Democracy in America? Partisanship, Polarization, and the Robustness of Support for Democracy in the United States

MATTHEW H. GRAHAM *Yale University*

MILAN W. SVOLIK *Yale University*

*I*s support for democracy in the United States robust enough to deter undemocratic behavior by elected politicians? We develop a model of the public as a democratic check and evaluate it using two empirical strategies: an original, nationally representative candidate-choice experiment in which some politicians take positions that violate key democratic principles, and a natural experiment that occurred during Montana's 2017 special election for the U.S. House. Our research design allows us to infer Americans' willingness to trade-off democratic principles for other valid but potentially conflicting considerations such as political ideology, partisan loyalty, and policy preferences. We find the U.S. public's viability as a democratic check to be strikingly limited: only a small fraction of Americans prioritize democratic principles in their electoral choices, and their tendency to do so is decreasing in several measures of polarization, including the strength of partisanship, policy extremism, and candidate platform divergence. Our findings echo classic arguments about the importance of political moderation and cross-cutting cleavages for democratic stability and highlight the dangers that polarization represents for democracy.

"*It is the function of public opinion to check the use of force in a crisis, so that men, driven to make terms, may live and let live.*"

Walter Lippmann, *The Phantom Public* (1925, 64)

INTRODUCTION

"It is nearly impossible to find an American who says that he is opposed to democracy or favors some alternative…On the contrary, nearly everyone professes to believe that democracy is the best form of government." This is how Robert A. Dahl, writing in 1966, summarized contemporary evidence for the support for democracy in the United States (Dahl 1966, 40). It remains conventional wisdom to this day. Research that traces its intellectual origins to Tocqueville's *Democracy in America* finds that the United States consistently exhibits some of the highest levels of support for democracy in the world (Almond and Verba 1963; Inglehart and Welzel 2010; Norris 2011).

We show that this conventional wisdom rests on fragile foundations. Rather than asking about support

for democracy directly, we adopt an approach that infers Americans' commitment to democratic principles from their choices of candidates in hypothetical election scenarios. Each candidate is experimentally assigned attributes and platforms that approximate real-world elections and, crucially, may endorse positions that violate core democratic principles, including free and fair elections, civil liberties, and checks and balances. In this framework, voters "support democracy" not when they say so, but rather when their choices *reveal* a preference for democratic principles over other valid but potentially conflicting considerations such as political ideology, partisan loyalty, or policy preferences.

This research design builds on the observation that elections represent a fundamental instrument of democratic self-defense: Especially in advanced democracies, voters have the opportunity to stop politicians whose positions violate democratic principles by defeating them at the polls. We argue that a key obstacle to the viability of such a democratic check is partisan, ideological, and policy-based polarization. Electoral competition often confronts voters with a choice between two valid but potentially conflicting considerations: partisan interests and democratic principles. Polarization raises the stakes of elections and in turn the price of prioritizing democratic principles over partisan interests. When faced with the choice between a co-partisan candidate whose positions violate democratic principles and a candidate who complies with democratic principles but is otherwise unappealing, a significant fraction of voters may sacrifice democratic principles to elect a candidate who champions their party or interests. In a sharply polarized electorate, even pro-democratically minded voters may act as partisans first and democrats only second.

In the next section, we formalize these intuitions and develop a model of the public as a democratic check. Building on Svolik (2020), we extend the classic, spatial framework for electoral competition to account for candidates who may hold positions that undermine democratic principles. The latter are conceptualized as

Matthew H. Graham [iD], Ph.D. Candidate, Department of Political Science, Yale University, matthew.graham@yale.edu.

Milan W. Svolik [iD], Professor, Department of Political Science, Yale University, milan.svolik@yale.edu.

negative valence attributes: while voters may differ over policy, ideology, or partisanship, they agree that electoral competition should be democratic and prefer candidates who comply with key democratic principles. This framework yields a number of predictions about the consequences of polarization for an electorate's resilience to undemocratic candidates: (i) voters who hold extreme or intense policy preferences will be willing to sacrifice democratic principles at higher rates than centrist and moderate voters, (ii) electorates that are polarized or lack cross-cutting cleavages will be less punishing of candidates that undermine democratic principles, and (iii) candidate platform polarization will be detrimental to democracy independent of voter polarization. Our model thus provides micro-foundations for classic arguments about the importance of political moderation and cross-cutting cleavages for democratic stability (Dahl 1956; Lipset 1960).

This framework guides the design and analysis of our candidate-choice experiment as well as a natural experiment that occurred during Montana's 2017 special election for the U.S. House. The following is a summary of our experimental findings:

1. **Americans value democracy, but not much**: A candidate who considers adopting an undemocratic position can expect to be punished by losing only about 11.7% of his overall vote share. When we restrict attention to candidate-choice scenarios with combinations of partisanship and policies that we typically see in real-world elections, this punishment drops to 3.5%.
2. **Support for democracy is highly elastic**: When the price of voting for a more democratic candidate is that candidate's greater distance from the voter in terms of her preferred policies, even the most centrist voters are willing to tolerate at most a 10–15% increase in such a distance.
3. **Centrists are a pro-democratic force**: "Centrist" voters who see small policy differences between candidates punish undemocratic behavior at four times the rate of "extremist" voters who strongly favor one of the candidates.
4. **Most voters are partisans first and democrats only second**: Only about 13.1% of our respondents are willing to defect from a co-partisan candidate for violating democratic principles when the price of doing so is voting against their own party. Only independents and partisan "leaners" support more democratic candidates enough to defeat undemocratic ones regardless of their partisan affiliation.
5. **Supporters of both parties employ a partisan "double standard"**: Respondents who identify as Republican are more willing to punish undemocratic behavior by Democratic Party than Republican Party candidates and vice versa. These effects are about equal among both Democrat and Republican respondents.
6. **Platform polarization is bad for democracy**: The larger the difference between the candidates' policy platforms, the weaker the punishment for their undemocratic behavior.
7. **Sensitivity to the menu of manipulation varies**: Voters are most punishing of undemocratic positions that undermine the free press and the rule of law; they are least punishing of restrictions on the freedom of assembly and executive aggrandizement. Nonetheless, when we benchmark these against extramarital affairs and underpayment of taxes—two negative valence attributes unrelated to democracy—we find that voters punish the latter *more* severely than they punish violations of democratic principles.
8. **Americans have a solid understanding of what democracy is and what it is not**: The vast majority of our respondents correctly distinguish real-world undemocratic practices—including those endorsed by our experimental candidates—from those that are consistent with democratic principles.

After examining these findings nonparametrically, we take advantage of the close connection between the design of our candidate-choice experiment and our theoretical framework and adopt a structural approach. This approach allows us to parsimoniously summarize the variation in our respondents' candidate choices using a few theoretically grounded parameters, including the weight that voters place on democracy relative to party and policy. We estimate this weight at 18% and find that candidates' democracy positions, economic and social policy platforms, and partisan affiliation jointly account for 80% of the systematic variation in voters' candidate choices. We also "price" democracy in terms of the candidate characteristics voters are willing to forgo to punish candidates who violate democratic principles. We find that this price corresponds to about half the difference between our respondents' favorite and least favorite policy positions. We then extend our theoretical model to account for our experimental findings about a partisan double standard in the punishment of candidates who violate democratic principles. We estimate a 48.4% co-partisan bias, implying that Americans are neither fully principled nor purely instrumental in their support for democracy.

The last portion of our analysis moves from hypothetical election scenarios to a real-world election: a natural experiment that occurred during the 2017 special election for Montana's only seat in the U.S. House of Representatives. On the eve of the election, one of the two major candidates assaulted a journalist, which we interpret as a negative public signal about his respect for a free press, or at a minimum, an undesirable valence attribute. Crucially, only election day voters saw this signal before they voted; absentee voters, who in Montana make up a majority of registered voters, had already cast their ballots. This allows us to adopt a difference-in-differences empirical strategy that compares precinct-level vote shifts between absentee and election day voters to infer their willingness to punish the assault. Our findings are consistent with both our theoretical expectations and experimental results. We find that only moderate precincts punished the assault on the journalist by voting across party lines. In heavily partisan precincts, partisan loyalty trumped valence considerations.

Our findings about the robustness of support for democracy in the United States contribute to a number of debates in comparative and American politics. Most

immediately, our article joins a growing body of work that has, especially in the wake of the 2016 presidential election, begun to reassess our knowledge about democratic stability in the United States[1] and other advanced democracies.[2] Theoretically, our arguments parallel Levitsky and Ziblatt (2018), who also highlight the dangers that polarization represents for democratic stability. Whereas their analysis focuses primarily on how polarization weakens the often informal norms that regulate interactions among political elites, our emphasis is on how polarization undermines the public's capacity to punish politicians for subverting the democratic process.

Our empirical methodology and substantive focus are closest to Carey et al. (2018), who also employ a candidate-choice experiment to study the commitment to democratic principles among the American public. They too report that although voters do punish candidates whose positions violate democratic principles, the magnitude of that punishment may be overshadowed by other political considerations—most notably partisanship. As we discuss in the conclusion, these findings suggest that conventional, direct-questioning measures of support for democracy may be flawed because they fail to account for voters' willingness to trade off democratic principles for partisan interests. In turn, our existing knowledge about the support for democracy in the United States and other advanced democracies is of limited utility when it comes to answering a key question: When can we realistically expect the public to check the authoritarian temptations of elected politicians?

Our theoretical framework helps us to address this question by proposing a new perspective on democratic stability: democracy is "self-enforcing" when politicians anticipate that, were they to behave undemocratically, their own supporters would punish them by voting for a competitor in large enough numbers to bring about their defeat. We explain why this electoral check may fail in polarized societies and even among voters who value democracy for its own sake: in sharply divided societies, voters put partisan ends above democratic principles.[3] The microfoundations that we develop thus combine insights from two lines of classic democratization research. The first views intense political cleavages as a threat to democratic stability (Dahl 1956; Lipset 1960); the second asks that explanations of democratic stability be explicit about the incentives of key actors to comply with the rules of democratic politics (Przeworski 1991; Weingast 1997). More broadly, we contribute to comparative politics research on democratic backsliding (Gandhi and Ong 2019; Haggard and Kaufman 2016;

Luo and Przeworski 2018; Nalepa, Vanberg, and Chiopris 2018; Waldner and Lust 2018).

Jointly, our theoretical framework and empirical findings suggest an explanation for the puzzling persistence in the United States of a number of deficiencies in the democratic process, especially at the state and local levels. Most frequent among these are gerrymandering (Chen and Rodden 2013; Cho and Liu 2016) and voter suppression (Grimmer et al. 2018). Our model implies that elected officials' incentives to comply with key democratic principles critically depend on the public's willingness to sanction those who violate them or neglect their enforcement.[4] Yet, our empirical analysis reveals that this check is at best limited in magnitude and subject to a partisan double standard. As we discuss in the conclusion, our findings imply that public officials may be effectively insulated from electoral sanction in states and districts where one party enjoys a significant electoral advantage.

Analytically as well as empirically, we examine the pernicious consequences of a number of distinct conceptions of polarization. At the level of individual voters, polarization may characterize voters who hold either extreme or intense preferences. At the level of the electorate, polarization may correspond to either a *U*-shaped distribution of voter preferences or one with a high correlation of preferences across issues or between issue areas and partisan affiliation. At the level of the candidates, polarization can be conceived of as corresponding to a large distance between candidate platforms. These distinct conceptions of polarization are rarely examined within a single theoretical framework. Our analysis demonstrates that each kind of polarization independently undermines the electorate's resilience to undemocratic candidates.

This comprehensive examination of the relationship between polarization and democratic stability contributes to a large research agenda that studies elite and mass polarization in the United States. Whereas most research on mass polarization focuses on characterizing its nature and origins (Abramowitz and Saunders 2008; Fiorina, Abrams, and Pope 2008; Iyengar et al. 2018; McCarty, Poole, and Rosenthal 2008), our focus is on a political consequence that this literature has yet to examine.[5] To our knowledge, our study is the first to examine what is possibly the most concerning consequence of increasing polarization in the United States: its potential to undermine the public's ability to check the undemocratic temptations of elected politicians.

A MODEL OF THE PUBLIC AS A DEMOCRATIC CHECK

Consider a model in which voters' preferences over two kinds of candidate attributes determine their electoral

---

[1] See especially Carey et al. (2019), Huq and Ginsburg (2018), Kaufman and Haggard (2019), Lieberman et al. (2019), Miller, Szakonyi, and Morgenbesser (2017), and Przeworski (2019).

[2] See Becher and Brouard (2019), Foa and Mounk (2017), and Voeten (2017).

[3] The trade-off between partisan interests and democratic principles can be seen as a special case of a more general trade-off between partisan and valence considerations (Ashworth and Bueno de Mesquita, 2009). Such an interpretation of our findings echoes Eggers (2014), who finds that voters punished politicians implicated in the 2009 U.K. expenses scandal less severely when the electoral stakes in their district were higher.

[4] Clark (2009) sees a similar role for public opinion in his analysis of court curbing.

[5] Because our primary focus is on mass polarization, we do not discuss research on polarization in Congress. For a review of this literature, see Barber and McCarty (2015).

choices: (i) positional issues, including candidates' policy positions, ideology, and partisan affiliation and (ii) candidates' compliance with key democratic principles. Building on Svolik (2020), we conceive of the latter as a valence issue: while voters may differ over policy, ideology, or partisanship, they all prefer candidates whose positions comply with key tenets of democratic electoral competition.

Formally, voter i's payoff from candidate j is

$$u_i(X_j, M_j) = -\sum_{k=1}^{K} \alpha_k (x_{ik} - x_{jk})^2 - \delta M_j, \quad \textbf{(1)}$$

where for each of K issues, x_{ik} is voter i's favorite position on issue k, x_{jk} is candidate j's platform on issue k, and α_k is the weight that i attaches to that issue. Meanwhile, M_j is candidate j's democracy position where M is increasing in how undemocratic j's platform is (M stands for "manipulation"). The term δ is the weight that voters attach to fair democratic competition—in effect, the intensity of their support for democracy.

This model yields several predictions that we evaluate throughout the article. First, voters who hold intense or extreme policy preferences are willing to tolerate undemocratic behavior by their favored candidate. To see the intuition behind these predictions, suppose there is only a single policy issue k. Then, i votes for candidate 1 as long as

$$x_{ik} \geq \frac{x_{1k} + x_{2k}}{2} + \frac{\delta(M_1 - M_2)}{2\alpha_k(x_{1k} - x_{2k})} \quad \text{for} \quad x_{1k} > x_{2k}, \quad \textbf{(2)}$$

where we are assuming that candidate 1's policy platform is to the right of candidate 2's platform.

Call the voter whose ideal policy x_{ik} barely satisfies the inequality in (2) the *swing voter* x_s^j. Note that the first term on the right-hand side of this inequality is the midpoint between the two candidates' policy platforms, separating the electorate into those who are policy-wise closer to candidate 1 and those who are closer to candidate 2. The swing voter x_s^j is located either to the right or the left of this midpoint, depending on which of the two candidates adopts an undemocratic position,

$$x_s^1, \ x_s^2 = \frac{x_{1k} + x_{2k}}{2} \pm \frac{\delta}{2\alpha_k(x_{1k} - x_{2k})}. \quad \textbf{(3)}$$

When candidate 1 adopts an undemocratic position ($M_1 = 1, M_2 = 0$), the swing voter x_s^1 is located to the right of the midpoint $\frac{x_{1k}+x_{2k}}{2}$. Voters to the right of the midpoint but to the left of x_s^1 favor candidate 1 based on their policy preferences, yet are sufficiently put off by his undemocratic position that they vote for candidate 2 instead. By contrast, voters whose policy preferences are extreme (large x_{ik}) or intense (large α_k) enough to be to the right of x_s^1 are willing to tolerate candidate 1's undemocratic position as their concern for democracy is outweighed by their proximity to his policies. The converse holds when candidate 2 adopts an undemocratic platform.

The segment between the two swing voters is related to another set of empirical predictions. It delineates the set of voters who always vote for the more democratic of the two candidates. We may therefore refer to this portion of the electorate as "democracy first" voters, as opposed to "policy first" voters. The fraction of voters in each category depends on both the length of this segment and the distribution of voters' ideal points. Its length $\frac{\delta}{\alpha_k(x_{1k}-x_{2k})}$ is increasing in the support for democracy δ; it is decreasing in the policy weight α_k and the distance between the candidates' platforms ($x_{1k} - x_{2k}$).

Meanwhile, the more polarized (U-shaped) the distribution of voters' ideal policies, the smaller the fraction of democracy first voters. The consequences of voter polarization are amplified when voters' preferences over the K policy issues correlate: when a voter's extreme position on one issue correlates with her extreme position on another, her preference for the more proximate candidate will compound across the two issues rather than cancel out.

The primary purpose of this framework is to guide the design of our candidate-choice experiment and the analysis of our data. This has shaped our theoretical analysis in two ways. First, we have intentionally treated the candidates' policy platforms and democracy positions as exogenous, as these will be randomly assigned in the candidate-choice experiment that we introduce in the next section.[6]

Second, to explicitly characterize the model's implications for our analysis of the candidate-choice experiment, we introduce a stochastic structure into the deterministic formulation of the voters' payoff in (1) and outline its consequences. We follow the classic random utility model of discrete choice and add to voter i's payoff from candidate j an error term ε_{ij},

$$u_i(X_j, M_j) = -\sum_{k=1}^{K} \alpha_k (x_{ik} - x_{jk})^2 - \delta M_j + \varepsilon_{ij}. \quad \textbf{(4)}$$

Assuming that the error terms ε_{ij} are (independently) drawn from type I extreme value distribution implies that the effect of candidate positions on voter i's probability of voting for a candidate can be estimated using the logistic regression.[7] We take advantage of this correspondence between our theoretical framework and the random utility model of discrete choice when we estimate the model's key parameters, including civic virtue δ and policy weights α_k.

Here, we highlight the aggregate empirical patterns implied by this stochastic formulation. First, in contrast to the deterministic case, the two swing voters no longer sharply separate "democracy first" from "policy first" voters. Rather, the segment between the two swing voters now delineates voters who, regardless of their policy proximity to the candidates, vote for the more democratic candidate with a probability greater than one-half. The left panel in Figure 1 illustrates this by plotting a voter's probability of voting for candidate 1 as a function of ideal policies x_{ik} over a single issue k.[8] The solid line plots the case when both candidates adopt

---

[6] For a model with endogenous policy and democracy positions, see Svolik (2020).
[7] This is only one of several plausible stochastic structures; see Cameron and Trivedi (2005, 476–8, 486–7).
[8] Figure 1 is based on parameter values $\alpha_k = 8$, $\delta = 2$, and $x_{1k} = -x_{2k} = 1/5$.

FIGURE 1. Model Predictions: The Fraction Voting for Candidate 1 (Left) and the Decline in the less Democratic Candidate's Vote Share (Right)
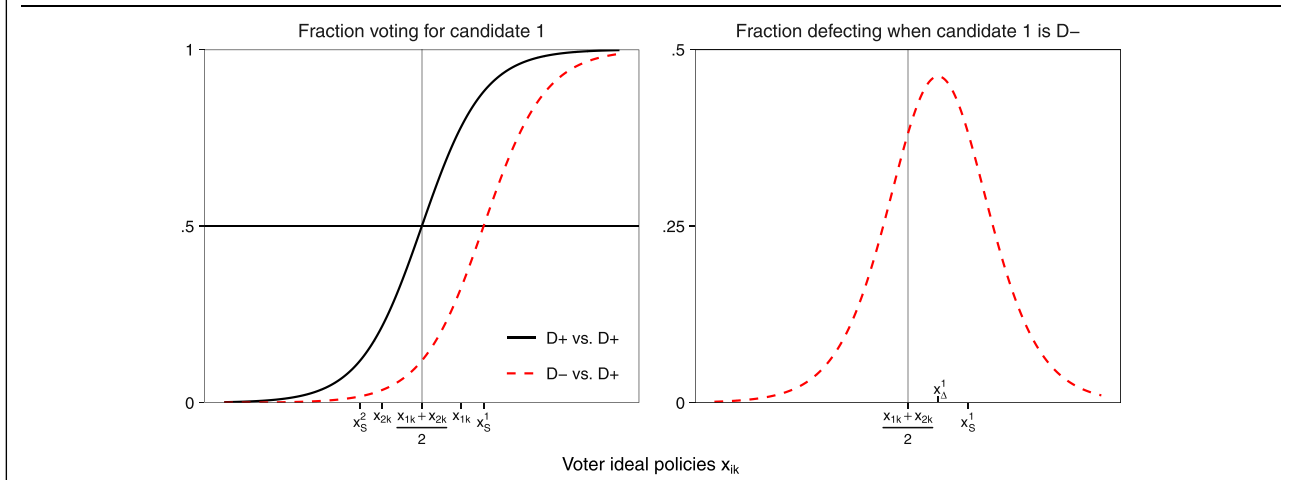
a democratic position; the dashed line plots the case when candidate 1 adopts an undemocratic position.

An alternative interpretation of the vertical axis in the left panel of Figure 1 that we employ when analyzing the candidate-choice experiment is that it corresponds to the fraction of voters in subgroups along the horizontal axis that support candidate 1. Furthermore, it will often be more practical to use as the horizontal axis the difference in voters' policy proximity to the two candidates rather than voters' ideal points alone. In that case, the intersection of the dashed curve with the 0.5 horizontal line will, instead of the swing voter x_s^1, mark the quantity $\frac{\delta}{\alpha_k}$, which we later interpret as voters' value for democracy in terms of issue k.

A key quantity that we estimate throughout the article is the overall fraction of voters who defect from a candidate that adopts an undemocratic position, which we denote by Δ. In effect, Δ measures the public's ability to serve as a democratic check. It results from a combination of two factors: the rate at which voters defect from the undemocratic candidate and the distribution of voters' ideal points.[9] The right panel in Figure 1 plots the former. It shows that while centrists will in general defect from undemocratic candidates at higher rates than extremists, the maximum defection rate (marked by x_Δ^1) will be located to the right of the midpoint $\frac{x_{1k}+x_{2k}}{2}$ but to the left of the swing voter x_s^1. Intuitively, it is voters who would otherwise narrowly break in favor of candidate 1 that abandon him at the highest rate once he adopts an undemocratic position.

To recapitulate, this theoretical framework yields five predictions about the relationship between a number of distinct conceptions of polarization and a decline in the public's ability to serve as a democratic check, which we denote by $\Delta\downarrow$:[10]

1. **Centrists are a pro-democratic force**: $\Delta\downarrow$ for voters whose policy preferences are more *extreme* relative to the candidates' platforms (x_{ik} far from the midpoint $\frac{x_{1k}+x_{2k}}{2}$);
2. **Moderation is good for democracy**: $\Delta\downarrow$ for voters who hold more *intense* policy preferences (a large α_k);
3. **Voter polarization is bad for democracy**: $\Delta\downarrow$ as the *distribution* of voters' ideal policies becomes more *U*-shaped (polarized) as opposed to inverse-*U*-shaped (centrist);
4. **Candidate polarization is bad for democracy**: $\Delta\downarrow$ as the *distance* between candidate platforms increases (large $|x_{1k} - x_{2k}|$);
5. **Cross-cutting cleavages are good for democracy**: $\Delta\downarrow$ when voter preferences over the K distinct issues are *aligned* as opposed to *cross-cutting* with respect to candidates platforms.

THE CANDIDATE-CHOICE EXPERIMENT

Our candidate-choice experiment investigates a key mechanism underlying the above predictions: even voters who value democratic principles may trade off those principles for partisan ends when confronted with a choice between the two. By asking respondents to choose between two candidates in an election, the experiment examines this mechanism at the same level at which it is hypothesized to operate—that of the individual voter. Moreover, because it entails one of the most essential and familiar actions that democratic citizens perform—voting—the experiment allows us to study our respondents' willingness to trade off democratic principles for partisan interests without alerting them to it.

In the candidate-choice experiment, respondents made a series of 16 choices, each between two candidates for a state legislature. The candidates were described by experimentally manipulated attributes typically seen in real-world elections: age, gender, race, profession, years of experience, partisan affiliation, two policy platforms, and a "democracy" position. This last attribute is the

---

[9] That is, $\Delta = \int_{-\infty}^{\infty} \left(\frac{1}{1+e^{-D}} - \frac{1}{1+e^{-(D-\delta M_1)}} \right) dF(x_i)$, where $F(x_i)$ is the distribution of voters and $D = -\sum_{k=1}^{K} \alpha_k (x_{ik} - x_{2k})^2 - \sum_{k=1}^{K} \alpha_k (x_{ik} - x_{1k})^2$.

[10] See the appendix for a formal statement and proofs of these predictions.

TABLE 1. Undemocratic Positions Endorsed by Candidates Assigned to the D^- Treatment Condition

D^-	Undemocratic position	Democratic principle
1a	Supported a redistricting plan that gives [own party]s 2 extra seats despite a decline in the polls	Electoral fairness
1b	Supported a redistricting plan that gives [own party]s 10 extra seats despite a decline in the polls	Electoral fairness
2	Supported a proposal to reduce the number of polling stations in areas that support [opposite party]s	Electoral fairness
3	Said the [own party] governor should rule by executive order if [opposite party] legislators don't cooperate	Checks and balances
4	Said the [own party] governor should ignore unfavorable court rulings by [opposite party]-appointed judges	Checks and balances
5	Said the [own party] governor should prosecute journalists who accuse him of misconduct without revealing sources	Civil liberties
6a	Said the [own party] governor should ban far-left group rallies in the state capital	Civil liberties
6b	Said the [own party] governor should ban far-right group rallies in the state capital	Civil liberties

focus of our analysis; we therefore describe its design and assignment below. We introduce most of the remaining attributes throughout the article. The appendix describes our design in detail and presents an example of a candidate-choice scenario as seen by our respondents.

Each candidate was assigned a democracy position that was either "undemocratic"—an action or statement by the candidate that violates a key democratic principle—or a democratically neutral, "generic" position. The undemocratic positions are listed in Table 1 and labeled D^-. There, [own party] refers to a candidate's randomly assigned political party (Democrat or Republican); [opposite party] denotes the complement. For instance, one possible realization of item 4 reads "Said the Republican governor should ignore unfavorable court rulings by Democrat-appointed judges."

In designing these undemocratic positions, we employed the following criteria:

Conceptual validity: The undemocratic positions capture violations of key democratic principles. Following classic scholarship on democratization (Dahl 1971), this includes measures that undermine electoral fairness (items 1a, 1b, and 2 in Table 1), checks and balances (items 3 and 4), and civil liberties (items 5, 6a, and 6b).

We also verify that our respondents indeed see the positions in Table 1 as undemocratic. One to three weeks before the candidate-choice experiment, each respondent was asked to evaluate a battery of both democratic and undemocratic practices, including the practices that would later appear as our D^- treatments. As we document in the appendix, the vast majority of our respondents correctly rated the D^- treatments as "undemocratic."

Contextual realism and partisan balance: The undemocratic positions approximate practices that have been used by politicians to subvert the democratic process in the United States and can be plausibly adopted by both major parties. Accordingly, the undemocratic positions are situated at the state level, where most attempts to subvert the democratic process for partisan gain in the United States occur and have historically been attempted by both major parties. The appendix provides real-world examples of each undemocratic position.

Incremental violations: A key feature of attempts to subvert the democratic process, both in the United States and around the world, is the use of ostensibly legal, incremental, and complementary measures (Levitsky and Ziblatt 2018; Waldner and Lust 2018). This has several consequences. First, such measures must often be implemented by or in conjunction with the executive. This is why some of our undemocratic positions refer to actions that the candidate suggests a co-partisan governor take. Second, because such measures are typically adopted through a constitutionally mandated process, they may undermine democratic principles without violating the law. This applies to items 1a, 1b, and 3. Finally, any single measure may allow for a partisan interpretation according to which it is consistent with some—often more majoritarian—conception of democracy or corrects an existing deficiency in the democratic process. For instance, proponents of stricter voter ID laws may respond to accusations of voter suppression by claiming that such measures are needed to prevent voter fraud, and proponents of gerrymandering may claim to be correcting an unfair status quo. Jointly and in their political context, however, such measures result in an uneven playing field that favors their proponents (Levitsky and Way 2010; Schedler 2002).

Neutral presentation: The undemocratic positions were presented in a manner that avoids conspicuousness or normatively leading language. To prevent candidates not assigned to hold an undemocratic position from appearing visually conspicuous, each was assigned one of seven democratically neutral, "generic" positions. For instance, one read: "Served on a committee that establishes the state legislature's schedule for each session"; we list the remaining six in the appendix. The content and length of such generic positions were designed to balance the cognitive effort required to distinguish a candidate who endorsed an undemocratic position from one that did not.[11] The wording of our undemocratic positions

---

[11] For the same reason, we randomized the order in which candidates' democracy and policy positions were listed.

also avoids negative connotations or normatively leading language. For instance, positions 1a, 1b, and 2 describe gerrymandering and voter suppression, respectively, but we intentionally avoided employing those terms. Put simply, we want respondents to decide for themselves whether or not a position violates a democratic principle.

In the experiment, each respondent made 16 distinct candidate choices, of which 11 were based on the following experimental design.[12] In four randomly chosen scenarios, both candidates adopted one of the democratically neutral, "generic" positions. Throughout, we treat these as our control scenarios and label them D^+ vs. D^+. In seven randomly chosen scenarios, one of the candidates adopted an undemocratic position, whereas the other held a neutral position. We refer to these as our treatment scenarios and label them D^- vs. D^+. Whether the undemocratic position was held by the candidate visually presented on the left or right was random. To simplify the presentation of our findings, we reshape our data so that candidate 2 always holds a neutral position (D^+) and, depending on the experimental condition, candidate 1 varies between D^+ and D^-.

The candidate-choice experiment was embedded in a nationally representative survey of American voters that took place in August–September 2018.[13] The 1,691 respondents made a total of 21,604 candidate choices.

Democratic Principles vs. Policy Preferences

We begin our analysis of the candidate-choice experiment by examining Americans' willingness to trade off democratic principles for their preferred policies. Each candidate proposed a platform in one economic and one social policy area. Economic policies concerned either state income taxes or state funding for local education; social policies concerned either immigration or marijuana's legal status. These policy areas were randomly assigned but identical across the two candidates. Within each policy area, candidates could adopt one of four platforms, ranging from extreme liberal to extreme conservative. The appendix lists all 16 policy platforms and justifies their selection.

One to three weeks before being presented with the candidate-choice scenarios, respondents rated each of the 16 policy platforms on a 0–100 proximity scale. This allows us to identify each respondent's ideal policy across the four areas and, following our theoretical framework, to compute the squared distance between a respondent's ideal policy and the two candidates' platforms for each of the candidate-choice scenarios

that the respondents faced.[14] The results that follow are robust to a range of alternative measures of policy proximity between respondents and candidates and account for the possibility that voters' policy preferences may be ideologically incoherent (Converse 1964), multidimensional (Treier and Hillygus 2009), or nonseparable from partisanship. We present these alternative measures in the appendix.

The left panel in Figure 2 plots the fraction of respondents voting for candidate 1 as a function of the difference in policy proximity to the respondent between candidate 1 and candidate 2.[15] On the horizontal axis, a value of 0 refers to scenarios when the two candidates are equally proximate to the respondent and a value of 1 (-1) to scenarios when candidate 1 is a full scale closer to (further away from) the respondent than candidate 2 on both policy areas. We treat the D^+ vs. D^+ scenario (solid line), when both candidates adopt neutral democracy positions but differ across other attributes, as our control condition; we treat the D^- vs. D^+ scenario (dashed line) as our treatment condition. Vertical bars denote 95% confidence intervals.[16] Figure 2 is thus the policy analog of Figure 1.

The D^+ vs. D^+ control scenario provides an initial plausibility check of our design. Consistent with our spatial framework, the closer candidate 1 is to a respondent's ideal policies relative to candidate 2, the more likely the respondent votes for candidate 1. Specifically, the fraction of respondents voting for candidate 1 increases from 13.0% when candidate 1 is a full scale less proximate to the respondent than candidate 2 to 87.0% in the opposite case. Furthermore, when the two candidates are equally proximate to respondents, the vote splits about evenly. Put simply, a respondent's proximity to each candidate's policy platform is a strong predictor of her candidate choices.

Figure 2 also provides an initial estimate of whether and how much Americans value democracy. Because all candidate attributes were randomly assigned, the only systematic difference between our control and treatment conditions is candidate 1's democracy position. We can therefore interpret a decline in the fraction of

---

[12] The remaining five scenarios were designed to provide extensions and robustness checks of our core design. We discuss them in the appendix.

[13] The first wave, which asked questions about partisanship, policy preferences, and support for democracy, took place on August 28–29, 2018. The primary focus of the second wave, which took place between September 4 and 25, 2018, was the candidate-choice experiment. Lucid fielded the survey. A pilot survey was fielded on Amazon Mechanical Turk in March 2018. The appendix benchmarks our sample against demographic data from the US Census Bureau and partisan and attitudinal questions from the 2016 American National Election Study.

[14] That is, the 0–100 proximity scale is a measure of $x_{ik} - x_{jk}$. For each candidate-choice scenario, this allows us to compute the squared distance $\sum_{k=1}^{2} \alpha_k (x_{ik} - x_{jk})^2$ between each respondent's ideal policies and both candidates' platforms, and in turn, candidate 1's policy proximity advantage $D = \sum_{k=1}^{2} \alpha_k (x_{ik} - x_{2k})^2 - \sum_{k=1}^{2} \alpha_k (x_{ik} - x_{1k})^2$.

[15] "Voting" here refers to the respondents' stated preference for one of the two candidates. We distinguish between respondents' candidate preferences and their stated intent to turn out to vote in the appendix and in Graham and Svolik (2019). We find that the stronger a respondent's preference for a D^- candidate, the more likely she is to abstain.

[16] Because each respondent made multiple choices, estimates that treat all observations as independent may understate statistical uncertainty. Unless otherwise noted, we compute all standard errors and confidence intervals using the block bootstrap, which accounts for dependence by resampling observations at the level of the respondent (see, e.g., Bertrand, Duflo, and Mullainathan 2004).

FIGURE 2. Fraction Voting for Candidate 1 (Left) and Fraction Defecting from the Less Democratic Candidate (Right) by the Difference between Candidates' in Policy Proximity to the Respondent
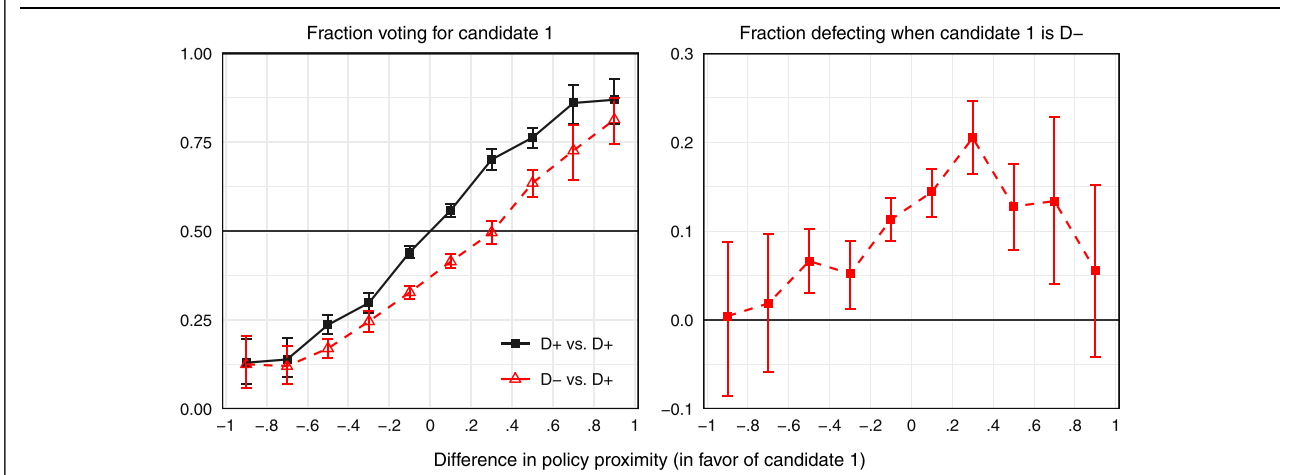
voters who support candidate 1 in the D^- vs. D^+ as opposed to the D^+ vs. D^+ condition as a measure of the public's ability to serve as a democratic check. Considering all candidate-choice scenarios in our experiment, the shift from the D^+ vs. D^+ to the D^- vs. D^+ condition decreases candidate 1's vote share from 50.0% to 38.3%—an 11.7% decline (CI: 10.0, 13.5). Put differently, just over 10% of Americans value democracy enough to punish an otherwise favored candidate for violating a democratic principle by voting against him.

Are Americans willing to trade off democratic principles in exchange for more appealing policies? Figure 2 allows us to address this question by partitioning our experimental electorate into two politically consequential subsets of voters anticipated by our theory: "democracy first" and "policy first" voters. A majority of the former vote for the more democratic candidate even when doing so goes against their policy interests. These respondents lie in the interval at the center of Figure 2 between the intersection of the D^- vs. D^+ line with the 0.5 horizontal axis and its mirror image along that axis. This interval corresponds to the values $(-0.3, 0.3)$ on the horizontal axis, and its limits are the empirical counterpart of the two swing voters in Figure 1. By contrast, voters to the left and right of this interval are "policy first" voters: a majority of these supports the more policy-wise proximate candidate, even if that candidate adopts an undemocratic position.

We gain additional insights into the robustness of support for democracy by examining differences in the severity with which respondents punish candidate 1 for adopting an undemocratic platform. The magnitude of this punishment is a combination of two factors: the baseline level of support for candidate 1 in each of the policy proximity subgroups and the rate at which respondents in a subgroup defect from candidate 1 after he adopts an undemocratic position.[17] The right panel in Figure 2 plots

the defection rate. Consistent with our theory, we see that the defection rate is highest among "bare supporters"—respondents who narrowly break in favor of candidate 1 in the D^+ vs. D^+ scenario—and declines as we move toward "policy extremists" on either side.

These policy-based differences in respondents' willingness to punish undemocratic behavior are consistent with our arguments about the pernicious consequences of polarization for democracy. Our representative sample allows us to simulate counterfactual electorates with increasing levels of policy polarization by varying the ratio of "policy centrists" to "policy extremists." As suggested by Figure 2, an electorate consisting entirely of "policy centrists"—the middle two bins—would result in a resounding defeat of a candidate who would adopt an undemocratic platform (12.9%, CI: 11.1, 15.1). In an electorate consisting entirely of the two most extreme subgroups, undemocratic candidates would only lose 3.0 percentage points of the vote share (CI: −4.1, 11.1), about one-fourth of the penalty among centrists.[18]

Does Partisanship Trump Civic Virtue?

Are voters willing to put democratic principles ahead of their partisan loyalties? To address this question, consider first only contests between candidates from different parties. Denote by Rep D^+ vs. Dem D^+ the

---

[17] That is, the defection rate is $\rho = \%_1(D^+ \text{ vs. } D^+) - \%_1(D^- \text{ vs. } D^+)$ where $\%_1(T)$ refers to the fraction of respondents voting for candidate 1 in treatment condition T.

[18] Because the candidates' two policies were independently and randomly assigned, respondents of all types—those who hold moderate as well as extreme policy preferences—faced some candidate-choice scenarios in which they were in the position of "centrists" and some in which they were "extremists." This results in an overall distribution of differences in policy proximity in Figure 2 that is heavy at the center: about 55% of scenarios are located on the interval $(-0.1, 0.1)$ and about 78% on the interval $(-0.3, 0.3)$. To probe the robustness of our claims, the appendix displays the distribution of policy proximity separated by respondents' strength of partisanship and replicates Figure 2 using a rank-based measure of respondents' proximity to candidates that, by construction, relies only on within-respondent differences in policy preferences.

control condition when both candidates adopt a neutral democracy position and by Rep D^+ vs. Dem D^- and Rep D^- vs. Dem D^+ the treatment conditions in which the Democrat or Republican, respectively, adopts an undemocratic position. Overall, undemocratic candidates are penalized by a loss of 9.9% (CI: 6.7, 13.0) and 10.9% (CI: 7.6, 14.0) of voters in the Rep D^+ vs. Dem D^- and Rep D^- vs. Dem D^+ scenarios, respectively.[19] Both effects are statistically different from zero but not statistically different from each other (difference: 1.0%, CI: −5.3, 3.3). Voters punish undemocratic behavior by both parties and they do so fairly evenly.

Are Americans willing to vote across party lines to punish a candidate for adopting an undemocratic position? 63.1% of our respondents support their own party in the control condition; this number declines to 54.8% when a respondent's co-partisan adopts an undemocratic position (−8.3%, CI: 5.0, 11.6). Put differently, only 13.1% of our respondents are willing to punish a co-partisan for violating democratic principles by voting against their own party. A majority of Americans are willing to give undemocratic co-partisans a pass.[20]

Figure 3 yields further insight into how our respondents' willingness to punish undemocratic candidates varies by the strength of their partisanship. It plots the fraction of respondents voting for a Republican Party candidate as a function of our respondents' party identification on the conventional 7-point scale, with the two treatment conditions plotted by dashed and dotted lines. As expected, stronger partisans vote for their party in greater proportion in the control condition, with independents breaking about evenly for the two parties. Crucially, only independents who lean toward a party defect from an undemocratic co-partisan in large enough numbers to defeat that candidate. By contrast, respondents who identify as a Democrat or Republican only break even. And a majority of strong partisans would rather elect a candidate who violates democratic principles than cross party lines.

Strong partisans' failure to punish undemocratic candidates from their own party is the result of two forces. First, strong partisans need to defect from undemocratic candidates at higher rates than moderates to compensate for their high baseline support for co-partisan candidates. Second, they do exactly the opposite: compared with other partisan subgroups, strong partisans are more lenient on violations of democratic principles by candidates from their party. Figure 4 shows this by plotting the fraction of respondents that defect from the D^- candidate, conditioning on both the respondent's and the D^- candidate's partisanship. We see that strong partisans are

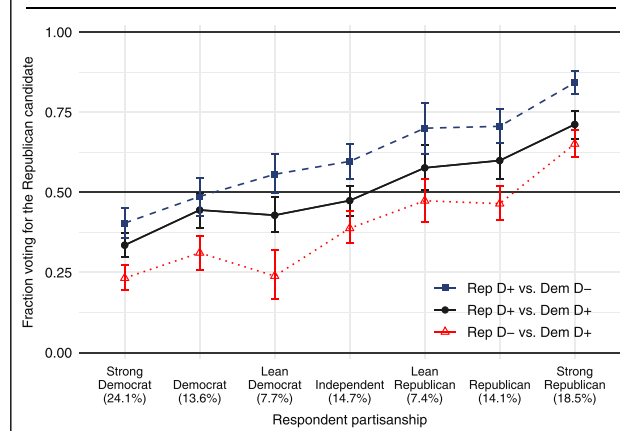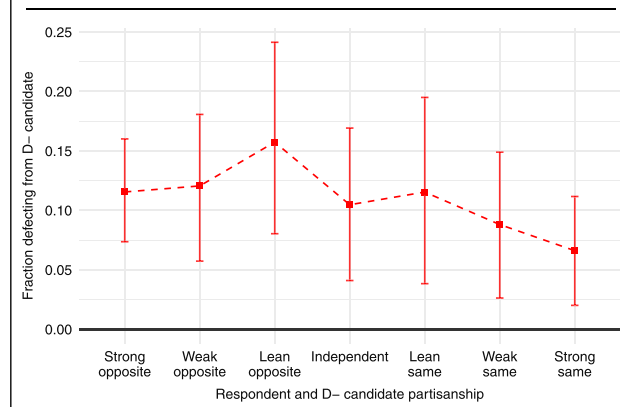**FIGURE 3. Different Party Contests: Fraction Voting for a Republican Candidate**

FIGURE 4. Different Party Contests: Fraction Defecting from the Less Democratic Candidate by the Respondent's Partisanship

less than half as likely to punish undemocratic candidates from their own party as are respondents who lean toward the opposite party. Put differently, strong partisans are willing to punish their own for violating democratic principles, but they hold candidates from the other party to a much higher standard.

This raises the question of whether contests between candidates from the same party—that is, primaries rather than general elections—are the more viable check on candidates who undermine democratic principles. Consistent with our theory, our respondents are more willing to punish undemocratic candidates in same party contests than in cross-party contests: a candidate who adopts an undemocratic position is penalized by a loss of 13.3% and 10.4% of voters, respectively (difference: 2.9%, CI: 0.0%, 6.1%). Figure 5, however, raises doubts about the promise of primaries as a democratic check. It plots the fraction of respondents voting for candidate 1 in contests between either two Democrats or two Republicans. When both candidates adopt a neutral democracy platform (either Dem D^+ vs. Dem D^+ or Rep D^+ vs. Rep D^+), that fraction is (by construction) 0.5. Now consider

---

[19] Both penalties imply that a D^- candidate is opposed by a majority respondents, regardless of that candidate's party.

[20] That is, among respondents who support their own party in the control condition, the treatment effect was 13.1% of the baseline, $(63.05 − 54.76)/63.05$. Among respondents who crossed party lines in the control condition, defection is higher: an undemocratic out-partisan's vote share declines from 37.0% to 24.4% (−12.6%, CI: 9.4, 15.9), which is 34.1% of the baseline, $(36.95 − 24.36)/36.95$.

the punishment for a Democrat who adopts an undemocratic position (Dem D^- vs. Dem D^+ in a dashed line): its severity is increasing as we move rightward along the horizontal axis, away from partisan supporters and toward partisan opponents. The reverse holds when a Republican candidate adopts an undemocratic position (Rep D^- vs. Rep D^+ in a dotted line). The promise of primaries as a democratic check is thus undermined by a partisan double standard: in most states, voters are not allowed to participate in the primary in which they would punish undemocratic candidates most severely.

In sum, both Democrats and Republicans employ a partisan double standard and they do so even when the partisanship of the winning candidate is preordained. A partisan double standard amplifies the consequences of polarization for the public's resilience to undemocratic candidates: an electorate consisting of strong partisans alone would provide only a tenuous check on candidates who violate democratic principles.

The Consequences of Candidate Polarization

A distinct advantage of our research design is that it allows us to explore the consequences of a number of distinct conceptions of polarization. In our empirical analysis so far, we have examined polarization as an individual or electorate-level phenomenon. We now shift to the candidate level, at which an increase in the distance between candidate platforms can be interpreted as an increase in a conceptually and empirically distinct kind of polarization—candidate polarization. Because candidates' policy platforms and partisanship were independently assigned in our experiment, we can examine how candidate polarization affects voters' ability to serve as a democratic check. Our theoretical model implies that greater candidate polarization results in a greater share of voters who are willing to tolerate undemocratic behavior. Crucially, these consequences of candidate polarization arise independently of voter polarization—they obtain even if we hold voter polarization fixed.

We now assess this prediction by examining the consequences of polarization in candidates' policy platforms. In Figure 6, the horizontal axis plots the mean absolute distance between the candidates' two policy platforms; the vertical axis plots the fraction of respondents defecting from a D^- candidate. We see that 15.9% of respondents defect from a D^- candidate when the two candidates take the same policy or have policy differences that cancel out (CI: 11.9%, 20.1%); the defection rate declines substantially when both policies are as far apart as possible.[21] Consistent with our theoretical framework, voters become more reluctant to punish candidates who violate democratic principles as candidates' policy platforms move apart.[22]

---

[21] See the appendix for a regression-based test of this relationship.
[22] Our earlier analysis of contests between candidates from the same versus different parties in effect found support for our prediction about the consequences of candidate polarization in terms of partisanship: we saw that respondents were less willing to punish undemocratic candidates in different-party than same-party contests.

FIGURE 5. Same Party Contests: Fraction Voting for Candidate 1 by the Respondent's and the D^- Candidate's Partisanship
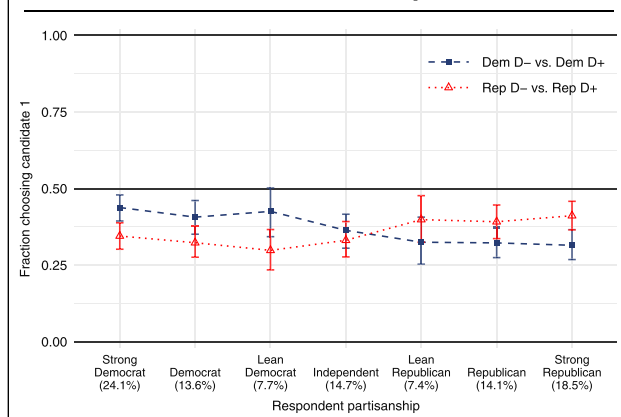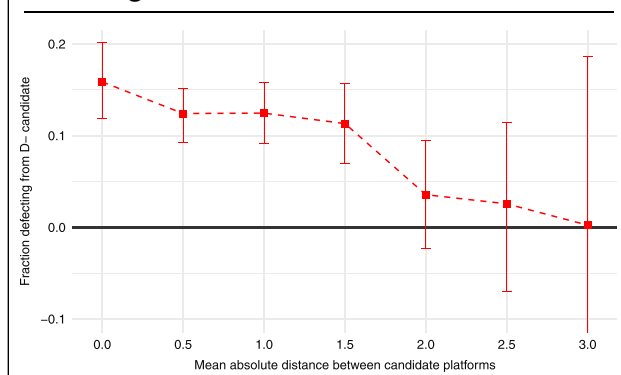

FIGURE 6. The Effect of the Mean Absolute Distance between the Candidates' Two Policy Platforms on the Fraction of Respondents Defecting from a D^- Candidate
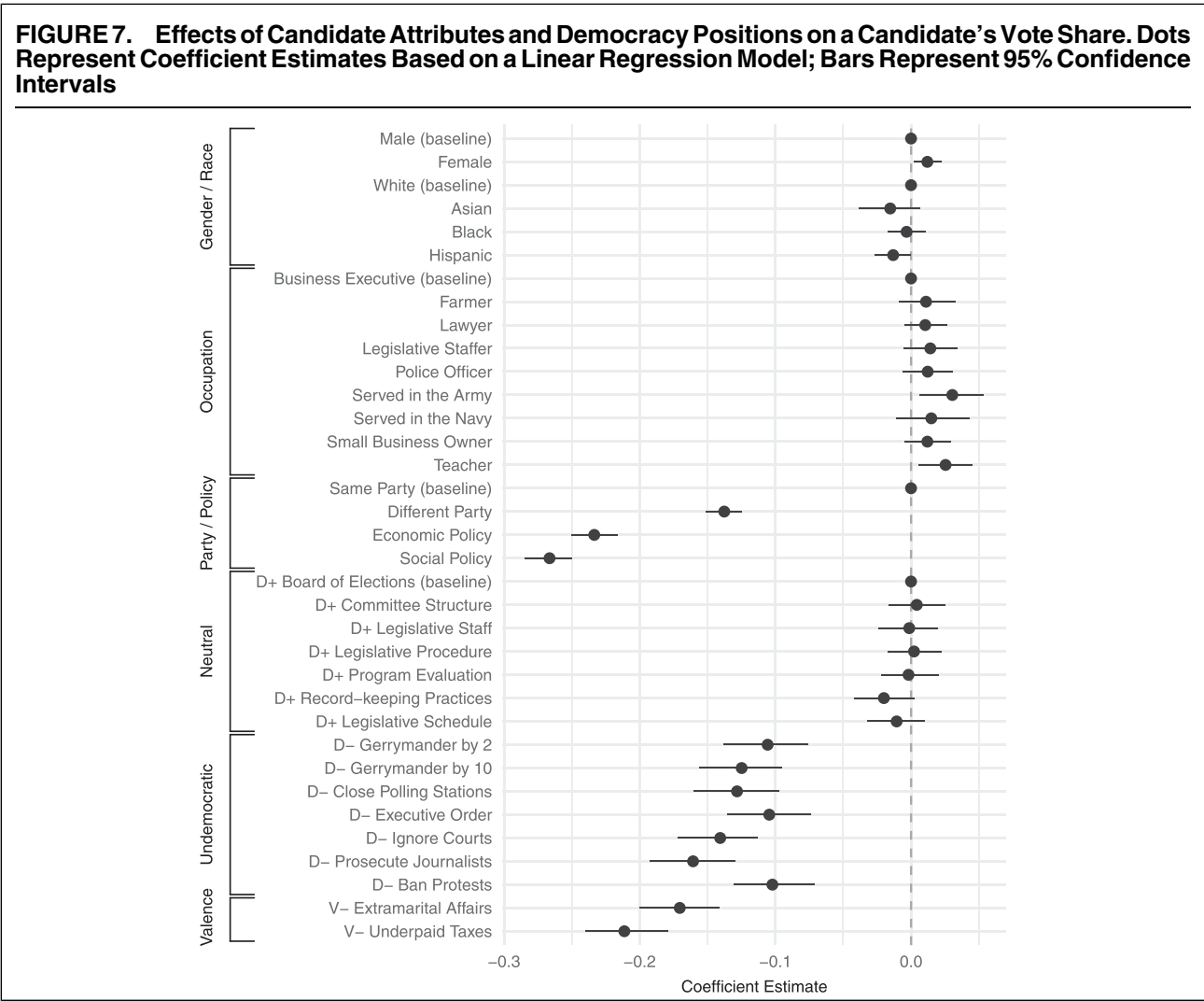
Resisting the Menu of Manipulation

When examining our respondents' willingness to punish candidates that violate democratic principles, we have so far pooled all democracy positions into two groups, neutral and undemocratic. We now examine Americans' willingness to tolerate the distinct undemocratic positions in Table 1 and interpret them in light of several benchmarks.

We estimate the following linear model:

$$\Pr(i \text{ votes for candidate } 1) = \alpha + \sum_k \beta_k (X_{i1k} - X_{i2k}) + \varepsilon_i.$$

(5)

In (5), X_{i1k} and X_{i2k} are dummy variables for all possible values of experimentally manipulated attribute k in choice i between candidates 1 and 2, respectively. Figure 7 plots the estimates of β_k. Bars represent the associated 95% confidence intervals; estimates without confidence intervals correspond to baseline categories. In the spirit of Hainmueller, Hopkins, and Yamamoto (2015), we interpret these coefficients as attribute k's average causal

FIGURE 7. **Effects of Candidate Attributes and Democracy Positions on a Candidate's Vote Share. Dots Represent Coefficient Estimates Based on a Linear Regression Model; Bars Represent 95% Confidence Intervals**

effect on a candidate's vote share, relative to the baseline attribute and averaging over the distribution of all other attributes.

Consider first the coefficients associated with the democratically neutral, "generic" positions. All seven are individually (and jointly) statistically indistinguishable from 0, implying that they do not affect a candidate's vote share. This validates our design and interpretation of these attributes as not only democratically neutral but also more generically inconsequential.

Figure 7 also demonstrates considerable variation in the effect of the individual undemocratic positions on a candidate's vote share. While all undemocratic positions impact a candidate's vote share negatively, the magnitude of that effect ranges from 10.2% to 16.1%.[23] Respondents most severely punish candidates who want to prosecute journalists (16.1%) and ignore court rulings (14.1%). Respondents are least sensitive to

candidates who endorse gerrymandering (by 2 seat, 10.6%) and suggest that the governor ban protests or rule by executive order (10.2 and 10.5%, respectively). Figure 8 differentiates these estimates by respondents' partisanship. Consistent with our earlier findings, we see few differences between supporters of the two major parties.[24]

To put the magnitude of these estimates in context, compare the effect of these undemocratic positions with that of other positional and valence candidate attributes. Consistent with our earlier analysis, the two main positional attributes—a candidate's party and policy platforms—have an impact on a respondent's candidate choice that is either comparable or greater in magnitude than individual undemocratic positions. The effects of the remaining attributes are an order of magnitude smaller than that of any undemocratic position.

---

[23] An *F*-test rejects the joint hypothesis that all coefficients associated with D^- positions in Figures 7 and 8 are equal at a *p*-value < 0.001. That is, there are statistically significant differences in how severely voters punish the distinct D^- positions.

[24] Figure 8 plots marginal vote shares (note the different scale of the horizontal axis than in Figure 7), ensuring that partisan differences in preferences toward the baseline attribute do not distort the comparisons (Leeper, Hobolt, and Tilley forthcoming).

FIGURE 8. Effects of Neutral, Valence, and Undemocratic Positions on a Candidate's Vote Share by Respondents' Partisanship. Dots Represent the Mean Vote Share among Candidates Adopting Each Undemocratic Position; Bars Represent 95% Confidence Intervals
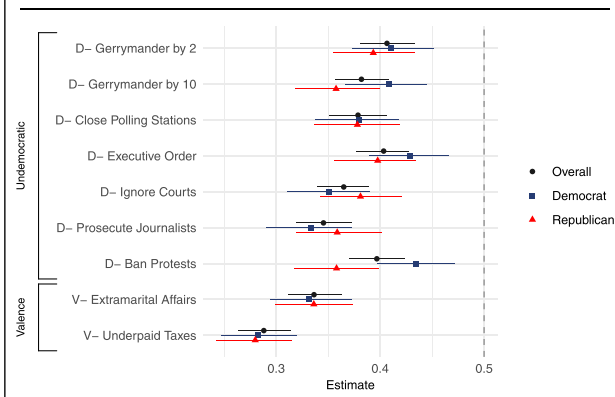
To help us further interpret the magnitude of undemocratic positions, two of each respondent's 16 choices included two negative valence attributes intentionally *unrelated* to democracy. According to the first, the candidate "was convicted of underpaying federal income taxes"; according to the second, the candidate "was reported to have had multiple extramarital affairs." Estimates associated with these two attributes appear at the bottom of Figures 7 and 8 and are labeled V^-. We see that voters punish candidates for extramarital affairs and underpaying taxes *more severely* than they punish them for undermining democratic principles.

STRUCTURAL ESTIMATES OF SUPPORT FOR DEMOCRACY

Throughout our analysis, we have systematically found a significant, if modest decline in the vote share of candidates whose positions violate democratic principles. This finding is consistent with our theoretical framework, which conceives of such positions as a negative valence characteristic and suggests that the civic virtue parameter δ in our model is positive.[25] We now shift from the nonparametric approach employed above to a structural approach that directly estimates our model's fundamentals: civic virtue δ and the weights α_k that voters place on other candidate attributes. This approach allows us to interpret δ as the *relative weight* that voters place on democracy in their voting decisions, to express voters' *value for democracy* in terms of other desirable candidate attributes that voters are willing to forgo to punish candidates who violate democratic principles, and to derive other theoretically informed quantities that yield insights about the robustness of support for democracy in the United States. In effect, the variation in our

respondents' candidate choices that we have seen so far can be parsimoniously summarized by a few fundamental, theoretically grounded parameters.

The estimation of these parameters is aided by a close correspondence between our theoretical model, the design of our candidate-choice experiment, and the random utility model of discrete choice. Recall that the assumption of type I extreme value distributed error terms ε_{ij} in (4) implies that the effect of candidate attributes on voter i's probability of voting for candidate 1 can be estimated using logistic regression.[26] For instance, our earlier discussion of voters' willingness to trade off democratic principles for their preferred policies and party suggests the estimation equation

$$\Pr(i \text{ votes for candidate } 1)$$
$$= \text{logit}^{-1}[\beta_0 + \beta_1 M + \beta_2 econ + \beta_3 social + \beta_4 party], \quad (6)$$

where $M, econ, social$, and $party$ are differences between candidate 1 and 2's democracy positions, economic and social policy platforms, and partisan affiliation. Estimates from (6) allow us to express the value that voters place on democracy vs. other candidate attributes k as normalized weights δ and α_k,

$$\hat{\delta} = \frac{|\beta_1|}{|\beta_1| + \sum_{l=2}^{L} |\beta_l|} \quad \text{and} \quad \hat{\alpha}_k = \frac{|\beta_{k+1}|}{|\beta_1| + \sum_{l=2}^{L} |\beta_l|},$$

where $l = 0, 1, \ldots, L$ indexes the regression coefficients in (6).

Table 2 presents estimates of the weights δ and α_k based on six alternative specifications. Column (1) estimates the logit model in (6) after the four input variables have been divided by twice their standard deviation. This standardization puts candidates' democracy positions, economic and social policy platforms, and partisan affiliation on a common, dispersion-based scale, allowing us to compare their importance (Gelman 2008). We see that the weight that voters place on democracy in their voting decisions is about 18.1% compared with 32.8% for social policy, 27.1% for economic policy, and 22.0% for partisanship.

Column (2) in Table 2 extends the model in (6) to include all experimentally manipulated candidate attributes.[27] The presented weights for candidates' democracy positions, economic and social policy platforms, and party imply that these four factors jointly account for 79.6% of the systematic variation in voters' candidate choices. These results are consistent with our earlier, nonparametric analysis: Americans value democracy—$\hat{\delta}$ is significantly greater than zero—but they value it less than major competing political considerations, especially economic and social policies and partisanship.

Estimates of the weights δ and α_k allow us to summarize an electorate's support for democracy in terms of voters' value for democracy. This concept expresses the

---

[25] This interpretation is bolstered by the fact that the vote share of D^- candidates declines across *all* respondent subgroups in Figures 2 and 3.

[26] See the appendix for evidence of goodness of fit.

[27] We present a complete set of results in the appendix.

TABLE 2. Structural Estimates of Our Model of the Public as a Democratic Check

	Standardized		Natural		Rating-based	
	(1)	(2)	(3)	(4)	(5)	(6)
δ democracy	0.181***	0.144***	0.155***	0.126***	0.211***	0.169***
	(0.161, 0.201)	(0.118, 0.164)	(0.137, 0.173)	(0.105, 0.143)	(0.191, 0.229)	(0.144, 0.188)
α_1 economic policy	0.271***	0.216***	0.312***	0.254***	0.292***	0.235***
	(0.254, 0.290)	(0.183, 0.236)	(0.291, 0.333)	(0.218, 0.276)	(0.273, 0.313)	(0.202, 0.257)
α_2 social policy	0.328***	0.261***	0.342***	0.279***	0.320***	0.256***
	(0.310, 0.346)	(0.222, 0.285)	(0.323, 0.363)	(0.239, 0.303)	(0.301, 0.339)	(0.222, 0.282)
α_3 party	0.220***	0.176***	0.190***	0.155***	0.177***	0.142***
	(0.202, 0.240)	(0.149, 0.197)	(0.173, 0.208)	(0.133, 0.173)	(0.161, 0.194)	(0.123, 0.161)
N	18,195	18,195	18,195	18,195	17,371	17,371

Note: Bootstrapped 95% confidence intervals in parentheses.
* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

trade-offs that voters are willing to make in terms of the implied marginal rate of substitution—the amount of other desirable candidate attributes that voters are ready to forgo to punish candidates who violate democratic principles. Specifically, the model in (6) implies that a voter's payoff remains unchanged as long as a shift in a candidate's democracy position M is accompanied by a shift of magnitude $\frac{\delta}{\alpha_k}$ in some other attribute k. Put simply, voters' value for democracy summarizes the robustness of support for democracy parsimoniously, in terms of the "price" voters are willing to pay to elect a more democratic candidate.

Estimates of value for democracy are most easily interpreted when we reestimate the model in (6) with candidate attributes entering in their most "natural" units. We do so by letting M and *party* enter as candidate-specific dummies and expressing economic and social policies on a 0–1 proximity scale. Columns (3) and (4) display the resulting estimates of the weights δ and α_k, paralleling the simple and extended models in (1) and (2). Estimates in (3) yield voters' value for democracy of 0.495 (CI: 0.424, 0.570) in economic policy, 0.454 (CI: 0.486, 0.523) in social policy, and 0.815 (CI: 0.687, 0.958) in co-partisanship. These quantities imply that to punish a candidate who adopts an undemocratic position, the voter is willing to tolerate a 0.5 and 0.45 increase in the distance from her ideal economic or social policies, respectively—about half the scale. The value for democracy in terms of co-partisanship is smaller than one, implying that the typical voter is not willing to vote across party lines to punish a co-partisan for adopting an undemocratic position.

Columns (5) and (6) probe the robustness of the estimates in columns (1)–(4) by letting M enter as each respondent's own pretreatment rating of "how undemocratic" the specific D^- treatment appearing in a candidate-choice scenario was. This approach allows for the possibility that only some respondents disapprove of the D^- positions or that they disapprove of the D^- positions at different rates. Our results remain essentially unchanged when we account for the respondents' own perception of the severity of each D^- position.

In our earlier, nonparametric analysis of partisanship, we found that our respondents punish platforms that

violate democratic principles less severely when adopted by co-partisan candidates. The present approach allows us to derive and estimate this tendency in terms of a single, theoretically informed quantity: the *partisan bias* π. Specifically, we extend our original model by multiplying candidate j's democracy platform M_j in (1) by the term $\delta[1 - \pi I_j(party)]$ instead of δ alone, where $I_j(party)$ is an indicator of co-partisanship between the respondent and candidate j. In turn, we may say that support for democracy is *principled* when $\pi = 0$ and voters punish candidates for violating democratic principles equally, regardless of their party; support for democracy is *instrumental* when $\pi = 1$, and voters only punish undemocratic positions when adopted by candidates from the other party.

This formulation implies that the degree of partisan bias π in the electorate can be estimated by adding to the logit model in (6) an interaction term between *party* and M. Denoting the coefficient on that interaction term by β_{int}, we have $\pi = 1 - \frac{\beta_{int}}{\beta_1}$. The estimate of π is 0.484 (CI: 0.358, 0.605), implying that Americans are neither fully principled nor purely instrumental in their support for democracy. Rather, voters are about 50% more lenient toward violations of democratic principles by candidates from their own party.

THE 2017 MONTANA NATURAL EXPERIMENT

Depending on whether they voted on the election day or by absentee ballot, voters in the 2017 special election for Montana's single U.S. House seat saw two different races.[28] Both races pitted Republican Greg Gianforte against Democrat Rob Quist. Absentee voters, who in Montana make up about 70% of registered voters, saw a small-government Republican with business credentials compete against a former musician Democrat who supported mainstream liberal positions. All three major

---

[28] The special election was held to fill the U.S. House seat vacated by Ryan Zinke, who became the Secretary of the Interior in the Trump administration.

newspapers in Montana initially endorsed Republican Greg Gianforte.

Election day voters saw the same race with one crucial difference: on the eve of the election, Gianforte assaulted *The Guardian* reporter Ben Jacobs after being repeatedly questioned about his position on health care reform.[29] The attack dominated the news coverage that evening and led the three major newspapers in Montana to rescind their endorsements of Gianforte on the morning of the election day. Gianforte nonetheless defeated Quist 50.2% to 44.1%, with 5.7% going to Libertarian Mark Wicks.

The timing of Gianforte's assault offers a real-world, quasi-experimental opportunity to evaluate our theoretical framework and corroborate our experimental results. We adopt a difference-in-differences empirical strategy that compares shifts in absentee and election day voters' support for the Republican U.S. House candidate between the November 2016 general election and the May 2017 special election.[30] Among absentee voters, shifts from 2016 to 2017 act as a control that reflects what would have happened if no voters observed the assault.[31] Vote shifts among election day voters, by contrast, reflect the causal effect of the assault. Even though absentee voters may be different from election day voters, such differences cancel out in a difference-in-differences framework where each group serves as its own, pretreatment baseline.

We interpret the assault as a public signal about Gianforte's respect for the free press or—at minimum—as a negative valence signal about his fitness for office. Our theory predicts that voters' willingness to punish Gianforte for the attack will be decreasing in the strength of the electorate's partisanship. In the context of Montana's partisan makeup, this implies that the precinct-level decline in Gianforte's vote share should be largest in moderate precincts and decreasing as Republican strength across precincts grows. This obtains because voters' willingness to tolerate a co-partisan who violates democratic principles increases in the intensity of their partisanship.[32]

---

[29] See, for example, Martin, Jonathan. "Montana Republican Greg Gianforte, Charged With Assault, Awaits Fate in Vote." *The New York Times*, May 24, 2017.

[30] On difference-in-differences estimation, see Bertrand, Duflo, and Mullainathan (2004).

[31] According to the Montana voter file, election officials had already received 92.5% of all absentee ballots by the time of Gianforte's assault. The corresponding figure for our sample of precincts is 94.1%.

[32] That is, we think of more Republican precincts as containing stronger partisans—a large x_{ik} in the voter's payoff function in (1)—and hence a greater fraction of "party first" as opposed to "democracy first" voters. Although this is only one of several possible voter-level interpretations of our precinct-level data, it is supported by actual associations between individual-level political attitudes and aggregate voting outcomes: in the appendix, we merge the 2016 county-level presidential vote data with the 2016 Cooperative Congressional Election Study to show that in more-Republican (Democratic) counties, Republicans (Democrats) tend to be stronger partisans, more conservative (liberal) on the liberal-conservative scale, more conservative (liberal) on issue position-based measures of ideology, and more disapproving (approving) of former President Barack Obama.

FIGURE 9. Differences in Precinct-Level Vote Shifts for the Republican U.S. House Candidate in Montana between November 2016 and May 2017
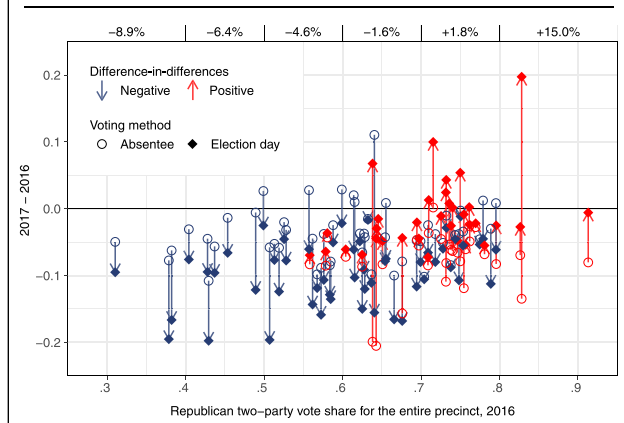
To investigate these predictions, we contacted election administrators in all of Montana's 56 counties and identified five counties that tallied absentee and election day ballots separately in 2016 and 2017. Figure 9 presents data based on the 87 precincts in these five counties. Separately for each voting method, it plots the precinct-level differences in Republican vote shifts between November 2016 and May 2017 as a function of the 2016 Republican two-party vote share in the entire precinct. Given the absence of extreme Democratic precincts in our sample, we treat the latter as an indicator for the intensity of a precinct's partisanship.[33] Absentee vote shifts are shown as circles and election day vote shifts as diamonds; positive differences in vote shifts are highlighted by upward-facing arrows and negative differences by downward-facing arrows. Consistent with our expectations, differences in vote shifts are negative and largest in moderate precincts; they decrease in magnitude and some even become positive as we move right along the horizontal axis, suggesting that punishment for the assault was smaller in more Republican precincts.[34]

To investigate this pattern formally, we estimate the following linear models:

$$R_{it} = \alpha + \beta_1 Y17_{it} + \beta_2 E_{it} + \beta_3 Y17_{it}E_{it} + \gamma \mathbf{X}_{it} + \varepsilon_{it}, \quad (7)$$

$$R_{it} = \alpha + \beta_1 Y17_{it} + \beta_2 E_{it} + \beta_3 Y17_{it}E_{it} + \beta_4 R_{16} + \beta_5 Y17_{it}R_{i16} + \beta_6 E_{it}R_{i16} + \beta_7 Y17_{it}E_{it}R_{i16} + \gamma \mathbf{X}_{it} + \varepsilon_{it}. \quad (8)$$

Above, R_{it} is the Republican candidate's vote share in precinct i in year t, $Y17_{it}$ is a dummy for the year 2017 (as opposed to 2016), E_{it} is a dummy for voting on election day (as opposed to absentee), R_{i16} is the Republican vote share for the entire precinct i in 2016, and $\mathbf{X}_{it}$ is a vector of control variables. The latter includes the

---

[33] The 2016 Republican two-party vote share ranges from 31% to 91%, with just nine precincts below 50% and three precincts below 40%.

[34] The top horizontal axis lists the mean differences between absentee and election day vote shifts for the corresponding subset of precincts.

TABLE 3. Difference-in-Differences Estimates

	Dependent variable: Republican two-party vote share			
	Full sample		Restricted sample	
	(1)	(2)	(3)	(4)
β_1 year 2017	−0.048***	−0.024	−0.063***	−0.111***
	(0.008)	(0.022)	(0.012)	(0.027)
β_2 election day	0.087***	0.171***	0.128***	0.301***
	(0.023)	(0.038)	(0.032)	(0.061)
β_3 2017 × election day	−0.036*	−0.237***	−0.042*	−0.174**
	(0.014)	(0.041)	(0.021)	(0.055)
β_7 2017 × $\%R_{i16}$ × election day		0.313***		0.220**
		(0.066)		(0.085)
N	348	348	164	164
Adjusted R^2	0.315	0.904	0.432	0.923

Note: Standard errors clustered by precinct.
* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

percentage of absentee voters, percentage living within the city limits, and mean age.[35]

Table 3 displays regression estimates for models in equations (7) and (8). Our main coefficients of interest are β_3 and β_7. In equation (7), β_3 refers to the overall effect of Gianforte's assault on his 2017 election day vote share and is estimated to be −0.036 of the two-party vote share (column 1). Thus, overall, Gianforte was punished by the loss of 3.6% of election day voters.

Our main interest is in β_7, which captures how the assault's effect varied with the 2016 Republican vote share across precincts.[36] A positive β_7 implies that Gianforte's 2017 vote loss was less severe in precincts with a higher 2016 Republican vote share. This is indeed what we observe (column 2): the more Republican a precinct was in 2016, the more forgiving election day voters were of Gianforte's assault in 2017. Columns 3 and 4 probe the robustness of these findings by restricting the sample to precincts most consistent with the parallel trend assumption.[37]

These findings are consistent with our theoretical framework and experimental findings: only moderate precincts punished Gianforte for assaulting the journalist by either abstaining or voting for a Democrat. In heavily partisan precincts, partisan loyalty trumped valence considerations.

---

[35] These controls are based on the voter file and account for time-varying factors that may differentially affect absentee and election day voters in the same precinct. In particular, there is a trend toward absentee voting in Montana: absentee ballots constituted 42.6% of all ballots cast in 2008, 47.2% in 2010, 58.9% in 2012, 60.2% in 2014, 65.4% in 2016, and 73.1% in 2017. Source: "Absentee Turnout 2,000-Present," Montana Secretary of State, accessed on November 16, 2018.

[36] In equation (8), β_3 estimates Gianforte's election day vote share in a precinct with the 2016 Republican vote share of 0 (i.e., $R_{i16} = 0$).

[37] These are precincts for which we can verify that the 2014 to 2016 difference-in-differences was less than 5%. This was the case for 41 of 68 precincts for which we have data from the year 2014. See the appendix for details and further plausibility checks for key assumptions behind difference-in-differences estimation.

CONCLUSION: IT CAN'T HAPPEN HERE?

This study investigates a conspicuous contradiction between the ideals and the practice of American democracy. At least since Tocqueville's *Democracy in America*, the United States has served as an archetype of democratic political development and an aspirational model for the rest of the world. Within the United States, a shared devotion to democratic ideals is a core part of the country's national identity. "To reject the democratic creed," to quote Dahl (1961, 317), "is in effect to refuse to be an American."

Yet the United States stands out as exceptional among advanced democracies in the perplexing recurrence of a number of democratic defects, especially at the state and local levels. Distortions of the democratic process long abandoned by most democracies—especially gerrymandering and the partisan administration and adjudication of elections—remain constitutional in the United States.[38] And even those that are indisputably unconstitutional—like voter disenfranchisement—resurface with unsettling regularity.

Our research addresses these contradictions by examining a fundamental question about the robustness of American democracy: Is allegiance to democratic principles in the United States strong enough for the electorate to serve as a check on undemocratic behavior by elected politicians? The relevance of this question rests on the key role that elections play as an instrument of democratic stability: the public can stop candidates with authoritarian tendencies by defeating them at the polls.

Our analysis of a candidate-choice experiment as well as a natural experiment consistently found that only a small fraction of Americans prioritize democratic principles in their electoral choices when doing so goes against their partisan identification or favorite policies.

---

[38] For a historical perspective on democratic development in the United States, see Bateman (2018) and Mickey (2015).

TABLE 4.   The U.S. Electorate's Resilience to Undemocratic Candidates: A Sensitivity Analysis

	Condition	% defecting from D^-	
	Overall	11.74	(9.99, 13.47)
1.	Cross-party contests	10.39	(8.12, 12.72)
2.	Platform divergence	7.27	(4.25, 10.36)
3.	Moderate party-policy alignment	3.46	(−2.64, 9.54)

We proposed that this is the consequence of two mechanisms: (i) voters are willing to trade off democratic principles for partisan ends and (ii) voters employ a partisan "double standard" when punishing candidates who violate democratic principles. These tendencies were exacerbated by several types of polarization, including intense partisanship, extreme policy preferences, and divergence in candidate platforms. Put simply, polarization undermines the public's ability to serve as a democratic check.

We conclude by discussing the implications of these findings for our understanding of democratic stability in the United States and the rest of the world. We saw that roughly 10–13% of our respondents—depending on the type of contests considered—value democracy enough to punish otherwise favored candidates for violating democratic principles by voting against them. We have interpreted this result in light of several benchmarks, including the punishment that voters meted out for two undesirable valence attributes unrelated to democracy—underpayment of taxes and extramarital affairs—as well as the trade-offs that voters made between democratic principles and their favorite party or social and economic policies. While our respondents valued democracy almost as much as some of these attributes, such comparisons potentially disguise our most concerning finding: because of the greater combined importance of policies and partisanship, most Americans will *not* cross party lines to punish a co-partisan for violating democratic principles, especially if the latter proposes appealing policies.

This conclusion should be interpreted in light of another benchmark: our candidate-choice experiment prioritized realistic, if incremental violations of democratic principles over those signaling a sharp turn toward dictatorship. This was a deliberate design choice, intended to capture key features of the process known as democratic backsliding, erosion, or degradation. This process starts in a democratic status quo, is typically initiated by or in conjunction with the executive, and entails violations of democratic principles that tilt the playing field in favor of the incumbent without being outright unconstitutional—often claiming to correct an ostensible deficiency in the democratic process (Huq and Ginsburg 2018; Levitsky and Ziblatt 2018; Svolik 2019). Accordingly, our experiment featured undemocratic practices that have been attempted or employed by politicians from both major parties, avoided normative

connotations in their presentation, and focused on departures from a democratic status quo (as opposed to improvements to democratic practice). In sum, we have attempted to mirror as realistically as possible the key features of democratic backsliding in the context of contemporary American politics.

We now consider what is arguably the most relevant benchmark: whether the American electorate values democracy enough to punish a *real-world* candidate for undermining democratic principles by defeating him at the polls. To address this question, recall that we based most of our analysis on a candidate-choice experiment in which all candidate attributes were independently and randomly assigned. A key advantage of this design is that it allows us to identify each attribute's causal effect. Its potential downside is limited external validity, which manifests in two ways: first, many of our candidates featured combinations of policies and partisanship rarely observed in the real world; second, the overall distribution of candidate policies was more centrist than is typically seen in contemporary U.S. elections.

To characterize the most plausible real-world implications of our analysis, we now progressively trim the least realistic scenarios from the more than 21,000 candidate choices made by our respondents. Table 4 lists our criteria and their implications. Condition 1 restricts attention to candidate-choice scenarios that pit a Democrat against a Republican—the most frequent type of a general election contest. Condition 2 requires at least some platform divergence by discarding any contests in which the two candidates adopted the same policy platform (a rare occurrence in real-world elections). Condition 3 generates a moderate party–policy alignment by asking that the Republican candidate be to the right of the Democrat on both economic and social policies.[39]

As we gradually restrict attention to candidate-choice scenarios that better approximate real-world elections, the punishment for candidates who violate democratic principles declines from 11.7% to 3.5%. This is because the progressive application of the three conditions in Table 4 explores the consequences of a conception of candidate polarization that we have anticipated theoretically but have not addressed empirically so far: a positive correlation between candidates' policies and partisanship. As each condition induces an increasing alignment between each candidate's policies and partisanship, the differences that respondents see between Democrats and Republicans compound, reducing voters' willingness to prioritize democratic principles over partisan interests. When candidate policy platforms and partisanship align—just as they do in the real world—the viability of the public as a democratic check declines.

In sum, our initial analysis yielded a *conservative* estimate of the U.S. electorate's vulnerability to candidates who violate democratic principles. Jointly with our findings from the 2017 Montana natural

---

[39] Note that condition 3 precludes Republican candidates from adopting the leftmost position on any issue and vice versa for Democrats.

experiment, this exploration of the most plausible real-world implications of our candidate-choice experiment suggests an explanation for the recurrence of democratic defects across the United States: in states and districts where one party enjoys a significant electoral advantage, politicians from the majority party may be effectively insulated from an electoral punishment for violating democratic principles. To get a sense of the real-world relevance of this implication, consider that in 2016 only 5.1% of US House districts were won by a margin of less than 6.9%—the smallest margin that Table 4 implies is necessary for violations of democratic principles to be electorally self-defeating. That share of districts was still only 15.2% in 2018. Put bluntly, our estimates suggest that in the vast majority of U.S. House districts, a majority-party candidate could openly violate one of the democratic principles we examined and nonetheless get away with it.

Our findings contrast sharply with political scientists' conventional measures of support for democracy. The latter are based on questions like "Democracy may have problems, but it is better than any other form of government. To what extent do you agree or disagree with this statement?" Just like citizens of other advanced democracies, Americans express by this measure an overwhelming support for democratic principles: More than 85% either "agree" or "strongly agree." In our framework, by contrast, respondents "support democracy" not when they *profess* a commitment to abstract democratic ideals but when their choices *reveal* a willingness to act on those ideals. We saw that, once confronted with candidate-choice scenarios that mirror the process of democratic backsliding by putting a price on democracy—as when a co-partisan candidate proposes a measure that violates democratic principles—only a fraction of Americans choose democracy over partisan loyalty. Our findings thus suggest that conventional measures of support for democracy have a fundamental blind spot: they fail to capture voters' willingness to act on their commitment to democracy precisely when democracy is at stake.

Our focus on the United States brings back into democratization research a case that featured prominently in the early, influential research on democratic culture (Almond and Verba 1963; Dahl 1966) but has since been conspicuously absent. Our findings raise questions about the reliability of our existing knowledge about support for democracy in both the United States and the rest of the world. On the one hand, our analysis reveals that the American voter is not an outlier: American democracy may be just as vulnerable to the pernicious consequences of polarization as are electorates throughout the rest of world (McCoy and Somer 2019; Svolik 2020).[40] On the other hand, because conventional measures of support for democracy around the world suffer from the same flaws that our study highlights, we may have been unduly confident about support for democracy worldwide. Put differently, if

only 3.5% of voters realistically punish violations of democratic principles in one of the world's oldest democracies, we should not be surprised by the public's failure to stop aspiring autocrats in new democracies (Ahlquist et al. 2018; Nalepa, Vanberg, and Chiopris 2018). Our findings suggest a sobering upper bound on what can reasonably be expected from ordinary people in defense of democracy.

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit https://doi.org/10.1017/S0003055420000052.

Replication materials can be found on Dataverse at: https://doi.org/10.7910/DVN/EEARKA.

REFERENCES

Abramowitz, Alan I., and Kyle L. Saunders. 2008. "Is Polarization and Myth?" *The Journal of Politics* 70 (2): 542–55.

Ahlquist, John S., Nahomi Ichino, Jason Wittenberg, and Daniel Ziblatt. 2018. "How Do Voters Perceive Changes to the Rules of the Game? Evidence from the 2014 Hungarian Elections." *Journal of Comparative Economics* 46 (4): 906–19.

Almond, Gabriel A., and Sidney Verba. 1963. *The Civic Culture: Political Attitudes and Democracy in Five Nations*. Newbury Park, CA: Sage.

Ashworth, Scott, and Ethan Bueno de Mesquita. 2009. "Elections with Platform and Valence Competition." *Games and Economic Behavior* 67: 191–216.

Barber, Michael J., and Nolan McCarty. 2015. Causes and Consequences of Polarization. In *Solutions to Political Polarization in America*, ed. Nathaniel Persily. New York: Cambridge University Press, 15–58.

Bateman, David A. 2018. *Disenfranchising Democracy: Constructing the Electorate in the United States, the United Kingdom, and France*. New York, NY: Cambridge University Press.

Becher, Michael, and Sylvain Brouard. 2019. "Executive Accountability beyond Outcomes: Experimental Evidence on Public Evaluations of Powerful Prime Ministers." Unpublished Manuscript, IAST Toulouse.

Bermeo, Nancy. 2003. *Ordinary People in Extraordinary Times: The Citizenry and the Breakdown of Democracy*. Princeton, N.J.: Princeton University Press.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics* 119 (1): 249–75.

Cameron, A. Colin, and Pravin K. Trivedi. 2005. *Microeconometrics: Methods and Applications*. New York: Cambridge University Press.

Carey, John M., Gretchen Helmke, Brendan Nyhan, Mitchell Sanders, and Susan Stokes. 2019. "Searching for Bright Lines in the Trump Presidency." *Perspectives on Politics* 17 (3): 699–718.

Carey, John M., Katherine P. Clayton, Gretchen Helmke, Brendan Nyhan, Mitchell Sanders, and Susan C. Stokes. 2018. "Party, Policy, Democracy and Candidate Choice in U.S. Elections." Report, Bright Line Watch.

Chen, Jowei, and Jonathan Rodden. 2013. "Unintentional Gerrymandering: Political Geography and Electoral Bias in Legislatures." *Quarterly Journal of Political Science* 8 (3): 239–69.

Cho, Wendy K. Tam, and Yan Y. Liu. 2016. "Toward a Talismanic Redistricting Tool: A Computational Method for Identifying Extreme Redistricting Plans." *Election Law Journal* 15 (4): 351–66.

Clark, Tom S. 2009. "The Separation of Powers, Court Curbing, and Judicial Legitimacy." *American Journal of Political Science* 53 (4): 971–89.

Converse, Philip E. 1964. "The Nature of Belief Systems in Mass Publics." *Critical Review* 18 (1–3): 1–74.

Dahl, Robert A. 1956. *A Preface to Democratic Theory*. Chicago, IL: University of Chicago Press.

---

[40] For a classic analysis of the relationship between polarization and democratic stability, see Bermeo (2003) and Sartori (1976).

Dahl, Robert A. 1961. *Who Governs? Democracy and Power in an American City*. New Haven: Yale University Press.

Dahl, Robert A. 1971. *Polyarchy: Participation and Opposition*. New Haven: Yale University Press.

Dahl, Robert A., ed. 1966. *Political Oppositions in Western Democracies*. New Haven: Yal.

Eggers, Andrew C. 2014. "Partisanship and Electoral Accountability: Evidence from the UK Expenses Scandal." *Quarterly Journal of Political Science* 9 (4): 441–72.

Fiorina, Morris P., Samuel A. Abrams, and Jeremy C. Pope. 2008. "Polarization in the American Public: Misconceptions and Misreadings." *The Journal of Politics* 70 (2): 556–60.

Foa, Roberto Stefan, and Yascha Mounk. 2017. "The Danger of Deconsolidation: The Democratic Disconnect." *The Danger of Deconsolidation: The Democratic Disconnect* 27 (3): 5–17.

Gandhi, Jennifer, and Elvin Ong. 2019. "Committed or Conditional Democrats? Opposition Dynamics in Electoral Autocracies." *American Journal of Political Science* 63 (4): 948–63.

Gelman, Andrew. 2008. "Scaling Regression Inputs by Dividing by Two Standard Deviations." *Statistics in Medicine* 27 (15): 2865–73.

Graham, Matthew, and Milan W. Svolik. 2019. "Turnout as a Democratic Check?" Unpublished Manuscript, Yale University.

Grimmer, Justin, Eitan Hersh, Marc Meredith, Jonathan Mummolo, and Clayton Nal. 2018. "Obstacles to Estimating Voter ID Laws Effect on Turnout." *The Journal of Politics* 80 (3): 1045–51.

Haggard, Stephan, and Robert R. Kaufman. 2016. *Dictators and Democrats: Masses, Elites, and Regime Change*. Princeton: Princeton University Press.

Hainmueller, Jens, Daniel J. Hopkins, and Teppei Yamamoto. 2015. "Causal Inference in Conjoint Analysis: Understanding Multi-Dimensional Choices via Stated Preference Experiments." *Poltical Analysis* 22 (1): 1–30.

Huq, Aziz Z., and Tom Ginsburg. 2018. "How to Lose a Constitutional Democracy." *UCLA Law Review* 65 (1): 78–169.

Inglehart, Ronald, and Christian Welzel. 2010. "Changing Mass Priorities: The Link between Modernization and Democracy." *Perspectives on Politics* 8 (2): 551–67.

Iyengar, Shanto, Yphtach Lelkes, Matthew Levendusky, Neil Malhotra, and Sean J. Westwood. 2018. "The Origins and Consequences of Affective Polarization in the United States Affective Polarization: An Outgrowth of Partisan Social Identity." *Annual Review of Political Science* 22: 129–46.

Kaufman, Robert R., and Stephan Haggard. 2019. "Democratic Decline in the United States: What Can We Learn from Middle-Income Backsliding?" *Perspectives on Politics* 17 (2): 417–32.

Leeper, Thomas J., Sara B. Hobolt, and James Tilley. Forthcoming. "Measuring Subgroup Preferences in Conjoint Experiments." *Political Analysis*.

Levitsky, Steven, and Daniel Ziblatt. 2018. *How Democracies Die*. New York: Crown.

Levitsky, Steven, and Lucan A. Way. 2010. *Competitive Authoritarianism: Hybrid Regimes after the Cold War*. New York: Cambridge University Press.

Lieberman, Robert C., Suzanne Mettler, Thomas B. Pepinsky, Kenneth M. Roberts, and Richard Valelly. 2019. "The Trump Presidency and American Democracy: A Historical and Comparative Analysis." *Perspectives on Politics* 17 (2): 470–9.

Lipset, Seymour M. 1960. *Political Man: The Social Bases of Politics*. Garden City, NY: Doubleday.

Luo, Zhaotian, and Adam Przeworski. 2018. "Subversion by Stealth: Dynamics of Democratic Backsliding." Unpublished Manuscript, New York University.

McCarty, Nolan, Keith T. Poole, and Howard Rosenthal. 2008. *Polarized America: The Dance of Ideology and Unequal Riches*. Cambridge, MA: MIT Press.

McCoy, Jennifer, and Murat Somer. 2019. "Polarizing Polities: A Global Threat to Democracy." *The Annals of the American Academy of Political and Social Science* 681: 1–273.

Mickey, Robert. 2015. *Paths Out of Dixie: The Democratization of Authoritarian Enclaves in America's Deep South, 1944–1972*. Princeton, NJ: Princeton University Press.

Miller, Michael K., David Szakonyi, and Lee Morgenbesser. 2017. Authoritarian Warning Survey.

Nalepa, Monika, Georg Vanberg, and Caterina Chiopris. 2018. "Authoritarian Backsliding." Unpublished Manuscript, University of Chicago and Duke University.

Norris, Pippa. 2011. *Democratic Deficit: Critical Citizens Revisited*. New York: Cambridge University Press.

Przeworski, Adam. 1991. *Democracy and the Market: Political and Economic Reforms in Eastern Europe and Latin America*. Cambridge: Cambridge University Press.

Przeworski, Adam. 2019. *Crises of Democracy*. New York: Cambridge University Press.

Sartori, Giovanni. 1976. *Parties and Party Systems: A Framework for Analysis*. Cambridge, U.K.: Cambridge University Press.

Schedler, Andreas. 2002. "The Menu of Manipulation." *Journal of Democracy* 13 (2): 36–50.

Svolik, Milan. 2019. "Polarization vs. Democracy." *Journal of Democracy* 30 (3): 20–32.

Svolik, Milan W. 2020. "When Polarization Trumps Civic Virtue: Partisan Conflict and the Subversion of Democracy by Incumbents." *Quarterly Journal of Political Science* 15 (1): 3–31.

Treier, Shawn, and D. Sunshine Hillygus. 2009. "The Nature of Political Ideology in the Contemporary Electorate." *Public Opinion Quarterly* 73 (4): 679–703.

Voeten, Erik. 2017. "Are People Really Turning Away from Democracy?" *Journal of Democracy Web Exchange*: 1–16.

Waldner, David, and Ellen Lust. 2018. "Unwelcome Change: Coming to Terms with Democratic Backsliding." *Annual Review of Political Science* 21: 93–113.

Weingast, Barry R. 1997. "The Political Foundations of Democracy and the Rule of Law." *American Political Science Review* 91 (2): 245–63.