes for additional detail.

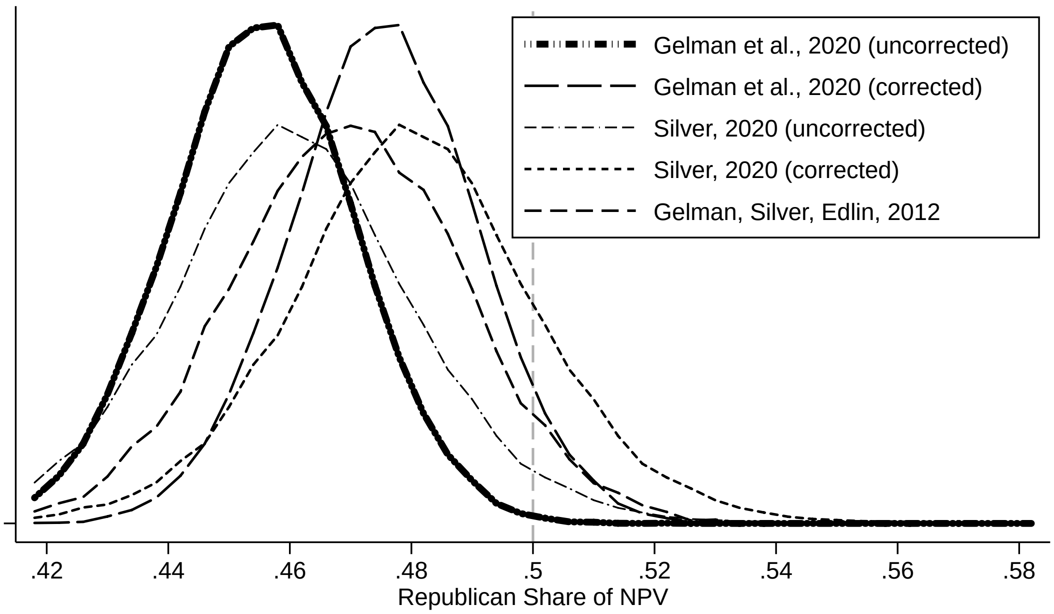
Figure A1: Main Result Replicated in 2008 Forecast Model and Uncorrected 2020 Forecast Models



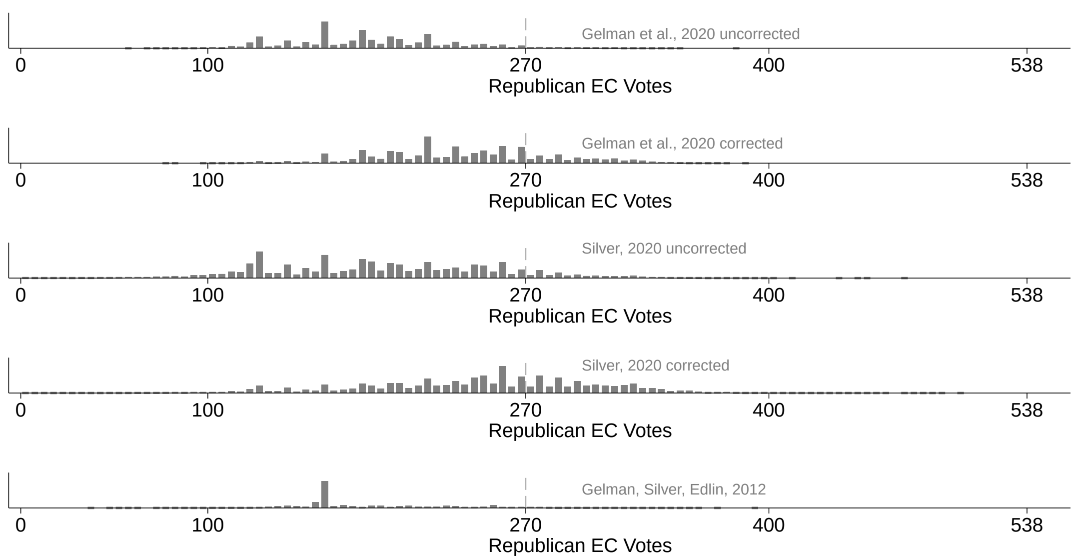
Note: Figure replicates the main result (Figure 2) for an extended set of models of the modern period. We plot both corrected and uncorrected versions of the 2020 forecasting models. The corrected versions subtract a 2 percentage point ex-post model bias favoring Democrats from every state-level simulated outcome in the *Economist* (Gelman, Heidemanns and Morris, 2020) and *FiveThirtyEight* (Silver, 2020) forecasts. The Gelman, Silver and Edlin (2012) forecast model characterizes the 2008 election. See Figure A2 for descriptive statistics of these models.

Figure A2: Data and Descriptive Statistics for 2008 Forecast Model and Uncorrected 2020 Forecast Models

(A) Implied Distributions of the National Popular Vote



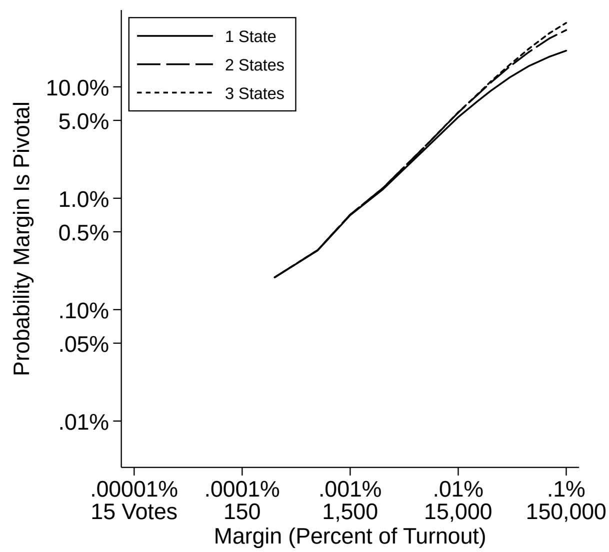
(B) Implied Distributions of the EC Vote



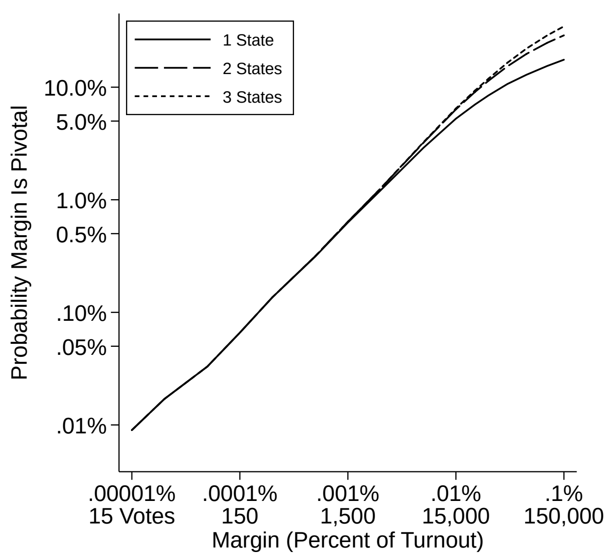
Note: Figure shows the probability distributions over electoral outcomes implied by three datasources, constructed from the 10,000 or 40,000 simulation draws in each datasource. Panel (A) plots the distribution of popular vote outcomes. Panel (B) plots the distribution of electoral vote outcomes. The corrected versions subtract a 2 percentage point ex-post model bias favoring Democrats from every state-level simulated outcome in the *Economist* (Gelman, Heidemanns and Morris, 2020) and *FiveThirtyEight* (Silver, 2020) forecasts. The Gelman, Silver and Edlin (2012) forecast model characterizes the 2008 election.

Figure A3: Allowing for Multiple States to Be Jointly Pivotal (Extended m Range for Figure 3)

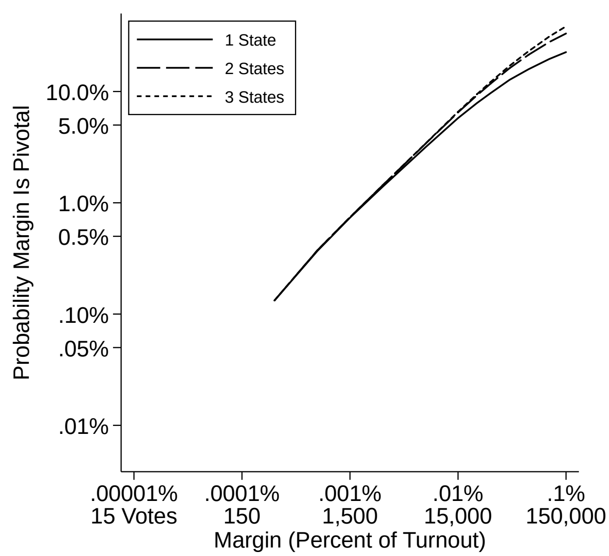
(A) Gelman et al. (2020)



(B) Geruso et al. (2020)

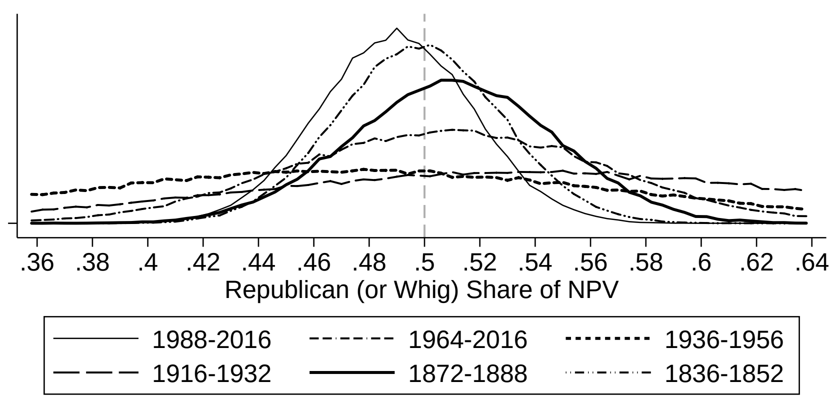


(C) Silver (2020)



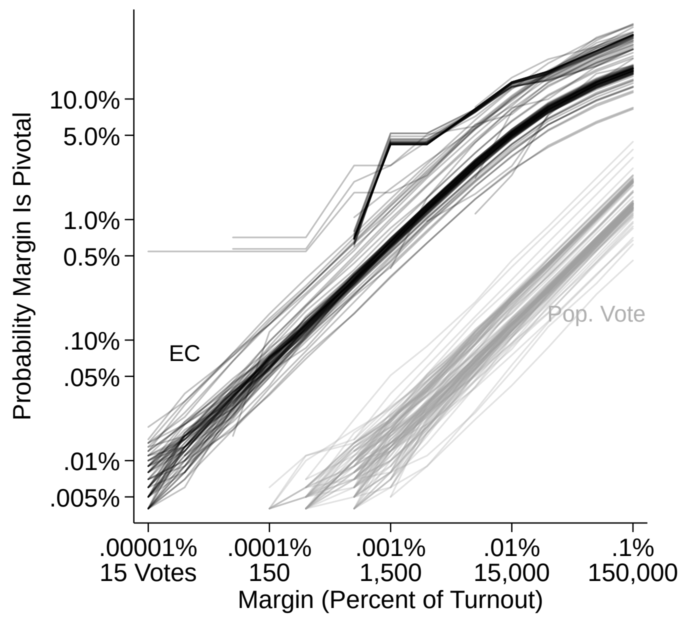
Note: Figure plots probabilities that election outcomes could be reversible by a small number of votes, extending the horizontal range in Figure 3. See Figure 3 for additional detail.

Figure A4: Popular Vote Distributions in the Historical Models in Geruso, Spears and Talesara (2020)



Note: Each trace represents a model from Geruso, Spears and Talesara (2020) describing a different time period, as indicated. See Figure 4 notes for additional detail.

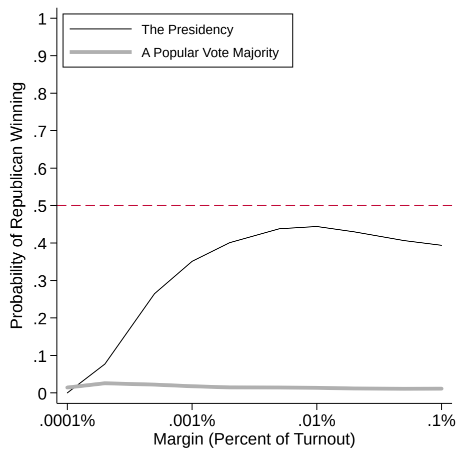
Figure A5: Robustness to 109 Alternative Election Models of the Modern Period (1964/1988–2016) from Geruso, Spears and Talesara (2020)



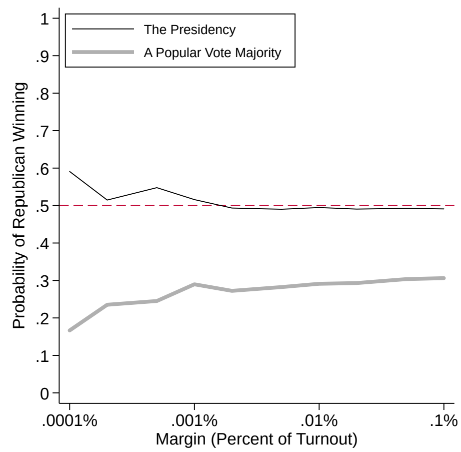
Note: Figure overlays the probability of a close outcome for 109 alternative models of the modern period. All models imported directly from Geruso, Spears and Talesara (2020). The models include: Parametric models (named M1, M2, M5 in GST), non-parametric bootstrap models (named M3, M4, M6 in GST), models that drop from the estimation sample elections in which an inversion occurred (2000, 2016), models that assign all third-party votes to Democrats (or Republicans) prior to estimation, models that vary the extent to which the shock process is correlated across states, and many other variants. See Geruso, Spears and Talesara (2020) for full detail.

Figure A6: Asymmetry in Close Races: Alternative Models

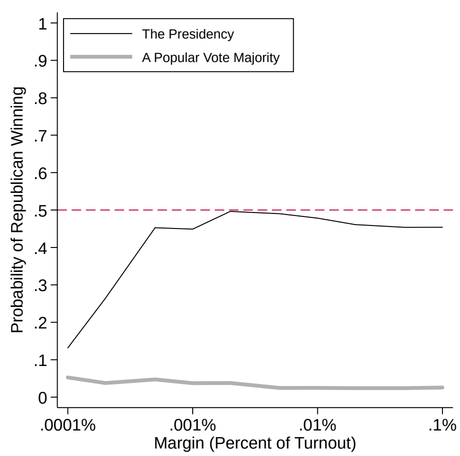
(A) Gelman et al. (2020)



(B) Geruso et al. (2020)

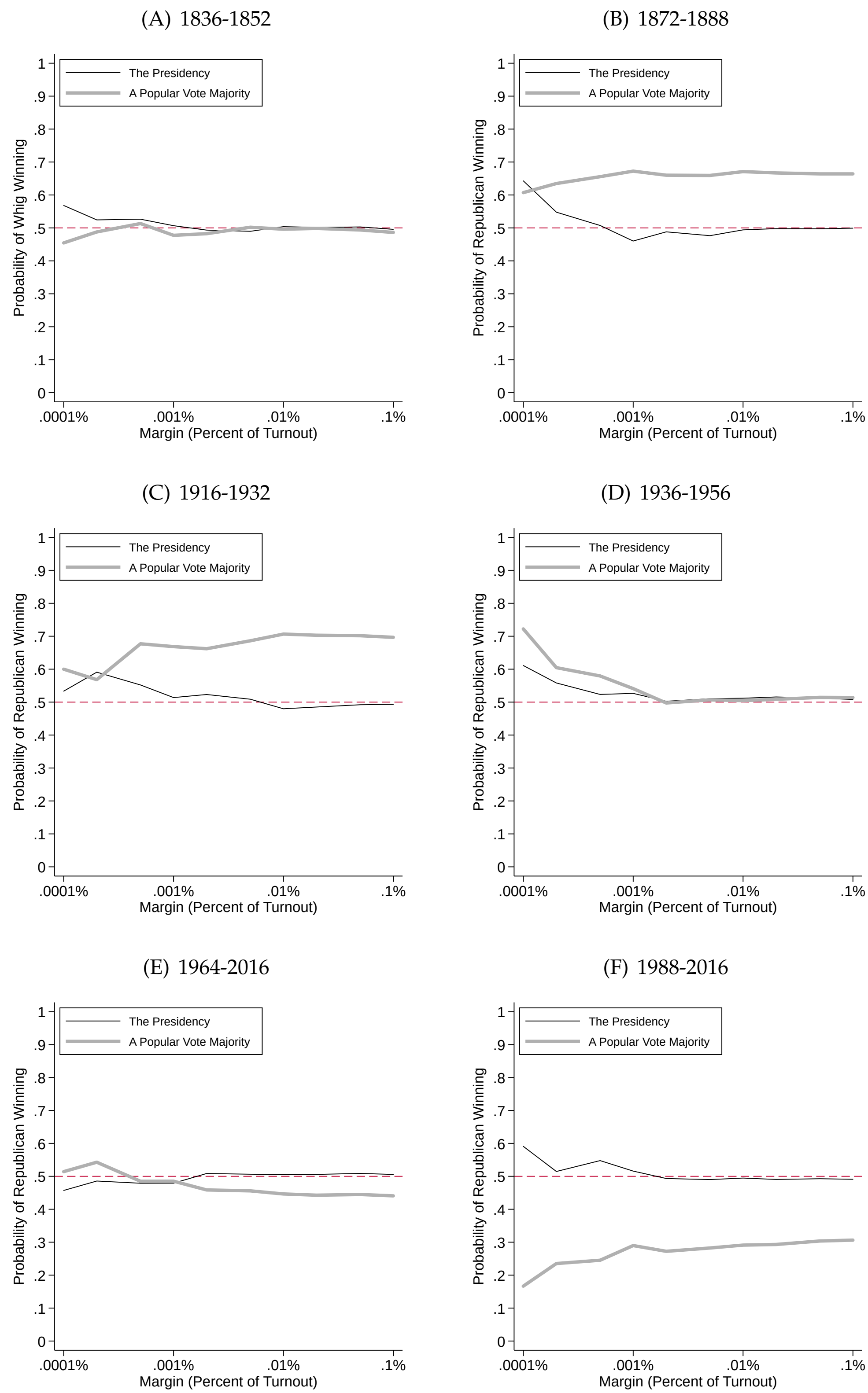


(C) Silver, 2020



Note: Figure plots the probability that the Republican candidate wins the presidency and the popular vote majority, conditional on a race decided in the EC by a margin indicated along the horizontal axis. Most simulated election outcomes are excluded from this figure because most outcomes are not close. The sources of data are indicated in the caption of each panel. See Figure 7 notes for full detail.

Figure A7: Asymmetry in Close Races: Historical Models



Note: Figure plots the probability that the Republican candidate wins the presidency and the popular vote majority, conditional on a race decided in the EC by a margin indicated along the horizontal axis. Most simulated election outcomes are excluded from this figure because most outcomes are not close. The models here are the GST models covering the indicated periods. See Figure 7 notes for full detail.