Democracy in America?

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

Matthew Graham and Milan W. Svolik*

First draft: March 2018, current version: March 2019

Abstract

Is 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.

*Department of Political Science, Yale University. Corresponding email: milan.svolik@yale.edu.

"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)

"I could stand in the middle of 5th Avenue and shoot somebody and I wouldn't lose voters."

Donald Trump at a presidential campaign rally in Iowa (January 23, 2016)

1 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).

In this paper, we show that this conventional wisdom rests on fragile foundations. We adopt an approach that, instead of asking about support for democracy directly, infers respondents' 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 include positions that violate core democratic principles, including free and fair elections, civil liberties, and checks and balances. In this framework, voters are said to "support democracy" when their

choices *reveal* a preference for democratic principles over other valid but potentially conflicting considerations such as political ideology, partisan loyalty, or favorite policies.

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 one who complies with democratic principles but is unappealing, a significant fraction of voters may be willing to sacrifice democratic principles in favor of electing 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 section 2, we formalize these intuitions and develop a model of the public as a democratic check. 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 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 microfoundations for classic arguments about the importance of political moderation and cross-cutting cleavages for democratic stability (Dahl, 1956; Lipset, 1960).

We employ this framework to guide 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. In the candidate-choice experiment, respondents were asked to choose between a series of candidates for a state legislature. Candidates were described by attributes typically seen in real-world elections, including their social and economic policy platforms, partisan affiliation, and – crucially – a "democracy" position. The latter endorsed an action that violates a key democratic principle and approximates real-world practices used by politicians in the United States to subvert the democratic process. Because all candidate attributes were experimentally manipulated, we can interpret the decline in undemocratic candidates' vote share as a revealed-preference measure of the U.S. electorate's ability to serve as a democratic check. The following is a summary of our experimental findings, which we present in section 3:

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% 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 5.5%.

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: Voters who are moderate in their policy preferences or partisanship punish undemocratic candidates at rates three-four times higher than more extreme voters.

4. Most voters are partisans first and democrats only second: Less than 15% of our respondents are willing to punish a co-partisan for violating democratic principles when the price of that punishment is voting against their own party. Only independents and partisan "leaners" support the more democratic candidate enough to defeat the undemocratic candidate regardless of his partisan affiliation.

5. Voters 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. In most of our specifications, these effects hold equally for partisans of both major parties.

6. Platform polarization is bad for democracy: The larger the difference between the candidates' policy platforms, the weaker the punishment for undemocratic behavior.

7. Sensitivity to the menu of manipulation varies: Voters are most sensitive to undemocratic positions that undermine the free press, checks and balances, and those that aim to disenfranchise opposition supporters. Nonetheless, when we benchmark these against extramarital affairs and underpaying 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. Strong partisans punish undemocratic behavior by abstaining: The stronger a respondent's preference for a candidate, the more likely she is to abstain rather than defect when that candidate adopts an undemocratic position.

9. 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 from those that are consistent with democratic principles.

We take advantage of the close connection between the design of our candidate choice experiment and our theoretical framework in section 4, where we shift from the primarily non-parametric analysis employed up to then to a structural approach that identifies the model's key primitives. The latter approach allows us to explicitly estimate the weight that voters place on democracy relative to other desirable candidate attributes. We estimate this weight at 18% and find that candidates' democracy positions, economic and social policy platforms, and partisan affiliation jointly account for almost 80% of the systematic variation in voters' candidate choices. This approach also allows us to put a "price" on voters' value for democracy in the terms of desirable candidate characteristics they are willing to forgo to punish candidates who violate democratic principles. Results from such a structural analysis are consistent with our non-parametric findings: while voters are willing to forgo about a two rank increase in the distance from their ideal economic and social policies to avoid a candidates who violate democratic principles, they are not willing to vote across party lines to do so. We also revisit our experimental findings about a partisan double standard in the punishment of candidates who violate democratic principles and extend our theoretical model to account for this phenomenon. We estimate a 43% co-partisan bias, implying that Americans are neither fully principled nor purely partisan in their punishment of candidates who violate democratic principles support for democracy.

We move from analyzing hypothetical election scenarios to a real-world election in section 5, where we examine 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 in-person voters saw this signal before they could cast a ballot; 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 attack on the journalist. Our findings are consistent with both our theoretical expectations and experimental results. We find that Montanans value a free press but only moderate partisans are willing to punish the assault on the journalist by voting across party lines. For strong partisans, partisan loyalty trumps 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 paper joins a growing number of papers that have, 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. While 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 willingness to punish politicians for subverting democratic competition in their favor.

Our empirical methodology and substantive focus are closest to Carey, Clayton, Helmke, Nyhan, Sanders and Stokes (2018), who also employ a candidate-choice experiment to study the commitment to democratic principles among the American public.

[1] See especially Carey, Helmke, Nyhan and Stokes (2018), Huq and Ginsburg (2017), Kaufman and Haggard (Forthcoming), Levitsky and Ziblatt (2018), Lieberman et al. (Forthcoming), Miller et al. (2017), and Przeworski (2017).

[2] See, for instance, the exchange between Alexander and Welzel (2017), Foa and Mounk (2017), Norris (2017), and Voeten (2017).

Just like we do, they report that while voters do punish candidates whose positions violate democratic principles, the magnitude of that punishment may be overshadowed by other political considerations – most notably partisanship (c.f. Barber and Pope, Forthcoming).[3] Taken together, this evidence suggests that 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 serve as a check on the authoritarian temptations of elected politicians?

Our theoretical framework helps us 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 ensure their defeat. We explain why this 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.[4] The microfoundations that we develop in section 2 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, 2018; Haggard and Kaufman, 2016; Luo and Przeworski, 2018; Nalepa et al.,

[3]While Carey, Clayton, Helmke, Nyhan, Sanders and Stokes (2018) examine a related set of democratic principles, their implementation is different from ours. Their democratic treatments probe voters' attitudes toward voter ID requirements, partisan influence in law enforcement investigations, elected officials' deference to court rulings, and willingness to compromise across party lines.

[4]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 UK expenses scandal less severely when the electoral stakes in their district were higher.

2018; Waldner and Lust, 2018).

This theoretical perspective also guides our empirical research design. The challenge in assessing the public's resilience to undemocratic candidates in advanced democracies is that we rarely observe outright violations of democratic principles in this class of regimes. To put it in the jargon of formal theory, undemocratic behavior by political elites is "off the equilibrium path" and so is the public's reaction to it. Yet if the elite's beliefs about the public's off-equilibrium reactions are what keeps them in check in the first place, then we lack an empirical basis for assessing the robustness of the public's support for democracy. Our candidate-choice experiment overcomes this difficulty: It allows us to investigate "out-of-equilibrium" responses by confronting voters with candidates who may adopt undemocratic positions that are plausible yet rarely encountered in advanced democracies.

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 level. These include gerrymandering (Chen and Rodden, 2013; Cho and Liu, 2016), misinformation (Albertson and Guiler, 2018), voter suppression (Hajnal et al., 2017; Grimmer et al., 2018), and voter fraud (Ansolabehere and Persily, 2008; Ahlquist et al., 2014).[5] 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.[6] Yet our empirical analysis reveals that this check is at best limited in magnitude and subject to a partisan "double standard." In turn, 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 effects of a number of distinct

[5] See Norris et al. (2018) for a recent assessment of electoral integrity in the United States. For a historical perspective on democratic development in the United States, see Mickey (2015).

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

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 an electorate's resilience to undemocratic candidates.

This broad look at 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 (Iyengar et al., 2018) and origins (Abramowitz and Saunders, 2008; Fiorina et al., 2008; McCarty et al., 2008), our focus is on a political consequence that this literature has yet to examine.[7] 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.

2 A Model of the Public as a Democratic Check

Consider a model according to which voters' preferences over two kinds of candidate attributes determine their electoral choices: i) positional issues, which may include candidates' policy positions, ideology, and partisan affiliation, and ii) candidates' compliance with key democratic principles. We conceive of the latter as a valence issue:

[7] This characterization of the literature intentionally omits work on the polarization in Congress. For a review of this literature, see Barber and McCarty (2015).

while voters may differ over policy, ideology, or partisanship, they all agree that electoral competition should be democratic and prefer candidates whose positions comply with key tenets of democratic electoral competition.

Formally, voter i's payoff from candidate j is

$$u_i(X_j, M_j) = -\sum_K \alpha_k (x_{ik} - x_{jk})^2 - \delta M_j \,, \tag{1}$$

where 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 *un*democratic j's platform is (i.e. M stands for "manipulation.")[8] The term δ is the weight that voters attach to fair democratic competition – in effect, the intensity of their support for democracy.

This simple model yields several predictions that we evaluate throughout the paper. 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, assume only a single policy issue. Then i votes for candidate 1 as long as

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

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 – it separates the electorate into those who

[8] In turn, M is equivalent to negative valence in models of electoral competition with valence; see e.g. Ashworth and Bueno de Mesquita (2009).

are policy-wise closer to candidate 1 and those who are closer to candidate 2. The swing voter x_s^j is in turn located either to the right or the left of this midpoint, depending on whether it is candidate 1 or candidate 2 who adopts an undemocratic platform,

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

When candidate 1 adopts an undemocratic platform ($M_1 = 1$, $M_2 = 0$), 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 between the two candidate's platforms 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 to vote for candidate 2 instead. By contrast, voters whose policy preferences are extreme (large x_{ik}) or intense (a 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 vote for the more democratic of the two candidates regardless their policy proximity. We may therefore refer to this portion of the electorate as "democracy first" voters, as opposed to "policy first" voters. How large that portion is depends on both the length of this segment and the fraction of voters located on it. 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 fraction of voters located within this segment is determined by the distribution of voters' ideal policies: the more polarized (U-shaped) the distribution is, the smaller the fraction of democracy first voters. The consequences of voter polarization are amplified when voters' preferences over the K policy issues are correlated: when a voter's

extreme position on one issue correlates with her extreme position on another, her preference for the more proximate candidate will be compounding across the two issues rather than cancelling 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 – these will be randomly assigned in the candidate-choice experiment that we present and analyze in sections 3-4. Second, in order 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 briefly discuss its consequences.

We follow the classic random utility models of discrete choice and add to voter i's payoff from candidate j an error term ϵ_{ij},

$$u_i(X_j, M_j) = -\sum_K \alpha_k (x_{ik} - x_{jk})^2 - \delta M_j + \epsilon_{ij} \,. \tag{4}$$

While only one of several plausible stochastic structures, assuming that the error terms ϵ_{ij} are (independently) drawn from the standard Gumbel 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.[9] We take advantage of this correspondence between our theoretical framework and the random utility model of discrete choice in section 4, where 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

[9]See Cameron and Trivedi (2005, 476-478, 486-487). The standard Gumbel distribution is also known as the generalized extreme value type-1 distribution.

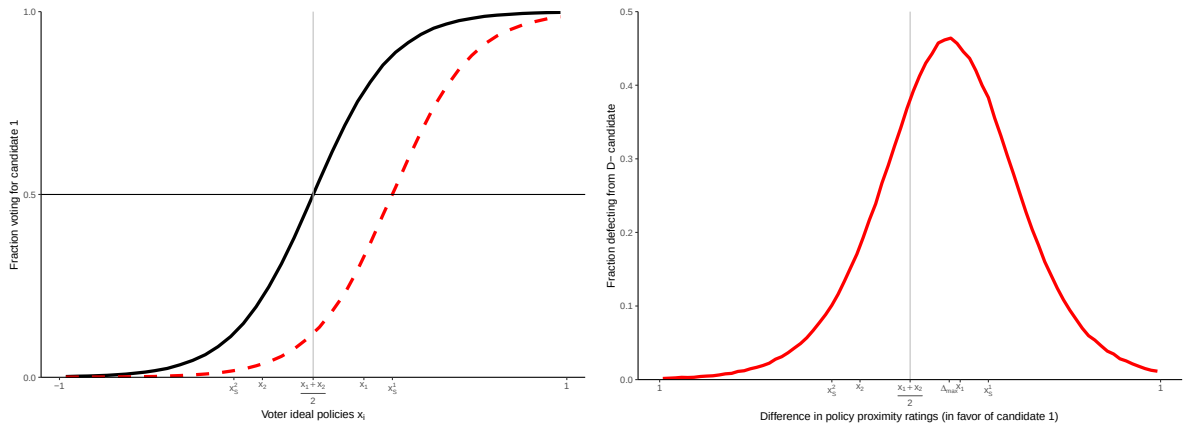

Figure 1: The fraction voting for candidate 1 (left) and the decline in the less democratic candidate's vote share (right)

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 voter ideal policies x_{ik} over a single issue k. An alternative interpretation of the vertical axis that we employ throughout is that it corresponds to the fraction of voters in subgroups along the horizontal axis that support candidate 1. The black solid line plots the case when both candidates adopt a democratic position; the red dashed line plots the case when candidate 1 adopts an undemocratic position. The plot is based on parameter values $\alpha_k = 8$, $\delta = 2$, $x_{1k} = 1/5$, $x_{2k} = -1/5$ and implies $x_s^1 = -x_s^2 = 5/16$.

When analyzing the candidate-choice experiment in sections 3-4, it will often be more practical to use as the horizontal axis the difference in voters' policy proximity to the two candidates. In that case, the intersection of the red dashed curve with the .5 horizontal line will, instead of the swing voter x_s^1, mark the quantity $\frac{\delta}{\alpha_k}$, which we interpret in section 4 as voters' value for democracy in terms of issue k.

The right panel in Figure 1 plots the decline in candidate 1's vote share when he adopts an undemocratic position. It shows that while voters with centrist policy preferences will in general defect from undemocratic candidates at higher rates than extremists, the maximum defection rate (marked by Δ_{max}) 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.[10] Intuitively, it is voters who would otherwise narrowly break in favor of candidate 1 that abandon him at the highest rates once he adopts an undemocratic position.

A key quantity that we estimate throughout the paper is the overall fraction of voters who defect from a candidate that adopts an undemocratic position, which we denote by Δ. Δ is a combination of two factors: the defection rates portrayed in the right panel in Figure 1 and the empirical distribution of voters.[11] For instance, a simulation that uses the symmetrical $Beta(\gamma, \gamma)$ density to vary the distribution of voters' ideal policies from one with a mass of centrists ($\gamma = 5$) to a U-shaped one ($\gamma = \frac{1}{5}$) yields a decline in Δ from .28 to .06. In effect, Δ measures the public's *resilience* to undemocratic candidates.

To recapitulate, this theoretical framework yields five predictions about the relationship between a number of distinct conceptions of polarization and a decline (denoted by $\Delta \downarrow$ below) in the electorate's resilience to undemocratic candidates:

1. **Centrists are a pro-democratic force:** $\Delta \downarrow$ for voters who hold more *extreme* policy preferences (large $|x_{ik}|$);

2. **Moderation is good for democracy:** $\Delta \downarrow$ for voters who hold more *intense* policy preferences (a large α_k);

[10] Given our stochastic structure, Δ_{max} has a closed form solution $\Delta_{max} = \frac{x_{1k}+x_{2k}}{2} + \frac{\delta}{4\alpha_k(x_{1k}-x_{2k})}$. This corresponds to the value $\frac{\delta}{2\alpha}$ when the horizontal axis refers to the difference in voters' policy proximity to the two candidates.

[11] 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 empirical distribution of voters and $D = \sum_K \alpha_k (x_{ik} - x_{2k})^2 - \sum_K \alpha_k (x_{ik} - x_{1k})^2$.

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 preference over distinct issue areas *correlate*.

3 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. The experiment examines this mechanism at the same level at which it is hypothesized to operate: that of the individual voter. By modelling one of the most essential and familiar actions that voters perform – the choice between two candidates in an election – the conjoint-based design we introduce below probes our respondents' willingness to trade off democratic principles for partisan interests without alerting them to our focus on that aspect of their choice.[12]

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 focus of our analysis; we therefore describe its design and assignment below. We introduce most of the remaining attributes throughout the paper. In the appendix, we outline the design and assignment of all attributes and

[12] Our candidate-choice experiment belongs to a broader category of survey-experimental techniques known as conjoint experiments (Hainmueller et al., 2015).

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

present 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.[13] 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 read "Said the Republican governor should ignore unfavorable court rulings by Democrat-appointed judges."

[13]The only difference between positions 1a and 1b is in the number of extra seats that a candidate's party obtains, a distinction in scale that we examine in section 3.4. The only difference between positions 6a and 6b is in whether the candidate advocates banning far-left or far-right group rallies. Republican candidates always advocated banning far-left group rallies, Democrats far-right rallies.

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).[14]

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 numerous 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, to be implemented, such measures must often be conducted 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

[14]For recent perspectives on how to conceptualize and measure democracy, see Boix et al. (2013), Cheibub et al. (2010), and Coppedge et al. (2011).

stricter voter ID laws respond to accusations of voter suppression by claiming that such measures are needed to prevent voter fraud, and proponents of gerrymandering may claim they are correcting an existing, 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. This entailed several steps. In order to avoid 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 of these positions 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. This is also why we randomized the order in which candidates' democracy and policy positions were listed.

The wording of our undemocratic positions also avoids negative connotations or normatively leading language. For instance, positions 1a, 1b, and 2 are instances of 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.

Each respondent made 16 distinct candidate choices of which 11 were based on the following experimental design:[15] In four randomly chosen scenarios, both candidates adopted one of the democratically neutral, "generic" positions. Throughout, we treat these

[15]The remaining five scenarios featured designs intended to provide extensions and robustness checks of our core design. We introduce them in section 3.4 and discuss in detail in the appendix.

as our control scenarios and label them D^+ vs. D^+. In seven randomly chosen scenarios, one of the candidates adopted one of our undemocratic positions while 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 and analysis 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.[16] The 1,692 respondents made a total of 21,151 candidate choices.

3.1 Democratic Principles versus 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 in a candidate-choice scenario. For each policy area, candidates were independently and randomly assigned to propose one of four possible platforms, ranging from extreme liberal to extreme conservative positions.[17]

[16]The survey was implemented via LUCID. 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 4-25, 2018, was the candidate-choice experiment. A pilot survey, implemented via Amazon Mechanical Turk, took place in March 2018. The appendix benchmarks our sample against demographic data from the US Census Bureau and partisan and attitudinal questions from the ANES.

[17]The four platforms on taxes, for instance, were "increase the state income tax on households earning over $250,000 and increase the state corporate tax," "increase the state income tax on households earning over $250,000," "cut the state income tax for all households," and "eliminate the state income tax. "

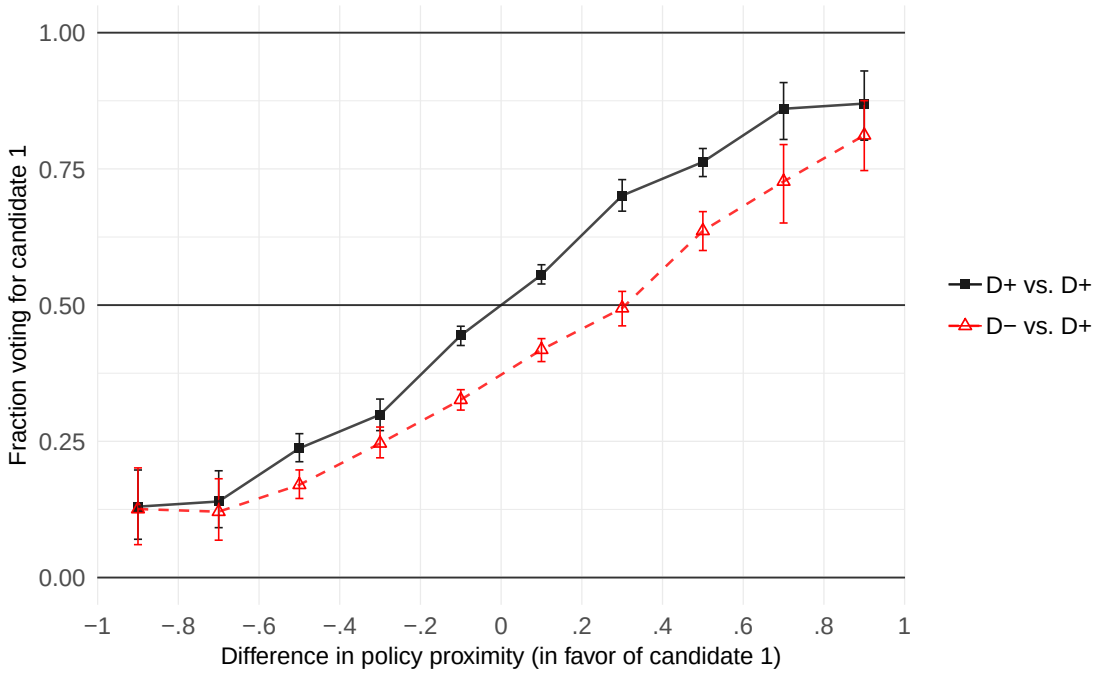

Figure 2: Fraction voting for candidate 1 by the difference in respondents' policy proximity between candidate 1 and candidate 2

The appendix lists all 16 policy platforms and discusses our reasons for their selection.

One to three weeks before being presented with the candidate-choice scenarios, each respondent was asked to express their support on a 0-100 proximity scale for each of the 16 policy platforms that the candidates might adopt. This allows us to identify each respondent's ideal policy in each of the four areas and, following the theoretical framework in section 2, to compute the squared distance between a respondent's ideal policy and each candidate's platform for each of the candidate-choice scenarios that the respondents will face. 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 non-separable from partisanship. We present and discuss these alternative measures in the appendix.

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.[18] On the horizontal axis, a value of 0 refers to scenarios when the two candidates are equally proximate to the respondent, 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.[19] We treat the D^+ vs. D^+ scenario (black 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 (red dashed line) as our treatment condition. Vertical bars denote 95% confidence intervals.[20] Figure 2 is thus the policy analogue of the left panel in 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 11% when candidate 1 is a full scale less proximate to the respondent than candidate 2 to 89% in the opposite case. Furthermore, when the two candidates are equally proximate to respondents, the latter act accordingly and split for the two 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 the only systematic difference between our control and treatment condition is candidate 1's democracy position, we can interpret a change in the fraction of

[18] "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.

[19] That is, policy proximity is the difference between a respondent's average squared distance from each of the candidates' two policy platforms, $\left[\sum_k (x_{ik} - x_{2k})^2 - \sum_k (x_{ik} - x_{1k})^2\right]/2$, normalized to range between -1 and 1. The term k refers to the assigned economic and social policy issue areas.

[20] Because each respondent made multiple choices, estimates that treat all observations as independent may understate statistical uncertainty. We therefore 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 et al., 2004). When they can be computed, we also report cluster-robust standard errors.

voters who support candidate 1 as a measure of the public's ability to serve as a democratic check. This change amounts to a 11.81% decline in candidate 1's vote share when he adopts an undemocratic position (CI: $10.64, 12.69$). All else equal, a candidate who adopts an undemocratic position can expect a virtually certain electoral defeat.

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 $(-.25, .25)$ 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 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.[21] Figure 3 plots the defection rate. Consistent with our theoretical analysis in section 2, 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.

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

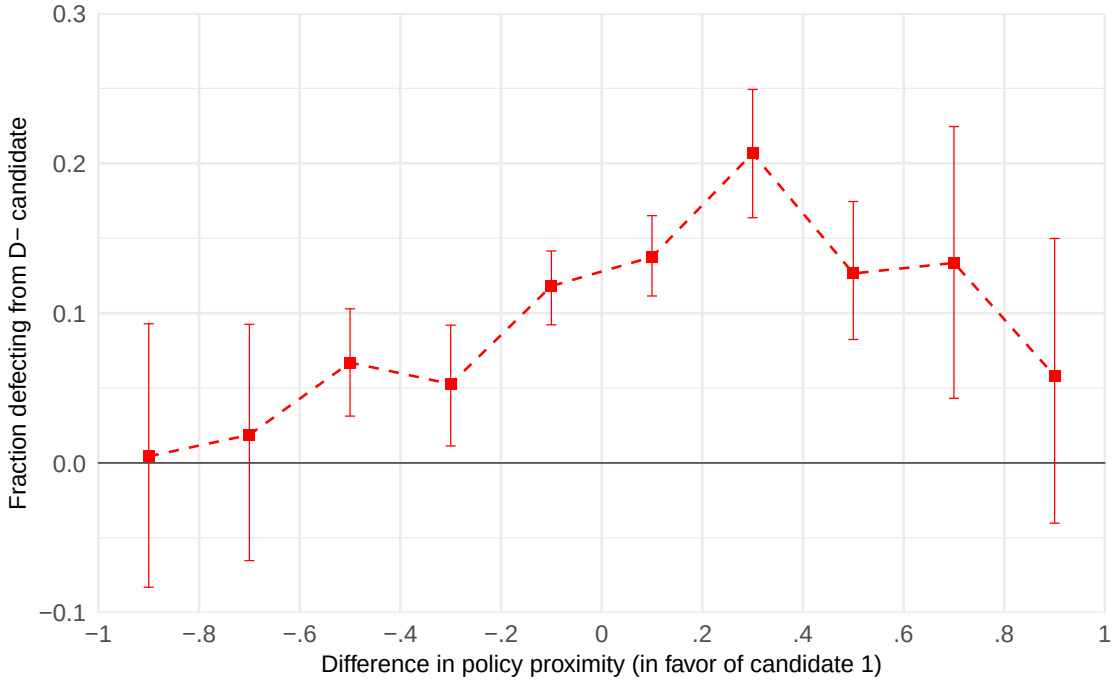

Figure 3: Fraction defecting from the less democratic candidate by the difference in respondents' policy proximity to candidate 1 (vertical bars denote 95% confidence intervals, the blue dotted horizontal line plots the overall fraction defecting)

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 3, an electorate consisting entirely of "policy centrists" would result in a resounding defeat of a candidate who would adopt an undemocratic platform. By contrast, an undemocratic candidate has a positive chance of prevailing in an electorate consisting entirely of the most extreme subgroups on each side of Figures 2 and 3.

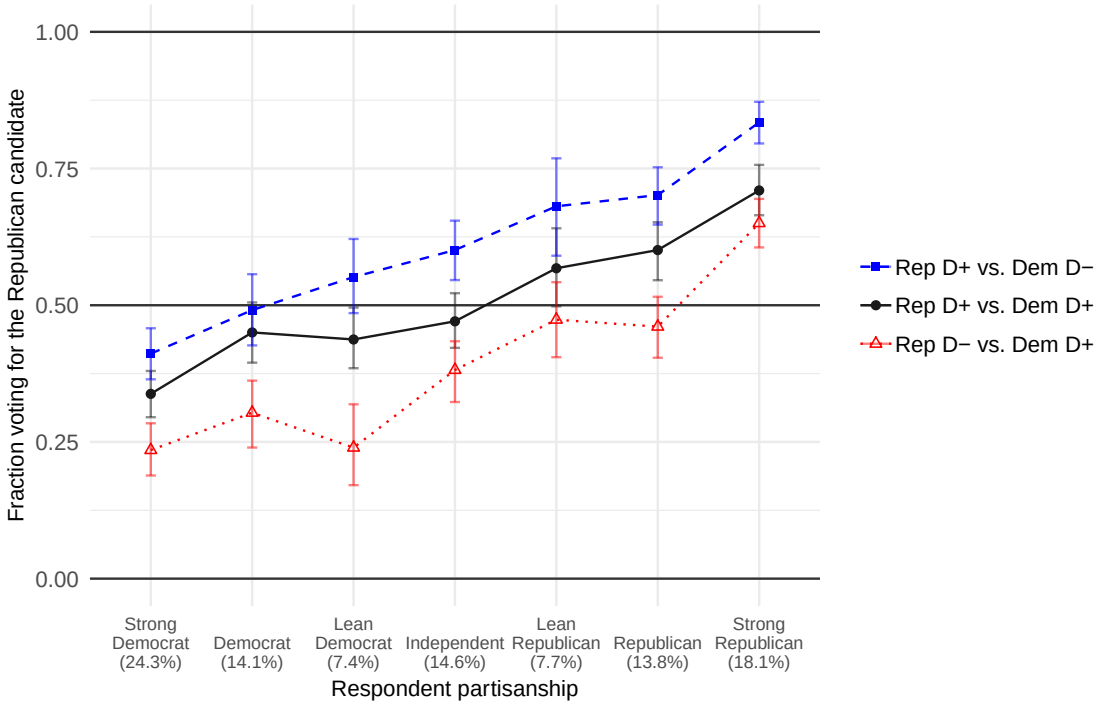

Figure 4: Different party contests: Fraction voting for a Republican Party candidate

3.2 Does Partisanship Trump Civic Virtue?

Are voters willing to put democratic principles ahead of their partisan loyalties? To address this question, consider first contests between candidates from different parties. Denote by Rep D^+ vs. Dem D^+ the 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 9.70% (CI: 7.04%, 12.22%) and 11.28% (CI: 8.44%, 13.84%) of voters in the Rep D^+ vs. Dem D^- and Rep D^- vs. Dem D^+ scenarios, respectively. Both effects are statistically different from zero but not statistically different from each other (difference: 1.58%, CI: -2.90%, 6.19%). 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? 62.69% of our respondents support their own party in the control condition; this number declines to 54.43% when a respondent's co-partisan adopts an undemocratic position. Put differently, only 13.18% of our respondents are willing to defect from a co-partisan for violating democratic principles when the price is voting against their own party (CI: 9.03, 17.25).[22]

Figure 4 yields further insights 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 blue dashed and red 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. Furthermore, 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 "Democrat" or "Republican" only break even. And a majority of "strong" partisans would rather elect a candidate that 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 to compensate for their high baseline support for co-partisans in the control condition. Second, strong partisans do exactly the opposite: they are more lenient on violations of democratic principles by candidates from their party. Figure 5 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

[22]The defection rate of 13.18% corresponds to $(62.69 - 54.43)/63.69$. Among respondents who voted across party lines in the control condition, this defection rate is much higher: opposing party candidates' vote share declines from 37.31% in the control condition to 24.78% in the treatment condition, implying a defection rate of 33.58% (CI: 26.78, 39.53).

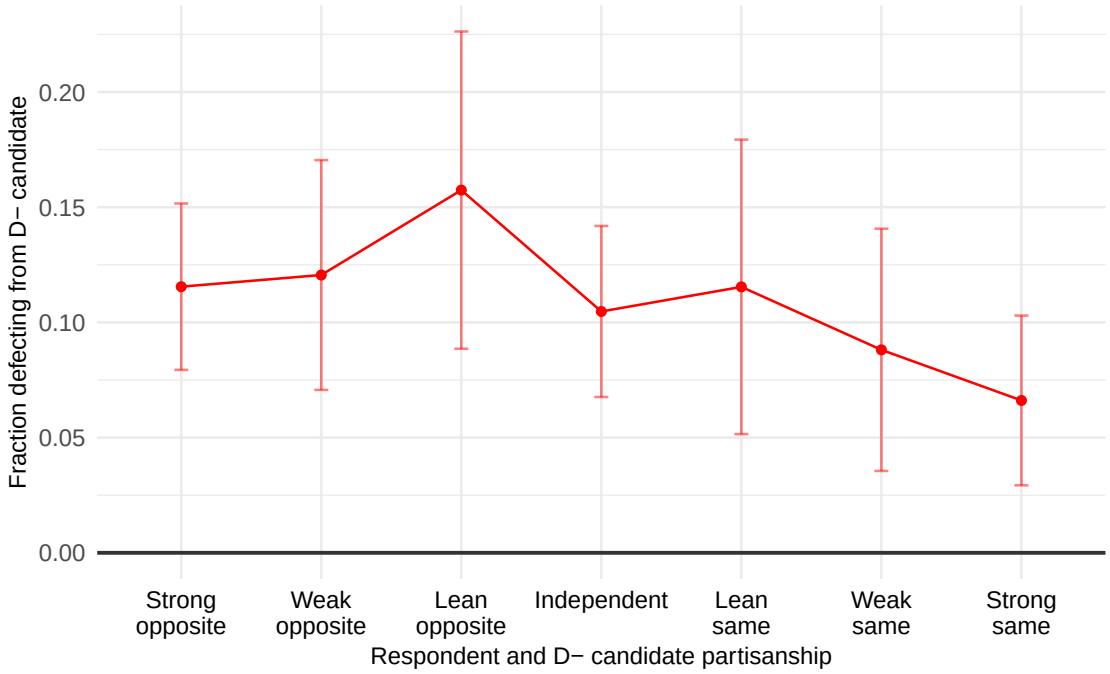

Figure 5: Different party contests: Fraction defecting from the less democratic candidate by the respondent's partisanship

are 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 most viable check on candidates who undermine democratic principles. Consistent with our theory, our respondents are more willing to punish undemocratic candidates in same-party than different-party contests: a candidate who adopts an undemocratic position is penalized by a loss of 13.34% and 10.48% of voters, respectively (difference: 2.86%, CI: 0.09%, 4.79%). Figure 6, 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

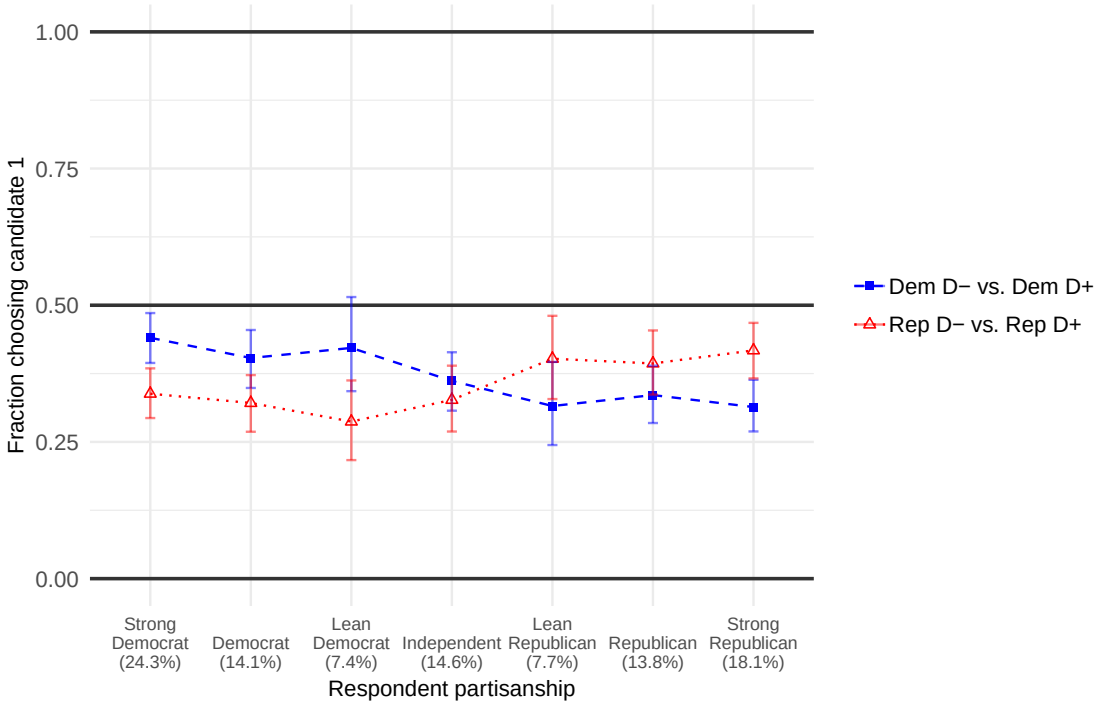

Figure 6: Same party contests: Fraction voting for candidate 1 by the respondent's and the D^- candidate's partisanship

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 the punishment for a Democrat who adopts an undemocratic position (Dem D^- vs. Dem D^+ in a blue 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 red dotted line.) The promise of primaries as a democratic check is thus undermined by a partisan double standard: voters who would punish undemocratic candidates the most are typically precluded from participating in the primary in which they would actually do so.

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 only strong partisans would provide only a tenuous check on candidates who violate democratic principles. This phenomenon was not anticipated by our theory. In section 4, we extend the framework in section 2 to account for a partisan double standard by distinguishing between principled and instrumental punishment for candidates that violate democratic principles and estimate the magnitude of a partisan bias in the electorate.

3.3 The Pernicious Consequences of Candidate Polarization

An important 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 model in section 2 implies that greater candidate polarization results in a greater share of voters who are willing to tolerate undemocratic behavior.[23] Crucially, in both the model and our experimental results, these consequences of candidate polarization are independent of voter polarization – they obtain even if we hold voter polarization fixed.

Our analysis in section 3.2 of contests between candidates from the same versus different parties found support for this prediction in terms of partisanship: we saw that respondents were less willing to punish undemocratic candidates in different-party than

[23] This is the effect of the term $(x_{1k} - x_{2k})$ in the denominator in (3).

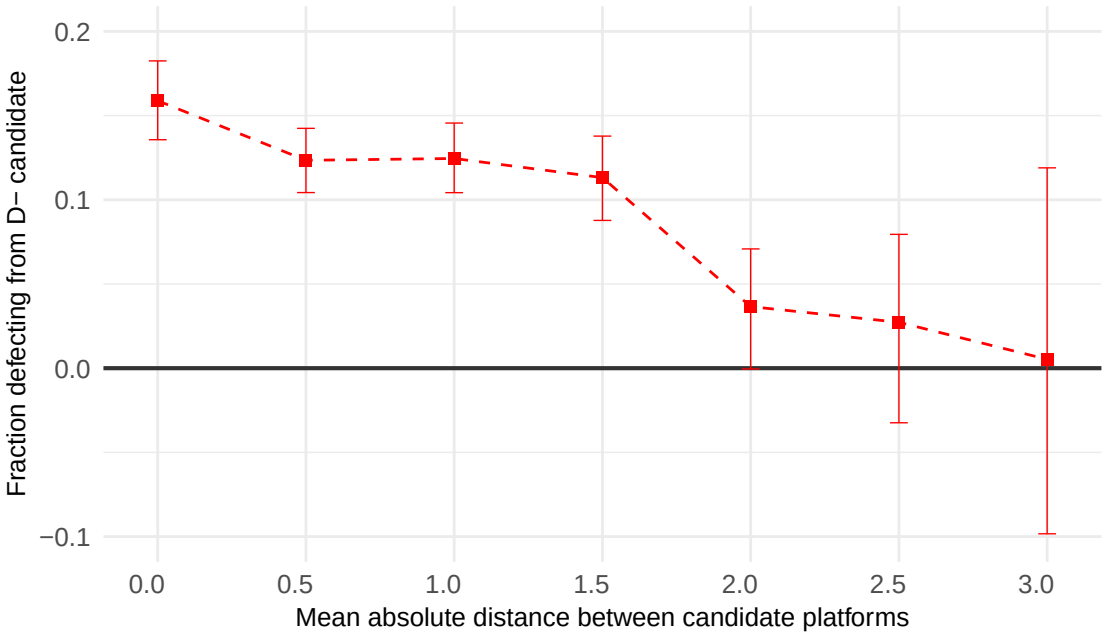

Figure 7: The effect of the mean absolute distance between the candidates' two policy platforms on the fraction of respondents defecting from a D^- candidate

same-party contests. Figure 7 provides further support for this prediction by showing the consequences of polarization in candidates' policy platforms. The horizontal axis plots the mean absolute distance between the candidates' two policy platforms;[24] the vertical axis plots the fraction of respondents defecting from a D^- candidate. We see that 32.2% of respondents defect from a D^- candidate when the two candidates' policies are identical (CI: 27.3%, 37.1%); the defection rate declines to levels that are statistically indistinguishable from 0 when both policies are as far apart as possible. Consistent with our theoretical framework, voters become more reluctant to punish candidates who violate democratic principles as candidates' policy platforms move apart.[25]

3.4 Resisting the Menu of Manipulation[26]

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 positions. We now examine the differences in Americans' willingness to tolerate the distinct ways in which the individual undemocratic positions violate democratic principles and interpret them in light of several benchmarks.

We estimate the following linear model:

$$\Pr(i \text{ votes for candidate 1}) = \alpha + \sum_k \beta_k (X_{i1k} - X_{i2k}) + \epsilon_{ij} \,. \tag{5}$$

In (5), X_{i1k} and X_{i2k} are dummy variables for all possible values of experimentally

[24] Specifically, the mean absolute distance between the candidates' two policy platform is $|\sum_k (x_{2k} - x_{1k})|/2$, where the term k refers to the assigned economic and social policies.

[25] In the appendix, we supplement Figure 3 with a formal test of whether punishment rates decline with the distance between the candidates. We find a statistically significant effect using each of our alternative measures of policy distance.

[26] This subtitle paraphrases the title of Schedler's (2002) seminal article.

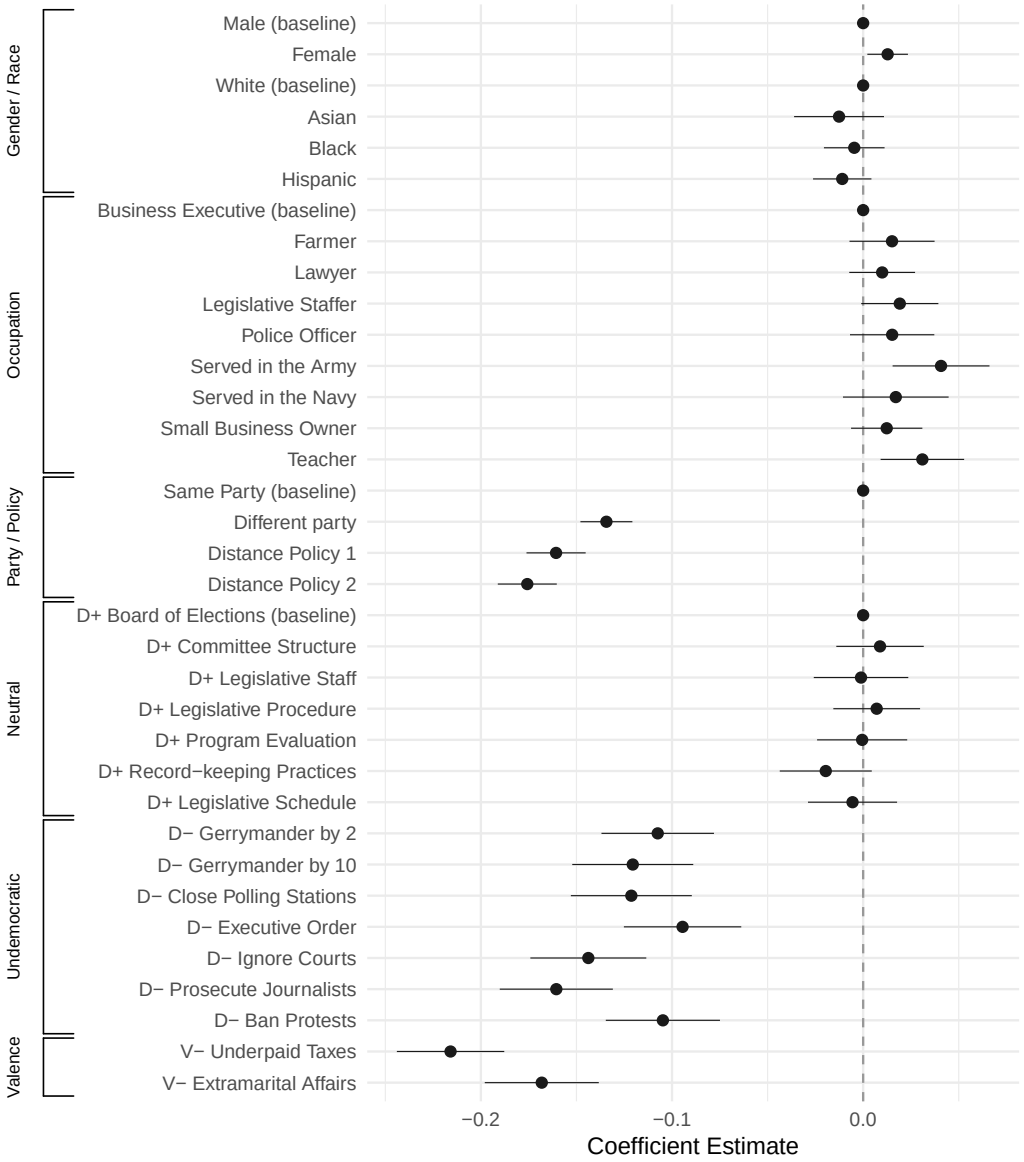

Figure 8: Effects of candidate attributes and democracy positions on a candidate's vote share. Dots represent coefficient estimates based on a linear regression model; bars represent 95% confidence intervals

manipulated attribute k in choice i between candidates 1 and 2, respectively.[27] Figure 8 plots the estimates of β_k. Bars represent the associated 95% confidence intervals;[28] estimates without confidence intervals correspond to baseline categories. We interpret these coefficients as attribute k's average effect on a candidate's vote share, relative to the baseline attribute and averaging over all other attribute levels (Hainmueller et al., 2015).

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 8 also demonstrates a 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 9.5% to 16.0%. Respondents punish most severely candidates who want to prosecute journalists (14.3%) and ignore court rulings (16.0%). Respondents are least sensitive to candidates who endorse gerrymandering (by 2 seats) and suggest that the governor ban protests or rule by executive order. Figure 9 differentiates these estimates by respondents' partisanship. Consistent with our earlier findings, we see few differences between supporters of the two major parties.

To put the magnitude of these estimates in context, compare the effect of these undemocratic positions to that of other positional and valence candidate attributes. Consistent with our discussion in sections 3.1 and 3.2, the two main positional attributes –

[27] Our estimates are substantively identical when we estimate the effect of each candidate's attributes separately rather than taking the difference as we do above. This latter approach is more concise and closer in interpretation to the model in section 2. We discuss it in detail in the appendix.

[28] These are almost identical when using block bootstrapped standard errors or normal approximations with clustered standard errors. The appendix presents complete numerical results with both sets of standard errors.

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.

Of the attributes assigned in the core 11 of the total 16 candidate choices that our respondents made, the most naturally interpretable as valence characteristics are candidate age, years of experience, and profession. From among the nine professions, only military service and being a teacher are statistically significant but their effects are an order of magnitude smaller than that of undemocratic positions. Due to space constraints, Figure 8 omits candidates' age and years of experience. With a few exceptions, effects of these attributes are also close to zero and not statistically significant.[29]

To help us further interpret the magnitude of undemocratic positions, we included in two of the 16 choices that our respondents made 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 Figure 8 and are labelled V^-. We see that, on average, voters punish candidates for extramarital affairs and underpaying taxes *more severely* than they punish them for undermining democratic principles.

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

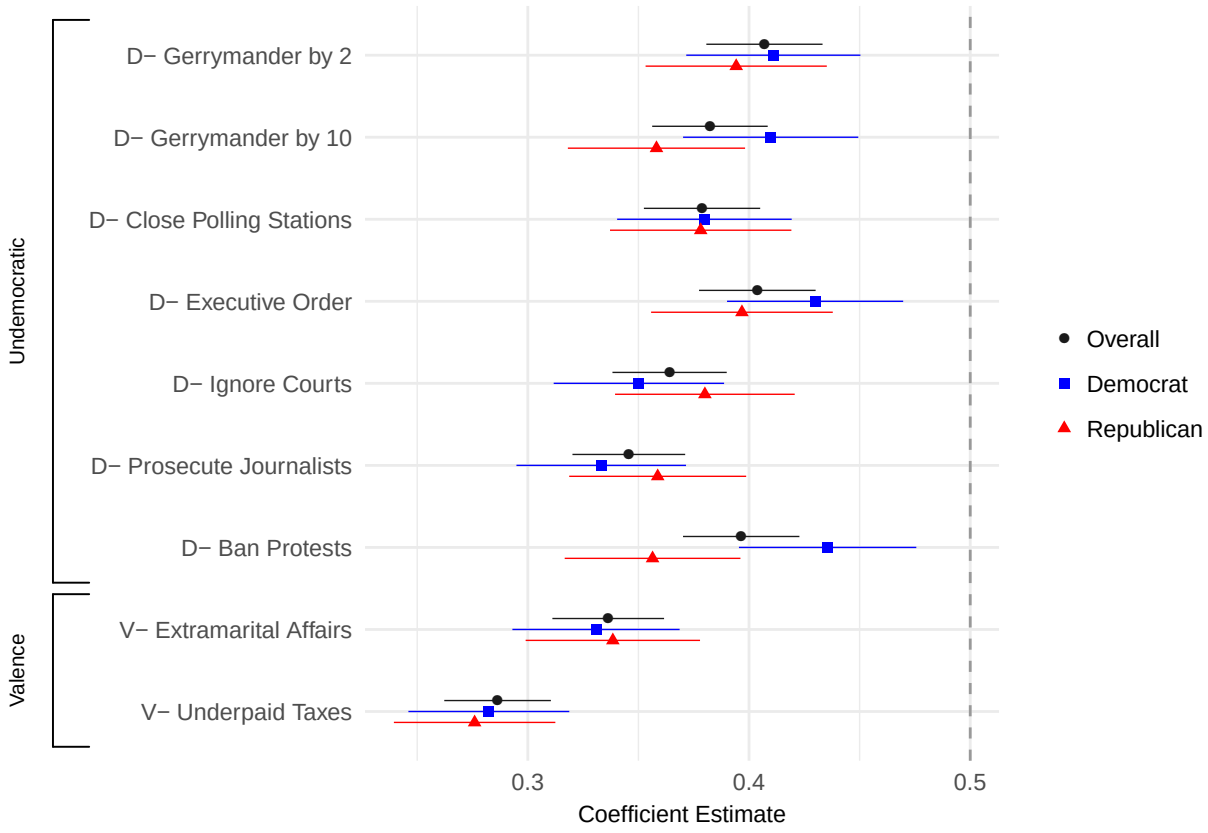

Figure 9: Effects of neutral, valence, and undemocratic positions on a candidate's vote share by respondents' partisanship. Dots represent coefficient estimates based on a linear regression model; bars represent 95% confidence intervals

4 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 that conceives of such positions as a "negative" valence and suggests that the civic virtue parameter δ in our model is positive. We now shift from the non-parametric approach employed so far to a structural approach that 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 the civic virtue parameter 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.

The estimation of these quantities 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 from section 2 that our assumption of standard Gumbel 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 the logistic regression.[30] For instance, our discussion in sections 3.1 and 3.2 of voters' willingness to trade off democratic principles for their preferred policies and party suggests the estimation equation

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

where M, *econ*, *social*, and *party* are differences between candidate 1 and 2's democracy

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

Table 2: Structural estimates of our model of the public as a democratic check

	Standardized		Natural	
	(1)	(2)	(3)	(4)
δ Democracy	0.183***	0.143***	0.157***	0.126***
	(0.168, 0.197)	(0.120, 0.158)	(0.143, 0.170)	(0.107, 0.139)
α_1 Economic policy	0.271***	0.213***	0.310***	0.249***
	(0.252, 0.290)	(0.179, 0.234)	(0.289, 0.332)	(0.213, 0.274)
α_2 Social policy	0.331***	0.261***	0.347***	0.279**
	(0.313, 0.350)	(0.220, 0.287)	(0.326, 0.369)	(0.238, 0.308)
α_3 Party	0.216***	0.171***	0.186***	0.150***
	(0.197, 0.235)	(0.142, 0.190)	(0.168, 0.204)	(0.127, 0.167)
N	17,341	17,341	17,341	17,341
Log-likelihood	-10,331	-10,315	-10,685	-10,666
Note:				*p<0.05; **p<0.01; ***p<0.001
				95% confidence intervals in parentheses

positions, economic and social policy platforms, and partisan affiliation. Estimates from (6) allow us to express the weights δ and α_k that voters place on democracy versus other candidate attributes k as

$$\hat{\delta} = \frac{|\beta_1|}{|\beta_1| + \sum_K |\beta_k|} \quad \text{and} \quad \hat{\alpha_k} = \frac{|\beta_k|}{|\beta_1| + \sum_K |\beta_k|}\,.$$

Table 2 presents estimates of the weights δ and α_k based on four alternative specifications. Column (1) estimates the logit model in (6) after the four input variables have been divided by two times 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% compared to 33% for social policy, 27% for economic policy, and 22% for partisanship.

Column (2) in Table 2 extends the model in (6) to include all experimentally manipulated candidate attributes.[31] The presented weights for candidates' democracy positions, economic and social policy platforms, and party imply that these four factors jointly account for 79% of the systematic variation in voters' candidate choices. These results are consistent with our earlier, non-parametric 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 the electorate's support for democracy in terms of voters' *value for democracy*. This concept expresses support for democracy in terms of other desirable candidate attributes that voters are willing to forgo in order to punish candidates who violate democratic principles. It is directly related to the marginal rate of substitution between democracy and policy that we approximated in section 3.1 by contrasting the control and treatment scenarios in Figure 2 and measuring at various points along the horizontal axis how much more attractive would candidate 1's policies have to become to compensate for his switching from D^+ to D^-. Voters' value for democracy summarizes the marginal rate of substitution between democracy and another candidate attribute parsimoniously – in terms of a single quantity. 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 differently, the "price" voters are willing to pay to elect a more democratic candidate is $\frac{\delta}{\alpha_k}$ of attribute k.

Estimates of value for democracy are most easily interpreted when we re-estimate 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

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

social policies on a 0-1 proximity scale. Columns (3) and (4) display the resulting estimates of the weights δ and α_k, parallelling the simple and extended models in (1) and (2). Estimates in (3) yield voters' value for democracy of 0.505 (CI: 0.443, 0.565) in economic policy, 0.451 (CI: 0.402, 0.507) in social policy, and 0.841 (CI: 0.734, 0.958) in co-partisanship. These quantities imply that in order to punish a candidate who adopts an undemocratic position, the voter is willing to tolerate a .51 and .45 increase in the distance from her ideal economic and social policies, respectively – about half the scale. A 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.

This finding about the value for democracy in terms of co-partisanship is consistent with our earlier, non-parametric analysis of partisanship in section 3.1. There, 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 the degree of this tendency in the electorate in terms of a single, theoretically-informed quantity: the *partisan bias* π. Specifically, we extend the simple framework in section 2 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; and 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

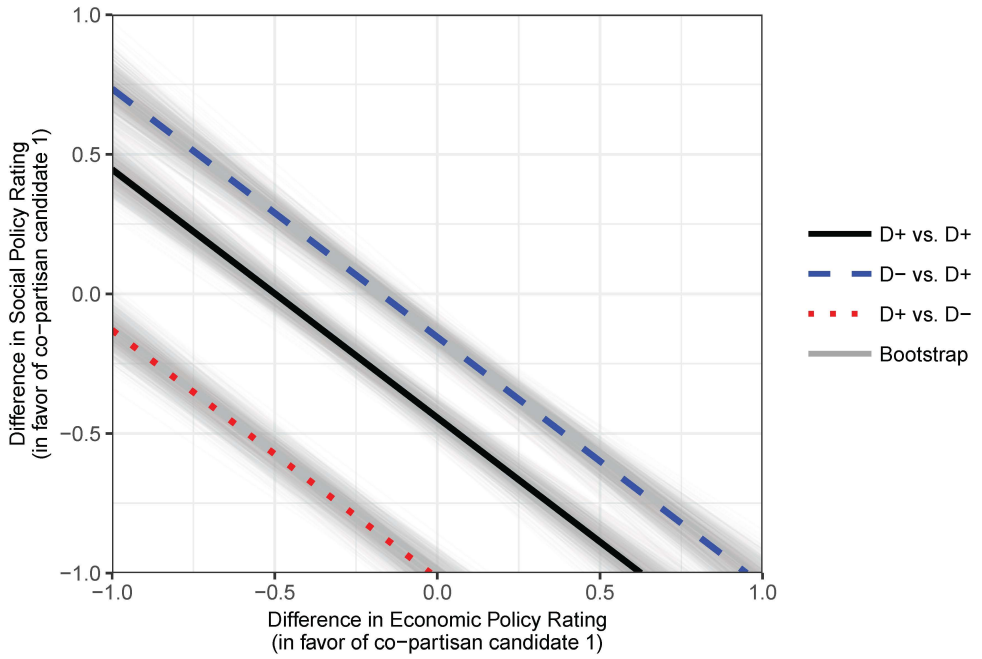

Figure 10: Isoelects depicting the combinations of economic and social policies that result in an equal probability of victory for both candidates. Candidate 1 is the respondent's co-partisan, candidate 2 is not; gray bands reflect statistical uncertainty

estimate of π is 0.495 (CI: 0.370, 0.613), implying that Americans are neither fully principled nor purely instrumental in their support for democracy. Rather, the average voter is about 50% more lenient toward violations of democratic principles by candidates from her own party.

The relationship between the structural parameters estimated in this section is graphically summarized by Figure 10. It plots the combinations of economic and social policies that result in an equal probability of victory for both candidates, assuming candidate 1 is the respondent's co-partisan but candidate 2 is not. We refer to these lines as *isoelects*. The solid black isoelect plots combinations of economic and social policies that result in an equal probability of victory for either candidate when both adopt a neutral democracy position (D^+ vs. D^+); the dashed blue and dotted red isoelects correspond to the D^- vs. D^+ and D^+ vs. D^- conditions, respectively. Combinations of economic and

social policies to the right and above the isoelects correspond to scenarios when candidate 1 is more likely to win; policy combinations to the left and below the isoelects correspond to scenarios when candidate 2 is more likely to win.

The slope of all isoelects is negative ($-\frac{\alpha_1}{\alpha_2}$) with an absolute value smaller than one ($|\frac{\alpha_1}{\alpha_2}| < 1$). This implies that voters value a candidate's greater proximity on both economic and social policies but place more weight on the latter.[32] Meanwhile, voters' value for democracy in terms of economic and social policies corresponds the distance between the D^+ vs. D^+ and the D^+ vs. D^- isoelect along the horizontal and vertical axis, respectively ($\frac{\delta}{\alpha_1}$ and $\frac{\delta}{\alpha_2}$).

The isoelects in Figure 10 summarize the impact of partisanship in two ways. First, note that the D^+ vs. D^+ isoelect is below the point $(0, 0)$, reflecting the advantage ($\frac{\alpha_3}{\alpha_2}$) conferred on candidate 1 by his co-partisanship with the respondent. Second, the 50% bias in the punishment of undemocratic positions that co-partisans benefit from is mirrored in the smaller distance between the D^+ vs. D^+ and the D^- vs. D^+ isoelects (compared to the distance between the D^+ vs. D^+ and the D^+ vs. D^- isoelects.)[33]

[32] A value of 0 on the horizontal and vertical axes refers to scenarios when the respondent rates the two candidates' economic or social platforms equally; positive values correspond to a higher rating of candidate 1's platform.

[33] This distance is $\frac{\delta}{\alpha_1}(1-\pi)$ and $\frac{\delta}{\alpha_2}(1-\pi)$ along the horizontal and vertical axis, respectively.

5 The 2017 Montana Natural Experiment

Depending on whether they voted on election day or by absentee ballot, voters in the 2017 special election for Montana's single U.S. House seat saw two different races.[34] Both races pitted the Republican Greg Gianforte against the 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 newspapers in Montana initially endorsed the 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 he repeatedly questioned the candidate about his position on Obamacare repeal.[35] The attack dominated the news coverage that evening and lead the three major newspapers in Montana to rescind their endorsement of Gianforte on the morning of election day. Gianforte nonetheless won by a 5.6% vote margin.

The timing of Gianforte's assault offers a real-world, quasi-experimental opportunity to evaluate the theoretical framework in section 2 and to corroborate the experimental results in section 3. We adopt a difference-in-differences empirical strategy that compares the (relative) shift among absentee and election day voters for the Republican U.S. House candidate between the November 2016 general election and the May 2017 special election.[36] That is, we think of vote shifts among absentee voters as a control that reflects what would have happened if no voters observed the assault;[37] vote shifts among election day voters

[34]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.

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

[36]On difference-in-differences estimation, see Angrist and Pischke (2009) and Bertrand et al. (2004).

[37]By the time of Gianforte's assault, election officials had already received 92.5 percent of absentee ballots that would ultimately be counted. According to the Montana voter file, 276,854 absentee ballots were

reflect the causal effect of the assault.[38] Even though absentee voters may be different from election day voters, those differences cancel out when we use each group as its own, pre-treatment benchmark.

In the context of our theory, 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.[39] In turn, we expect voters to punish Gianforte for the attack, but crucially, we predict that the severity of that punishment will be decreasing in the strength of a precinct'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 the strength of the Republican Party in a precinct grows. This obtains because voters' willingness to tolerate a co-partisan who violates democratic principles increases in the intensity of their partisanship.[40]

In order to investigate these predictions, we contacted election administrators in all of Montana's 56 counties and identified the five counties that tallied absentee and election day ballots separately for each candidate in the years 2014, 2016, and 2017.[41] Figure 11

processed and accepted in the 2017 special election on May 25, 2017. Of these, 256,156 have a received date of May 24, 2017 or earlier. For our sample of precincts, these figures are 33,191 of 35,264 (94.1 percent).

[38]As the *Billings Gazette* wrote, "To the voters who have not voted yet, we simply urge you to evaluate each candidate very carefully and make the best choice. To those who have voted: Unfortunately, Montana does not allow those who voted early to reconsider and vote again. We're one of the few states that does not. This would seem to be the best reason we should urge our state leaders to change that law." The Billings Gazette Editorial Board, May 24, 2017.

[39]This is, in fact, how one major newspaper, *The Billings Gazette*, interpreted the assault in its election day editorial: "First, we hope that Republican party members and leaders call this for what it appears to be, an inexcusable act. We hope that partisan politics has not eroded our decency to the point where leaders and supporters feel the need to defend the indefensible... This incident is not Montana. It's not America. It's not who we are, and attacking literally – those with whom we disagree cannot be justified, tolerated or explained away. We must adopt zero tolerance for such behavior if freedom of expression means anything." The Billings Gazette Editorial Board, May 24, 2017.

[40]We provide empirical support for this claim in the appendix, where we merge data from the 2016 Cooperative Congressional Election Study with county-level results from the 2016 presidential election to show that Republicans in more Republican counties indeed tend to be stronger partisans, more conservative on the liberal-conservative scale, more conservative on issue position-based measures of ideology, and more disapproving of former President Barack Obama.

[41]Whether a county tallied absentee and election day ballots separately for each candidate appears to be

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 a measure of the intensity of a precinct's partisanship.[42] Absentee vote shifts are shown as circles, election day vote shifts as diamonds; positive differences in vote shifts are highlighted by upward-facing red arrows, negative differences by downward-facing blue 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.[43]

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

$$R_{it} = \alpha + \beta_1 Y17_{it} + \beta_2 E_{it} + \beta_3 Y17_{it} E_{it} + \gamma \mathbf{X}_{it} + \epsilon_{it}\,, \tag{7}$$

$$\begin{aligned} R_{it} = \alpha &+ \beta_1 Y17_{it} + \beta_2 E_{it} + \beta_3 Y17_{it} E_{it} + \beta_4 R_{16} + \beta_5 Y17_{it} R_{i16} \\ &+ \beta_6 E_{it} R_{i16} + \beta_7 Y17_{it} E_{it} R_{i16} + \gamma \mathbf{X}_{it} + \epsilon_{it}\,. \end{aligned} \tag{8}$$

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 percentage of absentee voters, percentage living within the city limits, and mean age.[44]

primarily a function of the type of voting machine that it uses. In the appendix, we use results from 2014 to examine the parallel trends assumption and construct placebo tests.

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

[43] That is, some hardline Republican precincts appear to have rewarded rather than punish Gianforte's assault of the journalist.

[44] 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 secular trend toward

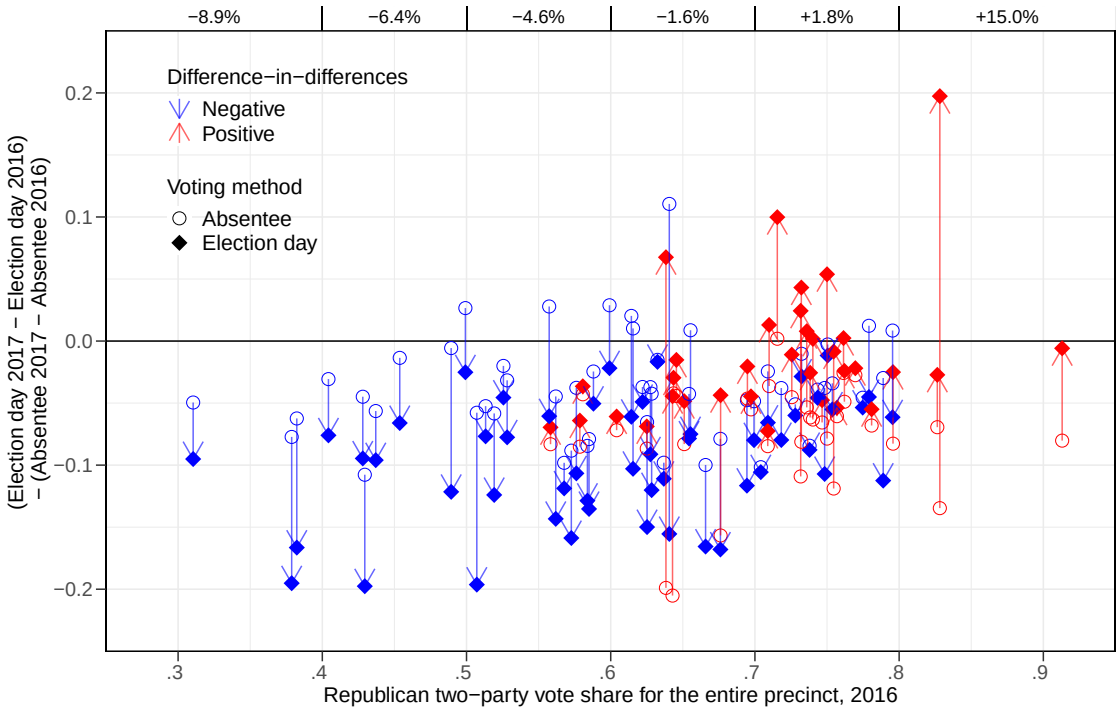

Figure 11: Differences in precinct-level vote shifts for the Republican U.S. House candidate in Montana between November 2016 and May 2017

Table 3 displays regression estimates for models in equations (7) and (8). Our main coefficients of interests 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 -.036 of the two-party vote share (column 1.) Thus, overall, Gianforte was punished by the loss of 3.6% election day voters. Montanans value a free press.

In equation (8), our main interest is in coefficient β_7, which captures how the assault's effect varies with the 2016 Republican vote share in the entire precinct.[45] A positive β_7 implies that Gianforte's 2017 election day vote share is increasing in a precinct's 2016 Republican vote share. This is indeed what we observe (column 2): the more Republican a

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 2000-Present," Montana Secretary of State, accessed on November 16, 2018.

[45]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. $R_{i16} = 0$).

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	-0.036^{*}	-0.237^{***}	-0.042^{*}	-0.174^{**}
$\times$ Election day	(0.014)	(0.041)	(0.021)	(0.055)
β_7 2017 $\times$ $\%R_{i16}$		0.313^{***}		0.220^{**}
$\times$ Election day		(0.066)		(0.085)
N	348	348	164	164
Adjusted R^2	0.315	0.904	0.432	0.923
Note:				*p<0.05; **p<0.01; ***p<0.001
				Standard errors clustered by precinct

precinct was in 2016, the more forgiving election day voters are 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 trends assumption.[46]

These findings are consistent with our theoretical framework and experimental findings: only moderate Republicans are willing to punish Gianforte for assaulting the journalist by either abstaining or voting for a Democrat; for strong partisans, partisan loyalty trumps valence considerations. Montanans value a free press, but not enough for most hardline Republicans to vote for a Democrat.

[46]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 42 out 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 the difference-in-differences framework.

6 Conclusion: It Can't Happen Here?

In this paper, we have addressed a fundamental question about democratic stability in the United States: When can we realistically expect the American electorate to serve as a check on undemocratic behavior by elected politicians? We consistently found that only a small fraction of our respondents prioritized democratic principles in their electoral choices when doing so went against their partisan identification, political ideology, or favorite policies. 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 they punish candidates who violate democratic principles. Findings based on our revealed preference approach thus indicate that Americans' commitment to democratic principles may be significantly weaker than suggested by conventional measures of support for democracy.

We conclude this paper by discussing the implications of our findings for democratic stability in the United States. Throughout sections 3-4, we based our analysis on a candidate-choice experiment in which all candidate attributes were independently assigned. A key advantage of this design is that it allows us to identify each attribute's causal effect (Hainmueller et al., 2015). One potential downside is limited external validity: Some of our candidates featured combinations of policies and partisanship rarely seen in the real world.

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

Table 4: The U.S. electorate's resilience to undemocratic candidates: A sensitivity analysis

	Condition	**% Defecting from D^-**	
	Overall	11.31^{***}	(9.79,12.84)
1.	Across-party contests	10.01^{***}	(7.92,12.09)
2.	Platform divergence	6.71^{***}	(3.94,9.49)
3.	Moderate party-policy alignment	5.50^{**}	(0.00,11.07)

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 policies. Condition 3 also precludes Republican candidates from adopting the left-most position on any issue and vice-versa for Democrats.

We see that as we gradually restrict attention to candidate choice scenarios with combinations of partisanship and policies that approximate real-world elections, the punishment for candidates who violate democratic principles declines from 11.31% to 5.50%.[47] This is consistent with our theoretical framework. The progressive application of the three conditions in Table 4 effectively explores the consequences of a conception of polarization that we have anticipated theoretically but have not addressed empirically so far: a positive correlation between policies and partisanship.[48] The three conditions induce an increasing alignment between each candidate's policies and partisanship, thus compounding the differences that respondents see between Democrats and Republicans. 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.

[47] Our analysis throughout sections 3-4 thus provides conservative estimates of the likely punishment for candidates who violate democratic principles.

[48] At the level of the electorate, this conception of polarization is closest to the "sorting" mechanism examined by Fiorina et al. (2008).

References

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

Ahlquist, John S., Kenneth R. Mayer and Simon Jackman. 2014. "Alien Abduction and Voter Impersonation in the 2012 U.S. General Election: Evidence from a Survey List Experiment." *Election Law Journal: Rules, Politics, and Policy* 13(4).

Albertson, Bethany and Akin Guiler. 2018. "Conspiracies, Electoral Fraud, and Support for Democratic Norms." Unpublished manuscript, University of Texas.

Alexander, Amy C. and Christian Welzel. 2017. "The myth of deconsolidation: Rising liberalism and the populist reaction." *Journal of Democracy* Web Exchange.

Almond, Gabriel A. and Sidney Verba. 1963. *The civic culture: Political attitudes and democracy in five nations*. Newbury Park, CA: Sage.

Angrist, Joshua David and Jorn-Steffen Pischke. 2009. *Mostly harmless econometrics: An empiricist's companion*. Princeton: Princeton University Press.

Ansolabehere, Stephen and Nathaniel Persily. 2008. "Revisiting public opinion on voter identification and voter fraud in an era of increasing partisan polarization." *Stanford Law Review* 121(7):1737–1774.

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. Cambridge University Press pp. 15–58.

Barber, Michael and Jeremy C. Pope. Forthcoming. "Does party trump ideology: Disentangling party and ideology in America." *American Political Science Review* .

Bertrand, Marianne, Esther Duflo and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-In-Differences Estimates?" *Quarterly Journal of Economics* 119(1):249–275.

Boix, Carles, Michael Miller and Sebastian Rosato. 2013. "A Complete Data Set of Political Regimes, 1800-2007." *Comparative Political Studies* 46(2):1523–1554.

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

Carey, John M., Gretchen Helmke, Brendan Nyhan and Susan C. Stokes. 2018. *Bright Line Watch*.

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." *Bright Line Watch* .

Cheibub, José A., Jennifer Gandhi and James R. Vreeland. 2010. "Democracy and Dictatorship Revisited." *Public Choice* 143(1):67–101.

Chen, Jowei and Jonathan Rodden. 2013. "Unintentional Gerrymandering: Political Geography and Electoral Bias in Legislatures." *Quarterly Journal of Political Science* 8(3):239–269.

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):351–366.

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.

Coppedge, Michael, John Gerring, David Altman, Michael Bernhard, Steven Fish, Allen Hicken, Matthew Kroenig, Staffan I. Lindberg, Kelly McMann, Pamela Paxton, Holli A. Semetko, Svend-Erik Skaaning, Jeffrey Staton and Jan Teorell. 2011. "Conceptualizing and Measuring Democracy: A New Approach." *Perspectives on Politics* 9(2):247–267.

Dahl, Robert A. 1956. *A Preface to Democratic Theory.* Chicago, IL: University of Chicago 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):441–472.

Fiorina, Morris P., Samuel A. Abrams and Jeremy C. Pope. 2008. "Polarization in the American Public: Misconceptions and Misreadings." *The Journal of Politics* 70(02):556–560.

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. 2018. "Committed or Conditional Democrats? Opposition Dynamics in Electoral Autocracies." Unpublished Manuscript, Emory University.

Gelman, Andrew. 2008. "Scaling regression inputs by dividing by two standard deviations." *Statistics in Medicine* 27(15):2865–2873.

Grimmer, Justin, Eitan Hersh, Marc Meredith, Jonathan Mummolo and Clayton Nal. 2018. "Obstacles to Estimating Voter ID Laws Effect on Turnout." *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.

Hajnal, Zoltan, Nazita Lajevardi and Lindsay Nielson. 2017. "Voter Identification Laws and the Suppression of Minority Votes." *Journal of Politics* 79(2):363–79.

Huq, Aziz Z. and Tom Ginsburg. 2017. "How to Lose a Constitutional Democracy." Unpublished manuscript, University of Chicago.

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

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* pp. 1–35.

Kaufman, Robert R. and Stephan Haggard. Forthcoming. "Democratic Decline in the United States: What Can We Learn from Middle-Income Backsliding?" *Perspectives on Politics* .

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. Forthcoming. "The Trump Presidency and American Democracy: A Historical and Comparative Analysis." *Perspectives on Politics* .

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 Univerisity.

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

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.

Norris, Pippa. 2017. "Is western democracy backsliding? Diagnosing the risks." *Journal of Democracy* Web Exchange.

Norris, Pippa, Sarah Cameron and Thomas Wynter, eds. 2018. *Electoral integrity in America: Securing democracy*. New York, NY: Oxford 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. 2017. "What's Happening?" Unpublished manuscript, New York University.

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

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.

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):245–263.