



Institutional Members: CEPR, NBER and Università Bocconi

WORKING PAPER SERIES

The Shift to Commitment Politics and Populism: Theory and Evidence

Luca Bellodi, Massimo Morelli, Antonio Nicolò, Paolo Roberti

Working Paper n. 714

This Version: May 3, 2025

IGIER – Università Bocconi, Via Guglielmo Röntgen 1, 20136 Milano –Italy
<http://www.igier.unibocconi.it>

The opinions expressed in the working papers are those of the authors alone, and not those of the Institute, which takes non institutional policy position, nor those of CEPR, NBER or Università Bocconi.

THE SHIFT TO COMMITMENT POLITICS AND POPULISM: THEORY AND EVIDENCE*

Luca Bellodi[†] Massimo Morelli[‡] Antonio Nicolò[§] Paolo Roberti[¶]

May 3, 2025

Abstract

This paper presents a theory of populism centered on commitment politics – a type of agency relationship in which political candidates promise specific and monitorable policies. The shift to commitment politics is driven by increased distrust toward government institutions, which is itself a consequence of the cognitive complexity and disinformation typical of modern social media environments. In equilibrium, candidates who adopt a commitment platform rationally choose all the complementary strategies associated with populism, including anti-elite rhetoric, misinformation, aversion to judicial independence, and bureaucratic expertise. The paper presents observational and experimental evidence from the United States on the supply (candidates’ rhetoric) and demand of commitment (voters’ preferences) consistent with the model’s key predictions.

Keywords: Commitment Politics, Trust, Populism, U.S. Elections

JEL codes: D72, D78, P16

*We wish to thank Alberto Bisin, Jack Blumenau, Luca Braghieri, Peter Buisseret, Steve Callender, Maria Carreri, Carlo Cusimano, Georgy Egorov, Ben Enke, Gabriele Gratton, Bard Harstad, Matt Jackson, Jenny Kim, Greg Martin, Pietro Ortoleva, Ken Shepsle, Ken Shotts, Carlo Schwartz, Konstantin Sonin, Richard Van Weelden, and Stephane Wolton for helpful comments and interesting discussions on earlier drafts. We thank Ginevra Aguiari, Francesco Migliore, Marvin Pappalettera, and Domenico Tripodi for excellent research assistance. Comments by seminar participants at Stanford GSB, University of Chicago Harris School, Northwestern MEDS, NYU, Bocconi, EUI, Luxemburg, Zurich, Leibniz Center, CESifo, IE-Rochester workshop, NERI Political Economy of Reforms Workshop, CEPR, APSA, EEA, and Econometric Society conferences are also acknowledged. The paper has been presented as the keynote address at the Central European University in Vienna, at the University of Auckland, at the Journées Louis-André Gérard-Varet in Marseille, and at the UniBZ Workshop on Political Economy in Bolzano. IGIER and Baffi research centers have provided useful support. The usual disclaimer applies.

[†]Hoover Institution and Department of Political Science, Stanford University

[‡]Bocconi University, LISER and CEPR. Email: massimo.morelli@unibocconi.it

[§]University of Padua and University of Manchester

[¶]Free University of Bozen-Bolzano

1 Introduction

Over the past decades, many democracies around the world have experienced a steady decline in citizens’ trust in representative institutions (Valgarðsson et al. 2025). In the United States, the share of individuals who trust the federal government dropped from 73% in 1958 to 22% in 2024 (Pew Research Center 2024). Among the sources of erosion of trust, scholars have focused mostly on economic and cultural factors: rising unemployment (Algan et al. 2017), globalization (Rodrik 2018*a*, Colantone & Stanig 2018), financial crises (Foster & Frieden 2017, Guiso et al. 2021), and immigration (Norris & Inglehart 2019) have all heightened economic and cultural insecurity, with direct negative consequences for citizens’ trust in politics. Moreover, the rise of the internet and the emergence of social media have contributed to declining confidence in government (Guriev, Sergei et al. 2021), facilitated the spread of misinformation (Allcott & Gentzkow 2017, Vosoughi et al. 2018, Denter & Ginzburg 2024), and reinforced in-group beliefs and simplistic narratives (Levy & Razin 2019, Aridor et al. 2024). Ultimately, these developments have contributed to an environment in which verifying the accuracy of information has become increasingly more difficult (Sunstein 2017, Zhuravskaya et al. 2020).

The most consequential implication of this crisis of trust has been the emergence and rise of populism, a multifaceted concept that frames politics as a fundamental struggle between a virtuous people and a corrupt elite (Mudde 2004). In electoral campaigns, populism is closely tied to anti-elite rhetoric, simplistic narratives, and the spread of misinformation (Hawkins et al. 2019, Guriev & Papaioannou 2022, Gennaro et al. 2024). When in office, the most common characteristics of populist governments are exclusionary policies, particularly anti-immigration and protectionist trade policies, as well as attempts to undermine democratic checks and balances, attacking the expertise of bureaucracies, the independence of courts, and traditional media (Funke et al. 2023, Gratton & Lee 2023, Bellodi et al. 2024).

This paper aims to provide a unifying theoretical perspective that connects the various elements of the populist phenomenon with this landscape of increasing distrust in a coherent model of political agency. Our theory shows how lower trust determines a shift away from the traditional “trustee” model of representative democracies – where politicians are entrusted to adjust policymaking to changing circumstances – towards what we call “commitment politics,” a situation in which voters demand commitments, often in the form of protection policies, as a response to the most pressing economic or cultural threats and to the altered information environment. These commitments are perceived by voters as directly actionable and verifiable, hence remedying the lack of trust in political representatives. Our primary contribution is to demonstrate theoretically and

empirically that such a shift to commitment politics provides a unifying rationalization of the main features of populism.

Illustrative examples of this shift toward commitment politics are observable across a wide range of political contexts. Over time, various politicians have proposed “contracts” with voters, signaling an explicit commitment to implement a predetermined set of policies.¹ Notably, the unprecedented volume of executive orders issued by President Trump during the first 100 days of his second term – three times larger than the number issued by former President Biden over the same period – highlights the commitment-oriented nature of policy delivery within the populist paradigm. Recent developments in Brazil, Argentina, Mexico, Hungary, and several other countries can also be interpreted through the lens of commitment politics. The widespread use of anti-immigration and protectionist pledges among populist candidates and parties in Western democracies reinforces the view that populism often entails a transformation of the political agency relationship – from a trustee model to one centered on a limited set of monitorable commitments.

We focus on how increasing informational complexity and fragmentation undermine voters’ ability to hold politicians accountable, thereby fueling distrust and heightening the demand for simple, monitorable policy commitments. In settings where politicians maintain discretion over policy decisions, trust rests on the belief that politicians will serve the interests of voters rather than being influenced or captured by elites. Accountability requires voters to infer the state of the world and assess the welfare implications of policy choices. This task is made more difficult by the technical complexity of the policy issues and the proliferation of conflicting narratives. Climate policy, AI regulation, and geopolitical strategies are domains that illustrate how difficult it can be for the electorate to evaluate policy ex-post. Similarly, the diffusion of the internet and the advent of social media have revolutionized the informational landscape and facilitated the spread of disinformation, thereby amplifying voter disorientation. In contrast, when candidates campaign on explicit policy commitments – with no conditioning on realized states of the world – voters can evaluate politicians’ performance more directly, reducing reliance on complex information environments. This shift is especially pronounced in policy areas where voters perceive high complexity or where there is a wide divergence between their preferences and those of powerful interest groups. In such environments, we show that it becomes strategically optimal for candidates to adopt observable commitments and to employ anti-elite rhetoric, as a way to discredit opponents who retain discretionary

¹The “Contract with the Italians,” launched by former Prime Minister Silvio Berlusconi in 2001—possibly inspired by Newt Gingrich’s “Contract with America” – exemplifies an early manifestation of commitment politics in Italy. The Five Star Movement further advanced this approach by promoting direct citizen engagement, enforcing mandate constraints, and imposing term limits to establish distance from traditional elites.

authority and to demobilize their electorate. Therefore, distrust increases turnout among supporters of commitment candidates and depresses turnout among supporters of trustee-type candidates. This dynamic provides electoral incentives for committed candidates to amplify distrust, including by reinforcing voters’ prior beliefs. In equilibrium, this creates a feedback loop: growing distrust increases support for commitment politics, and the rise of commitment-oriented candidates further entrenches that distrust.

Our framework also offers insights into populist behavior in office. Elected officials bound by policy commitments have incentives to reduce institutional checks and balances – including expert bureaucracies, the independence of the judiciary, and free media – that could obstruct the implementation of the populist commitments. Voters who support such commitments may similarly endorse efforts to weaken such constraints on executive power. Indeed, common populist pledges often face legal and regulatory barriers, including those imposed by international bodies. Consequently, a tendency to erode judicial independence and bureaucratic capacity frequently accompanies populist governance (for similar conclusions, see e.g., Sasso & Morelli 2021, Gratton & Lee 2023).

In the second part of the paper, we complement our theoretical account with novel evidence on the supply and demand for commitment. Although there is compelling evidence that changes in communication technology amplified the spread of misinformation and made it harder for voters to process and acquire information, the focus of our empirical analysis is on the relationship between the decline in trust (whatever its source) and the supply and demand for policy commitments.

Our empirical analysis focuses on U.S. congressional elections, which enable us to match candidates with their respective constituencies, thereby ensuring a precise mapping between the district electorate and candidates’ rhetoric. We provide evidence on both the supply and demand of commitment.

On the supply side, we analyze the social media campaign of candidates to the U.S. House of Representatives during the election years 2012, 2016, and 2020. First, we use a statistical model to create dynamic district-level estimates of voters’ distrust in the federal government. Second, we match the distrust estimates at the district level with the universe of tweets posted by more than 2,000 political candidates. To detect when a tweet contains a policy commitment or populist rhetoric, we use natural language processing and supervised machine learning techniques. We leverage within-candidate and within-district variation in distrust resulting from candidates running multiple times in the same district to identify the effect of distrust on candidates’ rhetoric. We find that voters’ distrust is strongly associated with candidates’ supply of commitments and populist rhetoric, and tweets about topics on which partisan voters display high levels of consensus and conviction about the right policy are more likely to take the form of policy

commitments.

We also test whether the supply of commitment has mobilization effects. We use a panel survey embedded in the American National Election Studies (2016-20), which asks the same group of respondents about their turnout intention during two election years. We exploit the differential timing of the interviews and create a measure of exposure to commitment for each respondent based on the candidates' actual supply of commitments in the same district as the respondents. Holding fixed individual respondents' and candidates' fixed characteristics, we find that an increase in distrust is associated with a higher probability of reporting intention to turn out to vote in support of one's own party's candidate when that candidate shows more commitment. The opposite holds true when the candidate runs a delegation campaign, lending empirical support to the prediction that distrust has divergent mobilization effects on the two groups of citizens.

On the demand side, we demonstrate that voters' preferences for commitment politics increase when they experience a negative, experimentally manipulated shock to their level of political trust. Through a survey experiment, we verify whether voters prompted with an informational treatment that induces distrust in elected representatives are more likely to believe that politicians should make and maintain clear commitments. We do find evidence of such a demand effect, with treated respondents more likely to agree with statements emphasizing the importance of commitment politics. Moreover, we provide additional survey data that compares voters' preferences for key democratic institutions before and after the 2020 presidential elections, which saw Donald Trump running against Joe Biden. We find that Trump supporters in the pre-election period (under Trump) are more likely to prefer weaker checks and balances compared to Biden supporters in the post-election period (under Biden). By approximating Trump with a commitment agent, this suggestive evidence is consistent with the demand for weaker agencies of restraints on behalf of voters who prefer commitment politics.²

The remainder of the paper is organized as follows. In the next section, we highlight the contributions we make to various literatures. Section 3 presents our theory, and Section 4 contains the empirical analysis. Section 5 concludes. The proofs of formal results and multiple robustness checks are reported in Section A in the Appendix.

2 Contributions to the literature

This paper contributes to several strands of literature in economics and political science.

First, empirically, our findings speak to a large body of work examining the economic

²The term “agencies of restraint” was introduced in Rodrik (2018*b*) to refer simultaneously to traditional media, independent judges and courts, expert bureaucrats.

(Algan et al. 2017, Guiso et al. 2024), technological (Campante et al. 2017, Manacorda et al. 2022, Petrova et al. 2024), cultural or moral (Norris & Inglehart 2019, Enke 2020) factors that drive support for populist parties and candidates. Second, several theories of populism rely on cognitive differences among voters. Levy et al. (2022), for instance, depict populist policies as simplistic ones preferred by unsophisticated voters who sometimes win elections due to strong dissatisfaction with the status quo. Crutzen et al. (2024) show that when voters are split between an informed minority and an uninformed majority, parties tend to cater to the elite, fueling disaffection among ordinary citizens and opening space for populist challengers. Egorov & Sonin (2021) present a model where an informed elite advises the uninformed majority about candidate competence, and examine when the latter follows the elite’s recommendations. In contrast to this literature, our model does not assume cognitive differences across groups. Instead, it emphasizes how cognitive complexity and disinformation endogenously generate distrust, thereby driving the shift toward commitment politics.

A parallel body of literature argues that the success of (radical right) populist parties stems from a shift in the main dimension of political conflict – from economic to cultural issues – which has reshaped patterns of social identification. (see e.g., Shayo 2009, Ford & Jennings 2020, Besley & Persson 2019, Bonomi et al. 2021, Gennaioli & Tabellini 2023). Consistent with this literature, when it becomes rational for parties to adopt a commitment strategy, such commitments may center on protecting national identity and communal values. Beyond the social-psychological explanation, cultural issues may gain salience because reforms in this domain are easier to monitor.

Our focus on complex information environments as exogenous sources of distrust and a key determinant of commitment politics is consistent with a large empirical literature on the political consequences of social media, especially as it connects to the surge of populism (Guriev & Papaioannou 2022). The proposed mechanisms connecting social media to populism are often centered on lower barriers to entry (Zhuravskaya et al. 2020) and direct contact with voters (Campante et al. 2017). We contribute to understanding the phenomenon by highlighting an alternative mechanism, namely the role of technological change in eroding trust in political agents and increasing demand for policy commitments that require no ex-post information. Our model also helps explain existing evidence on social media’s role in belief polarization (see e.g., Bakshy et al. 2015, Halberstam & Knight 2016). In this framework, committed candidates are incentivized to create narratives that reinforce the prior beliefs of voters who granted them a mandate based on specific commitments.

In our model, trust is the primary driver behind the shift to commitment politics. As such, our analysis engages with the broader literature on the economic and cultural roots

of political trust (for comprehensive reviews, see Berman 2021, Guriev & Papaioannou 2022). Specifically, we present an additional “endogenous” channel to the declining trust in traditional forms of delegation, which is rooted in the decline in accountability due to the increased costs and complexity in acquiring and processing fragmented information, often hard to verify. Di Tella & Rotemberg (2018) also link distrust to commitment, showing that betrayal aversion increases the demand for committed actions.

Our treatment of commitment politics builds on existing work in political agency and accountability. While voter-politician accountability is often modeled through a principal-agent framework (see e.g., Fearon 1999), the distinction between commitment and trustee models was introduced by Fox & Shotts (2009) in the context of optimal accountability. However, this paper is the first to use this distinction to explain populist behavior as a response to distrust and complexity. Our work also relates to Kartik et al. (2017), who model electoral competition with endogenous discretion. Like us, they include a post-election state realization, but unlike our model, they abstract from elite capture and its role in shaping distrust and demand for commitment. The trade-off between commitment and flexibility is discussed by Amador et al. (2006), who analyze optimal commitment devices when retaining some discretion is valuable. A similar distinction appears in Ghosh & Tripathi (2012) and Bueno de Mesquita & Friedenber (2011), though in their models, the committed agent is an “ideologue.” In contrast, our model does not tie commitment to ideology.

A key feature of policy commitments is their level of credibility (see e.g. Van Weelden 2013). We show that when information on the optimal policy is harder to obtain (higher complexity), then the wedge between the endogenous credibility of a committed candidate and that of a discretionary agent increases. If politicians have to design a multidimensional platform, they may choose to commit to issues where the distance from the elite or the representation gap between the preferences of voters and those of traditional parties is largest, and where the possibility of feedback to voters is difficult, in line with the theoretical insights in Buisseret & Van Weelden (2020) and empirical evidence in Günther (2024).

Policy commitments are themselves a form of pandering, for candidates who adopt a commitment strategy are more likely to commit on salient policy issues on which voters have similar preferences (Canes-Wrone et al. 2001). In fact, pandering to the ex-ante voter’s beliefs has been considered by itself an important feature of populism (Acemoglu, Egorov & Sonin 2013). However, we take several steps further and characterize the candidate’s complementary strategies to show that they are all consistent with the various facets of populism discussed in the literature. In particular, we show that a shift to commitment politics can generate demand for reduced checks and balances, hence ra-

tionalizing existing evidence on voters' willingness to trade off democratic principles for policy preferences (Graham & Svolik 2020, Alsan et al. 2023, Zaslove & Meijers 2023).

Finally, this paper aligns with empirical and theoretical work on the behavior of populist politicians while in office. The willingness to disregard economic constraints in implementing policy commitments is well established in the literature, which documents such a tendency in populist governments of Latin America (Kaufman & Stallings 1991, Dornbusch & Edwards 1991, Edwards 2019), and on a larger global sample of countries (Funke et al. 2023). Similarly, Bellodi et al. (2024) and Sasso & Morelli (2021) highlight detrimental implications of populist governments for the quality of bureaucracy and the retention of experts in government. Our paper sheds light on a clear reason why, when populists are in power, they replace competent bureaucrats with loyal bureaucrats to minimize the risk of resistance to the populist commitment agenda.

Populist governments frequently erode institutional constraints, including the bureaucracy, media, and judiciary. In Hungary, for example, checks and balances were curtailed under Orbán's rule.³ In Mexico, President Obrador passed a reform reducing the independence of the National Electoral Institute ahead of the 2024 elections.⁴ Poland's Law and Justice Party (PiS) similarly undermined judicial independence and media freedom following a campaign rich in policy commitments.⁵ The second Trump administration also showed efforts to weaken judicial independence and challenge the so-called deep state. The paper develops a theoretical perspective on this illiberal dynamic, which contributes to the literature on endogenous checks and balances and on the determinants of illiberal reforms. Acemoglu, Robinson & Torvik (2013) show how (economically disadvantaged) voters might prefer weaker checks on the executive when they expect the rich elite to be able to easily bribe the politician, as is the case in many weakly institutionalized polities. Voters are therefore willing to give up control of the executive as a guarantee that the politician will not be bought by the rich elite. Gratton & Morelli (2022) model checks and balances in a similar way to ours, for checks and balances regimes are treated as a veto player that can block executive decisions. Focusing on populist policies aimed at dismantling expert bureaucracies, Gratton & Lee (2023) show that demand for such illiberal reforms depends on the level of distrust voters have towards public servants.⁶

³The constitutional court can no longer block executive policy decisions. See <https://www.bbc.com/news/world-europe-21740743>.

⁴See <https://foreignpolicy.com/2023/03/23/lopez-obrador-electoral-reforms-mexico-democracy-ine/>.

⁵See <https://www.bbc.com/news/world-europe-65809525>.

⁶A potential demand for illiberal reforms is a pattern that has received sustained attention also among democratic theorists (for a recent review, see Urbinati 2019).

3 Theory

We want to derive the shift to commitment politics (demand and supply) from first principles, and then obtain the implications that relate to populist endogenous strategies.

We assume that the distribution of ideologies is fixed, in the sense that voters are either left-wing or right-wing, and this cannot change.⁷ Thus, we view parties and candidates as primarily focused on mobilization (affecting turnout incentives) as the main strategy to win office. The model will allow us, at the same time, to study the determinants of the demand for commitment within each party (and selection in primaries) and the strategic supply of policy proposals and rhetorical attempts to manipulate beliefs in general elections.

We begin by introducing the model, and then we present the results on the demand and supply of commitment, as well as the consequences for turnout, rhetoric, and populist attitudes.

3.1 Setup

There are two parties, L and R , corresponding to a division of the electorate into two ideological sides. Each party has mass equal one of voters. Within each party, citizens have the same preferences on policies and the same beliefs on the probability that elites can capture politicians, but they are heterogeneous in terms of cost of voting. Since preferences and beliefs are homogeneous within parties, we refer to the preferences and beliefs of the members of party $G = L, R$ simply as the preferences of party G .

Each party selects a candidate in a primary at time $t = 0$, and then the selected candidates $g = l, r$ compete in a general election at time $t = 1$. Then, at time $t = 2$ Nature chooses a state of the world that affects the value of different policies, and the elected policymaker decides what to do given their greater information with respect to voters and given the influence of elites. Before describing each of the three stages of the game, let us describe the policy-making problem and the different preferences and information of the relevant players.

There are n policy dimensions, $d = 1, \dots, n$, and on each dimension there is a binary choice between maintaining the status quo, s_d , and making a reform, r_d . Left and right voters may have different preferences on each dimension and, on each dimension, they lack information about whether the status quo or the reform is better for them.

⁷As already suggested by Campbell et al. (1960), ideologies tend to persist over people's life time (Krosnick & Alwin 1989) and are resistant to new information (Taber & Lodge 2006). However, "within" one's ideology, a significant economic shock can affect trust, willingness to participate in politics, and the decision to continue to support a traditional party or to embrace a new movement or candidate (see e.g., Golder 2016).

The ex-post optimal policy for party G voters depends on the state of the world at time $t = 2$, which is unknown before and is observed by the elected policymaker with probability 1.⁸ The potential candidates (agents) to be chosen in each primary election can either be given discretion on a policy or else can be asked by party voters (the principal) to commit to one, either the status quo or the reform. Given the asymmetric information about the state of the world at time 2, giving discretion to the agent is optimal if the principal trusts the agent to have aligned preferences and not to be possibly derailed by external influence. We will call “elite” the source of potential external influence on the agent. Giving discretion allows an informed agent to choose the ex-post optimal policy for the principal, but it exposes the principal to the risk that the agent is captured by the elite. Commitment to a specific policy ex-ante, on the contrary, reduces the probability that the elite distorts the agent’s choice, but it does not allow for adjusting the policy ex-post.

Principal G attaches probability γ_d^G that the best policy for her ex-post is the reform r_d on dimension d . Consequently, s_d is expected to be optimal with probability $1 - \gamma_d^G$. An elite with misaligned interests from the principal could influence the agent’s choice. This worry is captured by parameter $\lambda_d^G \in [0, 1]$, representing the probability that voters of party G assign to the event that a relevant elite prefers s_d over r_d . Thus, if we assume that the ex-post preferences of the principal and the elite are independent draws, then there is an ex-post conflict of interest with probability $\gamma_d^G \lambda_d^G + (1 - \gamma_d^G)(1 - \lambda_d^G)$. We also assume that the ex-post preferences of the two principals are independent draws. Thus, the two principals have ex-post different preferences on dimension d with probability $\gamma_d^G(1 - \gamma_d^{G'}) + \gamma_d^{G'}(1 - \gamma_d^G)$.

Parties have separable utility so that the utility of a n -dimensional policy is the sum of the utilities of the policy implemented in each dimension. The ex-post utility for principal G from dimension d is 0 if the implemented policy matches the state of the world, whereas she obtains -1 otherwise.

We now describe the three stages in detail.

1) Primary Election. The primary stage at $t = 0$ can be seen as a principal-agent problem. In the primary election, principal G (the representative member of party G) chooses whether to give discretion or not to the representative agent in each dimension. Before rewarding the agent at the end of time 2, the principal can observe the chosen policy, but observes the state of the world (i.e., whether the chosen policy is optimal for them) with probability $q_d < 1$, reflecting several factors that undermine voters’ ability to

⁸One could easily add an extension where the probability of figuring out the state of the world by the agent is less than one and related to her ability, hence adding to the model also the adverse selection dimension. Adverse selection is beyond the scope of our theory, hence our assumption that the elected policymaker observes the state with probability 1.

assess whether the chosen policy is ex-post optimal, such as the complexity of the policy issue or the spread of misinformation around the issue. The agent receives the reward W_d^G if and only if the policy committed to ex-ante has been implemented on dimension d when commitment was required, and if and only if there is no evidence of betrayal on this dimension when discretion was granted.⁹

2) General Election. The two candidates selected in the two primaries, each with their own mandate by their party principal, compete in a general election at $t = 1$. The winner of the general election chooses a policy in each dimension at time 2.

To determine endogenous turnout in the general elections, we make an “as if pivotal” assumption (Alesina & Rosenthal 1996), i.e., we assume that each citizen votes in the general elections if and only if the difference in her utility between the candidates of the two parties is higher than her cost of voting. The cost of voting for citizen v of party G is denoted by $c_v^G + \mu^G$, where c_v^G is distributed according to a uniform distribution on $[0, \bar{\phi}]$, with $\bar{\phi} > 0$, and $\mu^G \in [-\frac{1}{2\bar{\psi}}, \frac{1}{2\bar{\psi}}]$ is a party-specific voting cost common to all members of G that modifies the distribution of voting costs after the primaries and before the general elections. μ^G is, for simplicity, distributed uniformly. For each citizen, the cost of voting materializes after the primary elections and before the general elections. Elections are governed by plurality rule. If the candidates receive an equal number of votes, the election outcome is determined by a fair coin toss.¹⁰

In order to limit the attention to an interior and unique solution, we assume $\psi \leq \bar{\psi}$, with $\bar{\psi} > 0$ sufficiently low.

3) Policymaking. The elected agent chooses a policy in $t = 2$. On each dimension, we assume that if the policy that is optimal for the elected agent’s party differs from the policy preferred by the relevant elite, then the elite offers a monetary bribe m_d to the policymaker, and ex-ante such an offer is unknown and drawn from a uniform distribution from zero to $\bar{m}_d$.

3.2 Equilibrium analysis

We solve the game by backward induction.

We first analyze the decision of a policy-maker at time $t = 2$.

⁹In a dynamic model with re-election, the most intuitive interpretation of W_d^G is the expected number of party G members for which this dimension is salient; these party members will support the agent’s in the next party’s primary if and only if she adhered to the contract in dimension d . In our one-shot model, the principal has no cost in withholding W_d^G in the event of contract violation. Even in such a dynamic extension, voters may want to punish politicians who deviate from their electoral commitments without confirming them in office (see the survey by Duggan & Martinelli 2017), but they may not, depending on the replacement pool. These dynamic tradeoffs are avenue for future research.

¹⁰The cost of voting can be ignored for primaries, given that all voters of the same party have the same beliefs.

Suppose that the elite offers a bribe of a finite amount m_d to the agent to distort the policy in dimension d . The agent is captured only if the expected benefits of accepting a bribe exceed its expected cost. If the agent who has been given discretion in dimension d chooses not to implement the ex-post optimal policy in that dimension, she may incur punishment only when the principal observes the state of the world and realizes that the implemented policy is not the ex-post optimal one. Conversely, if a committed agent chooses not to implement the committed policy r_d , she is penalized with probability one.¹¹

Consider first an agent with discretion, and consider the two potential conflicts of interest scenarios, namely when state is r_d and elite prefers s_d and vice versa. The agent receives W_d^G when choosing the optimal policy, while she receives $W_d^G(1 - q_d) + m_d$ otherwise. The agent chooses to accept a bribe m_d when $m_d > q_d W_d^G$. Thus, the interim probability of the uncommitted agent's deviation from optimality in those situations where the optimal policy is not the one preferred by the elite is $p_d^G \equiv \max\{0, \frac{\bar{m}_d - q_d W_d^G}{\bar{m}_d}\}$. The endogenous probability of capture in dimension d , p_d^G , is decreasing in q_d . We can interpret a lower q_d as a policy with higher complexity or diffused misinformation, which makes it harder for voters to observe (and possibly understand) ex-post the correct policy. Hence, the probability of capture under discretion increases with either sources of complexity. In other words, cognitive complexity and/or misinformation that reduce q_d lead to distrust in uncommitted agents.¹²

Turning to commitment, the agent receives W_d^G when she fulfills the commitment, while she receives m_d in case she accepts a bribe to violate the commitment. The probability of capture is $\pi_d^G \equiv \max\{0, \frac{\bar{m}_d - W_d^G}{\bar{m}_d}\}$. π_d^G is therefore strictly less than p_d^G for every $q_d < 1$ if $\bar{m}_d > q_d W_d^G$, and they are both zero otherwise.¹³ For example, if $\bar{m}_d = W_d^G$ then π_d^G is zero while p_d^G is strictly positive and decreasing in q_d .

Remark 1. *The difference in the probabilities of capture, $p_d^G - \pi_d^G$, increases with complexity of figuring out the optimal policy in dimension d for a voter.*

We now look at voters' behavior in the general election at $t = 1$. The platform of a candidate $k = l, r$, which we denote by S_k , specifies whether the candidate commits to or retains discretion over each policy dimension. Each voter in the general election decides whether to abstain or to vote. If a citizen does not abstain, then she votes for the candidate who maximizes her expected utility. Let $EU^G(S_k)$ denote the expected

¹¹Commitment to punishing an agent who did not implement the commitment is optimal ex-ante in our one-shot model, but in a dynamic extension one would need to replace this assumption with the standard anticipation of the rational behavior of voters at the time of re-election.

¹²We can also note a complexity paradox: the more complex the salient policy dimensions are, the more distrust there is, and hence the greater demand (and supply, as we see below) of policy commitments.

¹³The inequality $\pi_d^G < p_d^G$ is important, since it is the simple answer to the frequently asked question "if there is little trust why should a voter trust a candidate who makes a commitment."

utility for a voter of party G if a candidate with platform S_k is elected. Voter v votes for candidate k rather than abstaining if $EU^v(S_k) - EU^v(S_{k'}) \geq c_v^G + \mu^G$.

We now move to analyze the primary elections which determine the policy platforms that are proposed by the candidates of party L and R respectively. Denote by $\hat{S}^G$ the electoral platform that provides the highest expected utility to the members of party G . The following lemma is a key result in our analysis, as it demonstrates that, under our assumptions, each party selects the platform that maximizes its expected utility. Consequently, the determination of the platform offered by each candidate can be analyzed independently.

Lemma 1. *In each party G 's primary elections, the winner proposes $\hat{S}^G$.*

The intuition for the proof is the following. Even in cases in which the most desirable candidate and platform conditional on winning ($\hat{S}^G$) could potentially not be the one that maximizes the probability of winning in the general elections against any given opponent, sufficient uncertainty about the realization of costs of voting and/or the strategy of the opponent makes $\hat{S}^G$ preferable for voters of party G .¹⁴

The immediate implication of Lemma 1 is that we can focus on each party's preferences in the choice of the platform made during the primaries.

3.3 Trust, complexity, and preferences for commitment

Now we can finally analyze the determinants of the choices of each party at the primary elections. Party G chooses an n -dimensional policy platform such that for each dimension d , the candidate is either committed to implement a policy or has discretion to choose the policy in case she becomes a policymaker.

The expected utility for principal G from giving discretion for dimension d is

$$EU_d^G(discretion) = -p_d^G[\gamma_d^G \lambda_d^G + (1 - \gamma_d^G)(1 - \lambda_d^G)], \quad (1)$$

where $\gamma_d^G \lambda_d^G + (1 - \gamma_d^G)(1 - \lambda_d^G)$ is the probability of conflict of interest between the elite and the voters of party G and p_d^G is the probability that an uncommitted agent is captured on dimension d .

Consider the case $\gamma_d^G \geq \frac{1}{2}$, i.e., the principal of party G in case of commitment com-

¹⁴This lemma holds due to the assumption that ψ is sufficiently low (capturing high uncertainty about the elections), but it would also hold without such an assumption if we assumed some other form of uncertainty, e.g., about the final outcome of the primaries in the other party. For instance, assuming that members of a party did not know the preferences for policies of members of the opposing party, or assuming that members of a party observe the electoral platform of the opposing candidate with a probability lower than 1.

mits to r .¹⁵ The expected utility of the principal G in case of commitment in dimension d is

$$EU_d^G(\text{commitment}) \equiv -(1 - \gamma_d^G)[(1 - \lambda_d^G) + \lambda_d^G(1 - \pi_d^G)] - \gamma_d^G \lambda_d^G \pi_d^G. \quad (3)$$

The first term captures the fact that when the optimal policy in dimension d for party G is s (probability $1 - \gamma_d^G$), then the principal suffers a loss when the commitment to r goes through, which is the case if the elite likes the reform (probability $1 - \lambda_d^G$) in which case she does not need to bribe, or when the elite would prefer the status quo but the agent does not find the temptation m sufficiently high (probability $\lambda_d^G(1 - \pi_d^G)$). The second term captures the fact that, when the state of the world is r , there is an expected loss when the elite manages to block it, which happens with probability $\lambda_d^G \pi_d^G$.

Comparing the two expected utilities, commitment is preferable if

$$1 - \gamma_d^G \leq p_d^G[\gamma_d^G \lambda_d^G + (1 - \gamma_d^G)(1 - \lambda_d^G)] - \pi_d^G \lambda_d^G (2\gamma_d^G - 1). \quad (4)$$

Proposition 1. *There always exists $\bar{q}_d^G(\lambda_d^G, \gamma_d^G) > 0$ such that for all $q_d < \bar{q}_d^G(\lambda_d^G, \gamma_d^G)$ the probability of capture $p_d^G(q_d)$ is sufficiently high to induce a preference for commitment in dimension d for party G .*

The link between complexity q_d and the probability of capture p_d^G shows that commitment candidates can serve as a safeguard in environments characterized by distrust. In reality, both committed politicians and those granted discretion risk being captured by special interests; however, the more complex it is to acquire information on the issue, the more tempting it becomes to opt for a simple commitment.

Let us now study how the threshold $\bar{q}_d^G$ and the corresponding threshold of distrust generating indifference between commitment and discretion change with the other parameters of the model.

Remark 2. *If $\gamma_d^G > 1/2$, then the threshold $\bar{q}_d^G$ increases with γ_d^G and λ_d^G . In words, a stronger ex-ante belief that the reform is the optimal policy, combined with a greater fear of elite obstruction, increases distrust and, consequently, the demand for commitment.*

On the other hand, if $\gamma_d^G < 1/2$, then $\bar{q}_d^G(\gamma_d^G, \lambda_d^G)$ decreases in both arguments. In this case, distrust and the demand for commitment grow as voters become more convinced that the status quo is desirable and more fearful that the elite is instead pushing for reform.

¹⁵If $\gamma_d^G < \frac{1}{2}$, the principal in case of commitment commits to s . The expected utility of the principal G in case of commitment is

$$EU_d^G(\text{commitment}) \equiv -\gamma_d^G[\lambda_d^G + (1 - \lambda_d^G)(1 - \pi_d^G)] - (1 - \gamma_d^G)(1 - \lambda_d^G)\pi_d^G. \quad (2)$$

The propositions always consider both cases: $\gamma_d^G \geq \frac{1}{2}$ and $\gamma_d^G < \frac{1}{2}$.

Corollary 1. *An increase in the conflict of interest, increases distrust and consequently the demand of commitment.*

The ex-ante considerations about γ_d^G and λ_d^G vary across dimensions and parties. Consider, for example, the issue of gun control in the United States. Republican voters are more likely to believe that the status quo is the correct policy to uphold (low γ), whereas Democratic voters are more inclined to support a reform that increases regulation (high γ), as they assign a higher probability to it being the optimal policy choice. For simplicity, the elite – which could be captured by the National Rifle Association (NRA) – prefers a policy that preserves the status quo. Hence, the probability of conflict of interest between Republicans and the elite is low, whereas it is high for Democrats. Conversely, when it comes to the issue of immigration, Republican voters are more likely to view a reform that restricts immigration as the correct policy and believe that the “elite” opposes such a reform. In contrast, Democratic voters are more inclined to believe that such a reform should not be implemented.

Consider again the case in which party G ex-ante believes that the reform is the correct policy, $\gamma_d^G \geq \frac{1}{2}$. The stronger the ex-ante conviction of party G ’s voters that the reform is the correct policy, the higher is γ_d^G , the more likely it is that the optimal policy platform includes commitment on that dimension. Moreover, the higher the perceived probability of a conflict of interests between the elite and party G ’s voters on a dimension d , the more likely it is that the optimal policy platform includes commitment on that dimension. In fact, the probability of a conflict of interests is equal to $\gamma_d^G \lambda_d^G + (1 - \gamma_d^G)(1 - \lambda_d^G)$ and, ceteris paribus, an increase in λ_d^G increases the probability of conflict of interests.

To reiterate, the probability of capture of a politician in dimension d can be interpreted as a measure of voters’ distrust: the higher the probabilities p_d^G and π_d^G , the stronger the belief among voters that politicians will not act in their best interest. However, a corollary of our analysis is that an increase in complexity increases distrust only towards the uncommitted politician, i.e. $\frac{\partial p_d^G}{\partial q_d} \leq 0$. This is the key reason why the internet and the advent of social media, by decreasing q across the board, are a fundamental cause of distrust in uncommitted representatives.

3.4 Turnout implications

In this section, we illustrate how changes in voters’ perception of the risk of elite capture modify parties’ turnout. We focus on the relevant case in which, on at least one dimension, one candidate is committed and one is not.

Keeping everything else fixed, assume that on dimension d candidate g is committed

to do r_d , given $\gamma_d^G > 1/2$. At the same time, the opponent g' is uncommitted. Given Proposition 1 and Remark 2 following the proposition, this means that the complexity q_d for voters is less than $\bar{q}_d^G$ but greater than $\bar{q}_d^{G'}$.

Assume also that $\lambda_d^G \geq 1 - \gamma_d^{G'}$, which implies that a voter in party G prefers a non-captured policy-maker of party G' to a captured one. Therefore, if candidate g' is elected, all voters prefer that politician g' is not captured by the elite.

Proposition 2. *Suppose that on dimension d , candidate g offers commitment to reform r_d and candidate g' has been given discretion. Suppose also that $\lambda_d^G \geq 1 - \gamma_d^{G'}$, meaning that all voters prefer that politician g' is not captured by the elite.*

If q_d decreases, then the turnout of party G increases, while the turnout of party G' decreases.

The intuition behind this result is as follows. Consider any shock occurring between the primary and the general election that leads to a decrease in q_d . Given that $\frac{\partial p_d^G}{\partial q_d} \leq 0$, distrust toward a politician who did not commit increases. Therefore, the expected utility of both principals of having an uncommitted policymaker decreases, and therefore the probability that the committed candidate wins the election increases.

The above result highlights how misinformation, one possible source of lower q_d , benefits committed candidates, for it makes it harder for voters to discern the true state of the world and, consequently, to hold a discretionary policymaker accountable.

3.5 From Commitment to Populism

After elucidating the mechanics of the determinants of the shift to commitment politics, we can show how the main features of populism, such as anti-elite rhetoric and aversion to checks and balances, emerge as complementary strategies of a commitment candidate. Therefore, through the lens of our theory of commitment politics, we are able to parsimoniously explain the recent rise of populism while keeping with the main findings in the literature on its main features – anti-elitism and illiberalism – and its causes, such as the advent of social media (and ensuing decrease in q_d) and the recent changes in the economic and cultural spheres which decreased trust in traditional tools or representative government (captured by p).

Anti-elite Rhetoric. When a candidate who commits to policy r on dimension d (and therefore $\gamma_d^G \geq \frac{1}{2}$) faces an opponent who retains discretion on that same dimension, the committed candidate g benefits from an increase in both γ_d^G and λ_d^G . A rise in either parameter boosts turnout among party G voters, thereby increasing the likelihood that candidate g wins the election. This implies that a candidate has strategic incentives to

use anti-elite rhetoric to persuade their voters that the elite holds opposing preferences and that the probability of a conflict of interest between the voters and the elite is high.

Proposition 3. *Suppose that on dimension d , candidate g commits to implementing reform r and candidate g' retains discretion. An increase in either λ_d^G or γ_d^G increases the turnout of party G and the probability that candidate g wins the election.*

A symmetric result holds if candidate g is committed to the status quo and $\gamma_d^G < 1/2$ while candidate g' retains discretion. In this case, a decrease in either λ_d^G or γ_d^G increases the turnout of party G and the probability that candidate g wins the election.

We remark that a vicious circle materializes between fear of the elite and distrust in politicians: when $\gamma_d^G > 1/2$ we know that $\bar{q}_d$ goes up, hence it is more likely that the true q_d falls below such a threshold, which determines sufficient distrust to determine in turn the choice of commitment; but we have now shown that a committed candidate has mobilization incentives to further increase the fear of the elite.

Aversion to Checks and Balances. A committed candidate is evaluated by their voters solely based on whether they fulfill their electoral promises, rather than on the actual merits or outcomes of the policies implemented. This fundamental feature of our analysis has important implications also for voters' preferences for checks and balances, broadly captured by free media, expert bureaucracies, and an independent judicial system. We argue that voters' preferences for commitment are correlated with voters' preferences for a limitation of the independence of agencies that can block a policy implemented by politicians. To understand why, consider a party G with $\gamma_d^G > \frac{1}{2}$. By Remark 2, we know that the threshold $\bar{q}_d^G$ is increasing in γ_d^G . Assume that – given λ_d^G and q_d – for $\gamma_d^G = \frac{1}{2}$ party G prefers delegation while for $\gamma_d^G = 1$ it prefers commitment to r_d . Consider an agency of restraint (for instance the judicial system) that can block a reform r_d . Suppose that the principal G assigns probability ρ_d^G that the agency blocks a reform when s_d is the preferred policy by G and probability σ_d^G that the agency blocks a reform when r_d is the preferred policy. Consider the expected utility of G from giving discretion. If the agency operates, then

$$EU_d^G(\text{discretion})|_{\text{agency}} = -p_d^G[\gamma_d^G \lambda_d^G + (1 - \gamma_d^G)(1 - \lambda_d^G)(1 - \rho_d^G)] - (1 - p_d^G)\gamma_d^G \sigma_d^G. \quad (5)$$

The expected utility of G from giving discretion if the agency does not operate is equal to

$$EU_d^G(\text{discretion})|_{\text{Noagency}} = -p_d^G[\gamma_d^G \lambda_d^G + (1 - \gamma_d^G)(1 - \lambda_d^G)].$$

The difference between the two utilities is equal to

$$p_d^G \rho_d^G (1 - \gamma_d^G) (1 - \lambda_d^G) - (1 - p_d^G) \gamma_d^G \sigma_d^G. \quad (6)$$

Similarly, the difference between the expected utilities from commitment with and without the agencies of restraint is equal to

$$\rho_d^G (1 - \gamma_d^G) [(1 - \lambda_d^G) + \lambda_d^G (1 - \pi_d^G)] - \gamma_d^G (1 - \pi_d^G) \sigma_d^G. \quad (7)$$

As γ_d^G increases, *ceteris paribus*, commitment becomes more attractive than discretion. An agency of restraint benefits a voter when it blocks a reform that is undesirable to her—either because it is promoted by politicians from the opposing party or by politicians captured by elites with opposing preferences. Conversely, it imposes a cost when it blocks a reform that the voter considers desirable. Whether the expected benefit outweighs the expected cost depends on the probabilities the voter assigns to these events. Voters who strongly believe that a reform on dimension d is beneficial to them - that is, when γ_d^G is high - assign a low probability to the reform being undesirable. Since high conviction increases the likelihood that voters prefer commitment over delegation (see Remark 2), we may observe a correlation between preferences for commitment and preferences for weaker checks and balances. The previous discussion is summarized in the following remark.

Remark 3. *Preferences for commitment are correlated with preferences for weaker checks and balances: the stronger the conviction among voters of party G that the reform is the correct policy (i.e., the higher γ_d^G is), the greater their support for commitment, and the lower their expected utility from the existence of an agency of restraint.*

4 Empirical Evaluation

We provide evidence in favor of the main predictions contained in our propositions, both on the supply of commitment by candidates and on voters' demand for commitment.

We test the predictions in the context of the U.S. House elections, which display a good fit with our theory. First, the electoral system for the U.S. House of Representatives allows for a clear partitioning of voters into electoral districts where candidates run for office, providing a wide scope for comparisons. Second, the first-past-the-post plurality rule and the presence of primary elections closely mirror key features of our model.

Before presenting our measurement strategy and empirical design, some clarifications on the link between theory and empirical implementation are warranted.

First, while the theory offers a dimension-specific analysis of the demand and supply of commitment across a multi-dimensional policy platform, it is empirically infeasible to construct reliable measures of distrust toward politicians along specific policy dimensions. As a result, although we can identify the policy dimensions on which candidates supply commitment, we cannot directly map dimension-specific commitments to corresponding measures of complexity (q_d) or distrust (p_d). We therefore rely on average levels of distrust toward the federal government as a proxy for p across dimensions. However, we can leverage available data on voters’ issue-specific opinions to test a key comparative statics prediction of the model—namely, that stronger convictions within a policy area lead to greater commitment by candidates.

Second, although Proposition 1 shows that sufficient policy complexity or disinformation within a dimension leads to heightened distrust and, in turn, a shift toward commitment “contracts,” we are unable to empirically map complexity to distrust at the dimension level. Our empirical strategy thus focuses on testing the second key implication: that higher levels of distrust—regardless of their source—are associated with increased use of and preference for commitment.

4.1 Supply Side Evidence

In this section, we study the supply of commitment. Our focus is on candidates’ communication on social media. We rely on Twitter, one of the social networks with the highest level of penetration in U.S. politics, to observe and measure how candidates appeal to voters by committing to policy. In particular, we show that

- (i) Distrust increases the supply of commitments and populist rhetoric, a strategy implemented by commitment candidates, as stated in Proposition 3.
- (ii) Commitment is more likely on topics where voters have high ex-ante policy convictions (Remark 2).
- (iii) Distrust increases turnout by supporters of parties running with a commitment candidate and decreases turnout by supporters of candidates running on discretion (Proposition 2).

4.1.1 Data and Measurement Strategies

First, we describe the data we use for the empirical evaluation of the predictions. To measure distrust (p), we combine public opinion surveys with geo-coded respondents and statistical models to derive estimates of distrust for each congressional district. We then apply natural language processing techniques to the universe of tweets posted by

candidates to detect policy commitments and populist rhetoric. Finally, we rely on survey data on voters’ policy preferences to select the policies with high within-party convictions.

Distrust in Government. We produce time-changing estimates of distrust in the federal government at the congressional district level by applying multilevel regression and post-stratification (MRP) techniques on the three most recent waves of the American National Elections Study, the only survey including questions on trust in government across multiple waves and with information on the congressional districts of respondents. MRP allows to provide more precise estimates for subgroups that may not have enough sample size in the survey data alone. Because ANES respondents are representative of the national voting-age population, we can obtain estimates of distrust for each congressional district by modeling individual survey responses as a function of individual-level variables and variables at the district- and state-level in a multilevel model and then post-stratify the predictions with census data (Gelman & Little 1997).

We use the three most recent ANES waves (ANES 2013, 2017, 2021) and produce estimates of distrust for the 435 congressional districts over three time periods, 2012, 2016, and 2020. To measure distrust, we use the question “How often do you trust the government in Washington to do what is right?” and re-code the variable in order to have dichotomous answers which allows us to estimate the share of distrustful individuals in each district. Respondents who “Never” trust the government are coded as 1, whereas respondents who answered “Always”, “Most of the time”, “About half of the time”, or “Some of the time” were coded as 0. This question is particularly suitable for evaluating the predictions of the model.¹⁶ By referring to the government in Washington, respondents are prompted to think about the same entity when answering the question and not state of local governments. In Table 1 we report the sample size and the share of respondents who report never trusting the government in Washington across each wave.¹⁷

We produce the estimates in two steps that we perform automatically and separately for any given year with the `AutoMrP` package in R, which uses machine learning algorithms to improve standard MRP models (Broniecki et al. 2022).¹⁸ First, we fit a multilevel model to the survey data to predict the level of distrust in government for specific categories of

¹⁶In the appendix (Table G.13) we show that the results are similar when using alternative survey questions: “How many of the people running the government are corrupt?” (Distrust = 1 for “All”/“Most” responses), “Would you say the government is pretty much run by a few big interests looking out for themselves or that it is run for the benefit of all the people?” (Distrust = 1 for “Run by a few big interests” responses).

¹⁷The size is large enough to create valid estimates of district-level responses. Warshaw & Rodden (2012) suggest that national samples of as many as 2,500 respondents produce reliable estimates for congressional districts. In Section C1 in the appendix we validate this approach by comparing the estimates about alternative quantities of interest obtained from the 2020 ANES sample and a much larger NationScape sample, which nonetheless does not include questions on trust in government (Tausanovitch & Vavreck 2021).

¹⁸We describe the estimation in greater detail in Section C in the appendix.

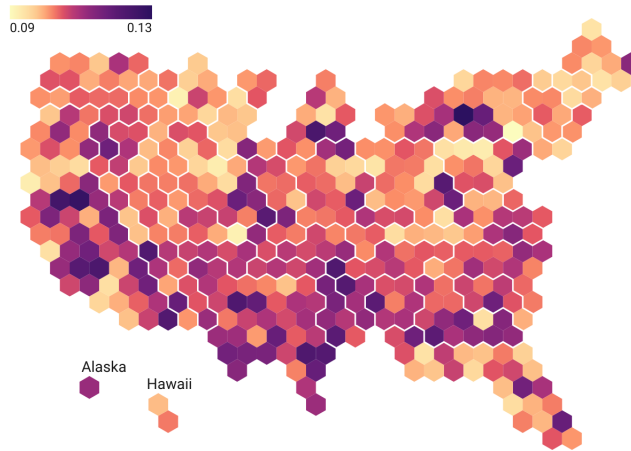
Table 1: ANES Surveys, Descriptive Statistics.

Year	N. Respondents	% Distrust
2012	5,663	5.9
2016	4,079	12.8
2020	7,805	9.6

Notes: Share of respondents who never trust the government in Washington to do what is right across the three ANES waves.

respondents (i.e., combining sex-education-age variables into unique categories of “ideal types” distributed across the congressional districts). Once obtained a predicted value for each ideal type in any given district, in the second step we calculate the weighted average level of distrust, where the weights are determined by how prevalent the ideal types are in the population of each congressional district (i.e., post-stratification step). Data for post-stratification comes from the American Community Survey.

In Figure 1, we plot the estimates for the year 2020. Each same-size hexagon represents a congressional district and darker shades indicate higher distrust. The average share of estimated distrustful individuals across the districts is 10%, compared to 6% in 2012 and 13% in 2016. Mississippi and Louisiana are the states where average distrust is highest in 2020, each with a share of distrustful individuals above 10%. There is also significant within-district variation in distrust, which is the identifying variation we leverage for the statistical analysis that follows. If we consider the standard deviation of the 2020 estimates as a benchmark – i.e., the cross-sectional variation displayed in Figure 1 – the standard deviation of the within-district change between 2016 and 2020 is twice as large.

Figure 1: MRP estimates of distrust.

Notes: MRP 2020 estimates of distrust in government (share of voting-age individuals who never trust the government in Washington to do what is right) for each congressional district.

Supply of Commitment and Populist Rhetoric. To measure the supply of commitments and populist rhetoric, we look at the Twitter activity of U.S. candidates to the House. We obtain data on the names of candidates and the districts where they ran from the *Candidates in American General Elections* database compiled by Cha et al. (2021). We then scrape the Twitter handles of candidates from several sources, which we describe in detail in Section D in the appendix.

3,579 unique candidates ran for office between 2012 and 2020. We recover at least one account for 64.1% of the candidates. Given the extensive scraping and manual verifications performed, we are confident that the 1,283 candidates for which we do not find an account indeed did not have an active Twitter account at the time of data collection. We downloaded 5.9 million tweets (in English and excluding retweets) from the Twitter API for the period 2012-2021 posted by candidates with an account. Democrats’ average number of tweets is larger than that of Republicans by 45%. Democratic candidates post 75 tweets per month on average, whereas Republican candidates 52.

To classify tweets based on whether they use policy commitments and populist rhetoric, we train and validate two machine learning classifiers that compute the predicted probability of a tweet containing policy commitments and populist rhetoric, respectively. Our approach follows the standard steps of classification tasks, which are described in greater detail in Section E in the appendix.

First, we represent tweets as numerical vectors through embeddings, which will serve as a matrix of predictors.

Second, we build and annotate two training datasets (3,000 tweets for commitment and 3,000 tweets for populist rhetoric). In particular, we use the `gpt-3.5-turbo` model developed by Open AI. We annotate the tweets in the training datasets with the assistance of large language models, which have been shown to outperform crowd-workers for standard annotation tasks Gilardi et al. (2023). To demonstrate the validity of GPT answers, we show that they are comparable to those of a human annotator trained specifically for this task.

Third, we train several machine learning classifiers on the vector representation of 75% of the tweets in the training datasets, with the goal of achieving good predictions on the 25% of held-out tweets. Among the various models we train, we select a gradient boosting classifier based on its good performance metrics.¹⁹

Finally, we use the classifiers to make out-of-sample predictions on the universe of tweets collected. Tweets with a large predicted probability of belonging to the commitment or populist rhetoric classes are assigned a score equal to 1, and 0 otherwise. We

¹⁹The classification exercises for commitment and populist rhetoric display very good performance metrics, with accuracy and F1 scores above .85.

select a high threshold of 0.80 to ensure that, for instance, references to policies without clear elements of commitment or vague policy promises are not classified as policy commitments, and that general praises of “the American people” are not classified as populist.²⁰ For instance, the classifier assigns a probability of belonging to the commitment class equal to .54 to the following tweet: “*I’m focused on solutions that benefit families in #MD6*” (John K. Delaney (D-MD) – 2016). This tweet includes a vague commitment to policy which will not be considered a policy commitment when the probability threshold is as large as .8. Conversely, the following tweet is assigned a .95 probability of being a commitment tweet.: “*My top priority will be ending illegal immigration. I will support the President in building a wall, and push immigration policies that put the American worker before foreign workers. If you’re tired of Americans being put last, donate to my campaign!*” (Greg Musselwhite (R-FL) – 2020).²¹

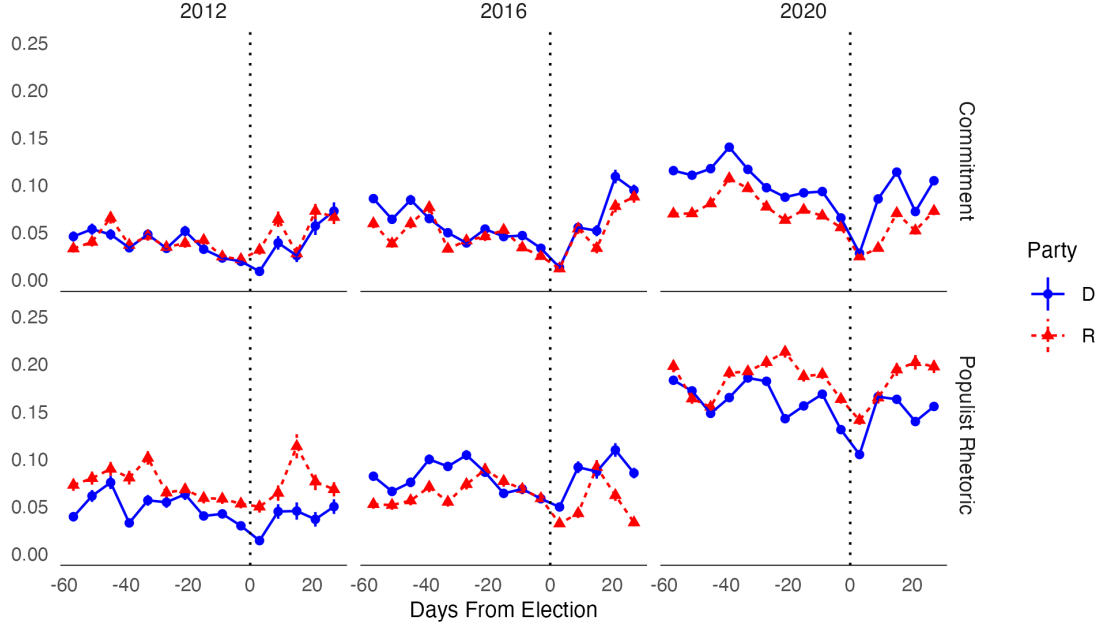
In Figure 2, we show the dynamic share of commitment and populist-rhetoric tweets posted by Democratic and Republican candidates during the 60 days before and the 30 days after the elections. Three patterns emerge from the data. First, except for the 2020 campaign – when Democrats’ share of commitments is larger than Republicans’ – the use of policy commitments on Twitter is similar for both Republican and Democratic candidates. Although slowly increasing from one election to the next, the supply of commitments remains relatively stable throughout the two months before the elections. Second, policy commitments rapidly increase in the days immediately following the elections. We interpret this spike as a confirmation of the commitments made by winning candidates – who are responsible for approximately 70% of tweets posted in the post-election period – before starting their term in Congress. In fact, it is in Congress that they will ultimately have the chance to translate those commitments into policy. Finally, with respect to populist rhetoric, there is a dramatic jump from the 2012 and 2016 elections to the 2020 elections, when the share of tweets classified as populist doubles for Democrats and becomes three times larger for Republicans.

Policy Convictions. Recall that our theory suggests that the likelihood of commitment is larger if policy convictions are high among voters (large γ_d). To test this prediction, we need to identify tweets about policies on which voters’ opinions are highly

²⁰As a robustness test, in the analysis we use the continuous predicted probability as outcome variable and find similar results (see Table G.12 in the appendix). Consistent with the standard measurement error framework, we find more precise estimates when using larger probability thresholds to dichotomize tweets, but the results are robust to different thresholds at intervals of .05, from .5 to .95 (see Figure G.4 in the appendix).

²¹In the appendix, we also report other examples of tweets posted by candidates of both parties whose probability of belonging to the commitment and populist-rhetoric class is large and medium to show how the classifier is able to distinguish between references to policies and more specific policy commitments, as well as tweets commemorating the people versus those portraying people and elite in a clear antagonistic (hence populist) fashion (see Table E.7 and Table E.8).

Figure 2: Share of commitment and populist rhetoric tweets.



Notes: Share of commitment and populist rhetoric tweets with bootstrapped confidence limits for tweets posted by Republican and Democratic candidates in the 60 days before and 30 days after the elections.

homogeneous.

To identify such policies, we use survey data from the Cooperative (Congressional) Election Study (Kuriwaki 2022) to select in a data-driven way the policy with the largest degree of agreement among Republicans or Democrats in any given state and year.²² The CCES survey regularly asks a representative sample of U.S. individuals 55 questions about policy preferences (2006-2021). We select the question with the lowest standard deviation in the responses across party-year-state combinations of respondents, which we easily map to seven different topics/policies based on the wording of the questions. The seven topics are immigration, gun regulation, abortion, military issues, environment, health care, and same-sex marriage. For instance, 95.4% of the 813 respondents identifying with the Democratic Party in 2016 in Virginia approve background checks for guns for all sales.²³

We then build topic-specific dictionaries with the 20 words most similar to the topic label. To learn the semantic relationship between words – and hence select the 20 words semantically most similar to the topic label (e.g., “gun” in the Virginia example) – we train a `word2vec` model (Mikolov et al. 2013) on the 5.9 million tweets we collected so that the model learns the relationship between words from the language used on Twitter by candidates. The model then produces a vector representation of each word and we

²²We group respondents at the state-level to ensure large samples of respondents and geographical variation in policy preferences.

²³In the appendix, we report the topics with the largest convictions by party-year-state combinations (Table F.10) and the mapping of survey questions to inferred topics (Table F.9).

select the 20 words whose vector representation is closest to the topic label. The words for each of the seven topics/dictionaries are reported in Table F.11 in the appendix.

We consider tweet i to be about topic j , with $J = j_1, \dots, j_7$, if tweet i contains at least one of the words in the dictionary of topic j . Finally, if the tweet is about topic j , and topic j is also the topic with the largest convictions in the state where the candidate posting the tweet is running, the tweet is considered to be about a high-conviction topic. To continue the example of Democrats in Virginia in 2016, a tweet given by a Democratic candidate on gun regulation in a race for a district in Virginia in 2016 will be considered to be about a high-conviction policy.

4.1.2 Evidence on Commitment and Populist Rhetoric

Several candidates ran for office more than once, allowing us to test the relationship between distrust and candidates' rhetoric by exploiting within-candidate variation in distrust. In particular, we estimate the following linear probability model:

$$y_{icsdt} = \alpha_c + \eta_d + \delta_{st} + \beta \text{Distrust}_{dt} + X'_{c dt} \zeta + \epsilon_{icsdt} \quad (8)$$

where y_{icsdt} is the probability of a commitment or populist rhetoric tweet, α_c , η_d , and δ_{st} are candidate, district, and state-by-election year fixed effects, $X'_{c dt}$ includes a set of additional time-changing covariates, namely the incumbency status of the candidate and the district-level demographics used to produce the estimates of distrust (share of black residents, median income (log), Gini index, and employment rate). Importantly, by including the same district-level covariates used to estimate distrust, the effect of distrust on the probability of commitment and populist rhetoric tweets is identified from the functional form of the distrust prediction model. Moreover, by including state-by-year fixed effects, we account for state-specific time trends. As a robustness test, we also present results without candidate fixed effects and simply conditioning on the party of each candidate. Standard errors are clustered by district.

To better characterize the use of Twitter as a communication tool during electoral campaigns, we analyze tweets posted during the two months preceding election day (included). This time frame also allows us to precisely match survey-based levels of distrust with Twitter data. In fact, ANES surveys are conducted in the run-up to the elections, from August to the election day.

In Table 2, we report the main results. The coefficients represent estimated changes in the probability of commitment and populist-rhetoric tweets resulting from a +1 standard deviation increase in distrust. In our preferred specification, a +1 SD increase in the share of individuals who never trust the government in Washington to do what is right is

Table 2: Distrust, supply of commitments and populist rhetoric.

	Commitment			Populist Rhetoric		
	(1)	(2)	(3)	(4)	(5)	(6)
Distrust	0.016*** (0.006)	0.021*** (0.008)	0.036*** (0.010)	0.021** (0.009)	0.029*** (0.010)	0.038*** (0.012)
Controls: Candidate	✓	✓	✓	✓	✓	✓
Controls: District	✓	✓	✓	✓	✓	✓
Mean DV	0.07	0.07	0.07	0.14	0.14	0.14
R ²	0.049	0.067	0.069	0.088	0.120	0.121
Observations	274,253	274,253	274,253	274,253	274,253	274,253
District FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓		✓	✓	
Candidate FE		✓	✓		✓	✓
State-Year FE			✓			✓

Notes: OLS estimates. SE clustered by district. DV is a dummy measure for commitment/populist tweets. Distrust is the standardized share of distrustful individuals in each district. Candidate covariates: party (only in columns 1 and 4, when candidate fixed effects are not included) and incumbent status. District covariates: Gini index, share of black individuals, employment rate, median income (log). Signif. codes: ***: 0.01, **: 0.05, *: 0.1.

associated with a 3.6 percentage point increase in the probability of a commitment tweet and a 3.8 percentage point increase in the probability of a populist rhetoric tweet. The effects are precisely estimated and sizable if compared to the sample mean (i.e., 7% of tweets are classified as commitments and 14% as populist-rhetoric).²⁴

Besides the positive effect of distrust on the probability of commitment and populist rhetoric tweets, we expect commitments to be more likely when the candidate tweets about topics on which voters have strong and uniform convictions. We therefore expect the average effect of a high conviction topic – i.e., holding distrust fixed – to be positive. We test this prediction by adding a term for high-conviction topic to Equation 8. As shown in Table 3, we observe a positive effect of high-conviction topics. Commitment tweets are about 17 percentage points more likely when the tweet is about a topic on which voters are highly in agreement with one another, even when distrust is held constant.

4.2 Evidence on Turnout

Proposition 2 shows that distrust increases turnout by supporters of parties running with a commitment candidate and decreases turnout by supporters of candidates running on discretion.

To illustrate the relationship between distrust, commitment, and turnout, we start

²⁴In Figure G.3 in the appendix, we show how the results are robust to using different time windows.

Table 3: High conviction topics and supply of commitments.

	Commitment				
	(1)	(2)	(3)	(4)	(5)
High Conviction Topic	0.183*** (0.013)	0.175*** (0.012)	0.175*** (0.012)	0.174*** (0.012)	0.174*** (0.012)
Distrust				0.035*** (0.010)	0.035*** (0.010)
Controls: Candidate		✓	✓	✓	✓
Controls: District		✓	✓	✓	✓
Mean DV	0.07	0.07	0.07	0.07	0.07
R ²	0.045	0.069	0.071	0.071	0.071
Observations	255,196	255,196	255,196	255,196	255,196
District FE	✓	✓	✓	✓	✓
Year FE	✓	✓			
Candidate FE		✓	✓	✓	✓
State-Year FE			✓	✓	✓

Notes: OLS estimates. SE clustered by district. DV is a dummy measure for commitment tweets. Candidate covariates: party (only in columns 1 and 4, when candidate fixed effects are not included) and incumbent status. District covariates: Gini index, share of black individuals, employment rate, median income (log). High conviction topic is a dummy equal to 1 if the tweet’s topic matches the topic of the CCES survey question with the largest share of agreement in any given state-party-year combinations. Distrust is the standardized share of distrustful individuals in each district. Signif. codes: ***: 0.01, **: 0.05, *: 0.1.

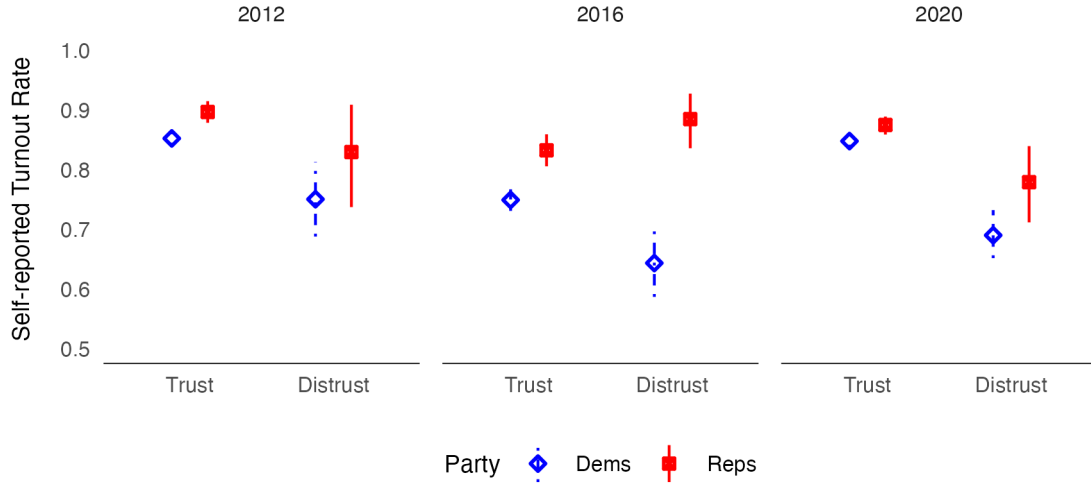
by displaying self-reported turnout for distrustful voters under a clearly populist and committed candidate: Donald Trump.²⁵ In Figure 3, we display self-reported turnout rate by party, election year, and level of trust in government. Data for party identification, distrust, and self-reported turnout are from the ANES. We cover three elections, one without Trump (2012) and two when Trump was running as the Republican nominee.²⁶

In keeping with our proposition, distrustful Democrats – whose presidential candidates were not commitment ones – are less likely to turnout compared with trustful Democrats. Conversely, Republicans who never trust the government in Washington to do what is right are more likely to vote in the 2016 presidential elections compared to both trustful Republicans in the same year and distrustful Republicans in the 2012 elections. Interestingly, we do not detect a “Trump” effect on distrustful republican voters in

²⁵As a reality check, we applied our classifier to the tweets posted by Hillary Clinton and Donald Trump during the run-up to the 2016 presidential elections and found that Trump posted 66% more commitment tweets compared to Clinton.

²⁶ANES self-reported turnout – as is generally the case for self-reported measures of turnout (e.g. Enamorado & Imai 2019) – is markedly higher than actual turnout. In our data it is as high as 86% in 2012, 77% in 2016, and 84% in 2020, against actual figures at 57%, 60%, and 67% (data from <https://www.electproject.org/>, accessed on 29 March 2023).

Figure 3: Distrust, Trump, and Turnout.



Notes: Average self-reported turnout rate in the presidential elections and 95% bootstrapped confidence limits for survey respondents identifying as Republicans and Democrats with different levels of distrust (Distrust = 1 when respondent some of the time/never trusts the government in Washington to do what is right, and = 0 when respondent always/most of the time/about half of the time trusts the government in Washington to do what is right).

2020.²⁷ To confirm the descriptive evidence from Figure 3, in the appendix we estimate the effect of the triple interaction between the Trump periods (i.e., election year dummies), self-identifying as a Republican, and reporting distrust in the federal government, and the results are widely consistent with the visual representation (see Section G1).

To provide a more rigorous test of whether the supply of commitments mobilizes distrustful voters, we return to the House elections. We rely on a panel design including a subset of ANES 2016 respondents re-interviewed in 2020, and compare individual self-reported turnout as a result of changes in the supply of policy commitments by their party's candidate. We observe self-reported turnout intentions twice, in 2016 and 2020, and we focus on respondents in districts where we observe candidates posting on Twitter, for a total of 1,038 unique individuals who answered in 2016 and 2020. 134 respondents display variation in the level of trust across the two time periods analyzed. We match each respondent with the supply of their own candidate's policy commitments from the beginning of each of the two election years until the date when the respondent was interviewed. For instance, we match a democratic respondent from the 30th district of California interviewed on August 27th, 2020 with the 71 commitment tweets that the incumbent Democratic candidate Brad Sherman had posted in the period from January

²⁷There could be several reasons why distrustful Republicans did not turn out to vote at the same rate in 2020. It is possible that incumbents – who can hardly deliver on all of their commitments – do a poorer job of winning distrustful voters compared to challengers. This finding suggests an interesting avenue for future research, where temporal dynamics and the probability of fulfilling promised commitments influence voters' incentives to participate in presidential elections.

1st, 2020 until the date of the interview. For each respondent, we then record whether they report they intend to vote for the party they identify with in the elections for the House of Representatives (in this case, the Democratic candidate). By exploiting the panel structure of the data, we can hold fixed all the time-invariant characteristics of respondents (including fixed propensity to misreport self-reported turnout intentions) and districts while also accounting for state-level time-varying shocks.

In particular, we estimate the following model of turnout:

$$y_{icdst} = \phi_i + \eta_d + \alpha_c + \delta_{st} + \tau_1 \# \text{ Own Candidate's Commitment Tweets}_{ct} + \tau_2 \text{ Distrust}_{it} + \tau_3 \# \text{ Own Candidate's Commitment Tweets}_{ct} \times \text{Distrust}_{it} + X'_{ict} \zeta + \epsilon_{icdst}, \quad (9)$$

where ϕ_i , η_d , α_c , and δ_{st} are individual-respondent, district, candidate, and state-by-election-year fixed effects, $\# \text{ Own Candidate's Commitment Tweets}$ is the logarithm of the number of commitment tweets posted by the candidate of the respondent's party. Distrust equals 1 if the respondent never trusts the government in Washington to do what is right and 0 otherwise. X' is a vector of individual-level (age, employment status, and university degree) and candidate-level (total number of tweets posted and incumbency status) covariates. τ_3 is our parameter of interest, which captures the effect of a +1% increase in commitment tweets on turnout intention of distrustful individuals.

Table 4 reports the results. There is a marked and precisely estimated increase in the probability of turnout when the supply of commitments increases, especially when respondents have low trust, which determines an increase in the probability to turnout for the respondents' own party candidate by 6-7 percentage points, depending on the specification.

4.3 Demand Side Evidence

In this section, we study the demand for commitment. To study voters' demand for commitment, we first present the results of a survey experiment where U.S. respondents were randomly assigned an informational treatment in the form of an article showing that U.S. politicians are captured by economic elites. Second, we use existing survey evidence to show that voters who prefer commitment also prefer weaker checks on executive power if their preferred candidate is in power. In particular, we show that

- (i) An exogenous increase in the perception of conflict of interest between voters and elites increases voters' distrust and consequently the demand for commitments, as stated in Corollary 1.
- (ii) Voters who prefer commitment may also prefer fewer checks and balances on their

Table 4: Supply of Commitment and Turnout of Distrustful Voters.

	Own Candidate's Turnout			
	(1)	(2)	(3)	(4)
# Own Candidate's Commitment Tweets	0.031* (0.017)	0.036** (0.018)	0.035 (0.022)	0.062** (0.026)
Distrust	-0.216** (0.098)	-0.225** (0.091)	-0.254** (0.111)	-0.318** (0.127)
# Own Candidate's Commitment Tweets $\times$ Distrust	0.047* (0.027)	0.054** (0.025)	0.057* (0.031)	0.070** (0.034)
Respondent Covariates	✓	✓	✓	✓
Candidate Covariates	✓	✓	✓	✓
Survey Weights	No	No	No	Yes
R ²	0.786	0.807	0.731	0.761
Observations	2,868	2,868	2,002	1,947
Respondent FE	✓	✓	✓	✓
Year FE	✓			
Candidate FE		✓	✓	✓
State-Year FE		✓	✓	✓
District FE			✓	✓

Notes: OLS (Columns 1-3) and WLS (Column 4) estimates. SE clustered by individual respondent. The outcome is self-reported turnout intention for the candidate of the respondent's party. Columns 4 reports estimates from using survey weights. Individual-level covariates include: age, unemployed (dummy), and university degree (dummy). Candidate-level covariates include the number of tweets posted by the respondent's own party candidate (log), and the candidate's incumbency status. Distrust = 1 if respondent never trusts the government in Washington to do what is right. # Own Candidate's Commitment Tweets is the logarithm of the sum of commitment tweets in the year of the election until the date of the interview. Signif. codes: ***: 0.01, **: 0.05, *: 0.1.

candidate once elected, compared to voters who prefer delegation (Remark 3).

4.3.1 Preferences for Commitment

We test whether an increase in the perception of conflict of interest between voters and elites increases the demand for commitment with a pre-registered online survey experiment administered to a sample of U.S. respondents ($N = 1,825$). Respondents were recruited through LUCID, and the sample is representative of the national population with respect to age, education, and gender. The experiment was conducted in February 2024. In the experiment, participants first answered a battery of pre-treatment questions and were then randomly assigned to either treatment or control conditions with equal probability. In keeping with our characterization of conflict of interest as the perception that powerful elites distort elected politicians' choices, treated respondents received a short vignette summarizing a *Vox.eu* article whose title provocatively states, "Study:

Politicians listen to rich people, not you”, and where the author summarizes a scientific study that finds that “economic elites and organized interest groups play a substantial part in affecting public policy, but the general public has little or no independent influence.” We describe the details of the experiment and report the full vignette in Section B of the appendix. Respondents in the control group receive no content.

The survey experiment allows us to test two pre-registered hypotheses. We first check whether the vignette indeed altered respondents’ level of trust. In the appendix, we show that treated respondents report lower levels of trust in political institutions (see Table B.1). For our second and main test, we estimate the causal effect of being exposed to the vignette on respondents’ preferences for commitment and, symmetrically, delegation. We ask respondents to select their level of agreement on a 1-5 strongly disagree-agree scale with the following statements:

1. Politicians should publicly commit to a clear set of policies before they enter office.
2. Politicians should try as hard as they can to implement their policy commitments once in office.
3. Politicians should receive free mandate to adjust policies that best address future challenges.
4. Politicians should have the freedom to exercise their best judgment when confronted with new information once in office.

We then estimate the following model with OLS:

$$y_i = \alpha + \beta D_i + X'_i \zeta + \epsilon_i \quad (10)$$

where y_i is the reported level of agreement with one of the four commitment/delegation statements, D_i is treatment status, and X' is a set of pre-treatment covariates. β estimates the average treatment effect of the distrust vignette. Table 5 presents our estimates for each of the four outcome statements.

Our pre-registered hypotheses are partially supported. The effect is positive but not significant when looking at the first commitment statement, but we find a positive and significant effect on the second commitment statement, which states that politicians should try as hard as they can to implement their policy commitments once in office. Simply reading about politicians being responsive to economic elites and not the general public increases preferences for commitment by 3% compared to the sample mean.

4.3.2 Preferences for Checks and Balances

In Remark 3, we showed that voters who prefer commitment may also prefer fewer checks and balances on their candidate once elected, compared to voters who prefer delegation.

Table 5: Survey Experiment: Distrust and Preferences for Commitment.

	Politicians should [...]			
	Commitment		Delegation	
	publicly commit to a clear set of policies before they enter office (1)	try as hard as they can to implement their policy commitments once in office (2)	receive free mandate to adjust policies that best address future challenges (3)	have the freedom to exercise their best judgment when confronted with new information once in office (4)
Treatment	0.02 (0.05)	0.12** (0.05)	-0.01 (0.05)	0.04 (0.05)
Covariates	✓	✓	✓	✓
Mean DV	4.09	4.04	3.39	3.69
R ²	0.08	0.05	0.09	0.07
Observations	1,715	1,714	1,714	1,715

Notes: OLS estimates. Robust SE in parentheses. Outcomes are survey items capturing preferences for commitment and delegation statements with possible answers on 1-5 Strongly disagree - Strongly agree scale. Individual-level pre-treatment covariates include: race, sex, education, religiosity, age category, salary, party identification, Trump and Biden supporter, and turnout intention. Signif. codes: ***: 0.01, **: 0.05, *: 0.1.

However, in the real world, voters cannot choose their preferred agent, and preferences for checks and balances may vary depending on who is in power. For instance, voters might be exposed to the electoral victory of commitment or delegation candidates from parties different than their own, with policy proposals that they do not support, further complicating a direct test of our Remark.

The closest possible approximation of our theory in a world with political parties and endogenous preferences is to observe preferences for checks and balances for two groups of voters over two time periods, those preferring commitment with their preferred commitment candidate in power, and those preferring delegation with their preferred delegation candidate in power. We would expect preferences for checks and balances to be weaker among committed voters-committed politician compared to discretionary voters-delegation politician. In this section, we leverage existing surveys that approximate this setting.

Our real-world case is the 2020 U.S. presidential elections, which saw Donald Trump running against the Democratic nominee Joe Biden. As previously discussed, we can consider Trump a commitment candidate and Biden a discretion candidate. Trump voters are those who prefer commitment over delegation. Biden voters prefer delegation over commitment. We shall expect Trump supporters to report weaker preferences for checks and balances in the pre-election period under Trump compared to Biden supporters in the post-election period under Biden.

We make this comparison with two sources of survey data. First, for the pre-election

period, we use the pre-election wave of the ANES 2020 (ANES 2021), which features more than 8,000 respondents who answered several batteries of questions, including voting intentions and preferences for some key democratic institutions. Second, for the post-election period, we use the ANES-GSS 2020 Joint Study ANES (2022), a collaborative survey involving the ANES and the General Social Survey (GSS) that includes 1,164 participants of the GSS 2016-2020 Panel Study who completed the same ANES 2020 questionnaire. This survey, administered online in the post-2020 election period between November 2020 and January 2021 (i.e., before the January 6th Capitol attack), hence contains the same questions on voting intentions and preferences for checks and balances as the ANES 2020 study.

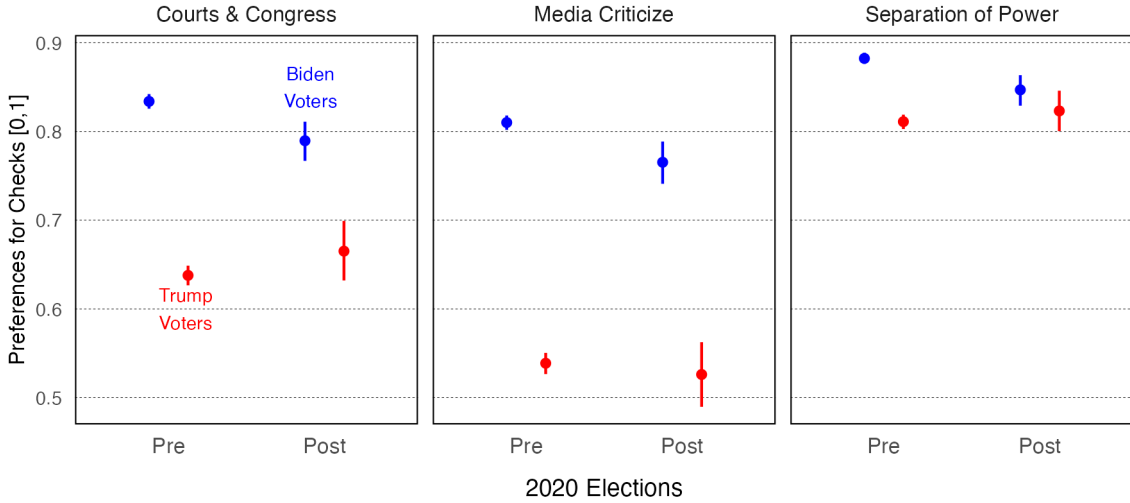
To measure preferences for commitment, which we equate with voting intentions, we use a survey question that asked respondents for whom they intended to vote (pre-election) or did vote (post-election) for President. To capture preferences for checks and balances, we focus on the following three survey questions:

1. “Would it be helpful, harmful, or neither helpful nor harmful if U.S. presidents could work on the country’s problems without paying attention to what Congress and the courts say?”
2. “How important is it that news organizations are free to criticize political leaders?”
3. “How important is it that the executive, legislative, and judicial branches of government keep one another from having too much power?”

These questions aptly capture key institutions tasked with limiting executive aggrandizement. We recode and scale all questions to range from 0 to 1, with larger values signifying stronger preferences for checks and balances.

Figure 4 shows the average preferences for checks and balances for each of these two groups of voters and over the two time periods. First, Trump voters consistently prefer weaker checks and balances, irrespective of who is in power. In fact, the pre-post election difference for Trump voters is never statistically significant for each of our three survey questions. Second, Biden voters prefer weaker checks and balances in the post-election period, when Biden won the presidency compared to the Trump period before the 2020 elections (all differences statistically significant at 95% confidence level). Third, no group-time period combination displays weaker preferences for checks and balances than Trump voters under Trump. Importantly, Trump voters in the pre-election period prefer smaller checks and balances compared to Biden voters under Biden. The difference between the average preferences for checks and balances for Trump supporters under Trump is .15, .23, and .04 points smaller compared to the average for Biden voters under Biden across the three survey questions on courts and Congress, free media, and the separation of

Figure 4: Preferences for Checks and Balances among Trump and Biden Voters.



Notes: Average preferences for checks and balances with bootstrapped confidence limit as measured by the three survey questions reported on the panel labels. Answers range from 0 to 1, with larger values signifying stronger preferences for checks and balances.

power, respectively; and these differences are always distinguishable from 0 at standard confidence levels.

These comparisons provide suggestive evidence that voters who prefer commitment do not wish their candidates' commitments to be undermined by the actions of agencies of restraint.

5 Concluding Remarks

In this paper, we have proposed a novel logical mechanism that links the erosion of trust with the observed populism wave. This mechanism is the rational increase in the demand and supply of policy commitments. More distrust induces voters, and consequently politicians, to shift from the full delegation model of representative democracy to demand and supply of firm commitments that, in their most extreme manifestations, span commitments to building walls and closing borders, to anti-immigration or protectionist policies, or other forms of identity protection pledges. In turn, such a shift to commitment politics produces a rational adoption of anti-elite rhetoric, which is the gold standard for most definitions of populist candidates and parties. Moreover, we have shown that the shift to commitment politics is not only a mechanism that determines a rationalization of the rhetorical component of populism, but it can also explain the crisis of liberal democracy in terms of preferences for reduced checks and balances and for illiberal reforms more generally.

Finally, we have shown that the demand and supply effects can create vicious cycles:

rapid changes in information environments determined by the advent of social media facilitated the spread of disinformation and complicated fact-checking, thereby amplifying voters' confusion and leading to greater distrust and, as a result, higher demand for commitment candidates. However, stronger demand for commitments creates incentives for political agents to further push their voters' beliefs in the same direction and to make information acquisition more difficult. As a result, populism becomes deeply intertwined with polarization and disinformation.

The novel empirical evidence on the U.S. validates the shift to commitment politics and its sources on the demand side, confirming the rational best response by politicians in terms of commitment as well as anti-elite rhetoric in their electoral campaigns. The supply results also confirm that politicians react even more promptly when the most salient topics display high levels of conviction, hence confirming the connection between populism, commitment, and belief polarization. Finally, our empirical findings on turnout confirm that an increase in distrust has significant asymmetric impact on turnout in case of asymmetric reliance on commitment by the two parties in electoral competition.

Future theoretical research on the dynamics of populism will need to investigate several potential sources of cycles and persistence. Occupational choice models, following Caselli & Morelli (2004), could be used to anticipate the effects of populism on the expected quality of the endogenous political class, opening up a range of additional problems for study. We have also seen that committed populists have rational incentives to amplify disinformation in order to defeat traditional candidates — another vicious circle that could foster persistence. At the same time, the tendencies to weaken checks and balances described in this paper may lead to greater concentration of blame on the executive for policy failures, potentially fueling political instability and reversals. Future research on the dynamics of populism could build on these insights, and policy debates on reforming liberal representative democracies must place “building trust” at the center of the discussion to ensure their sustainability.

References

- Acemoglu, D., Egorov, G. & Sonin, K. (2013), ‘A Political Theory of Populism’, *Quarterly Journal of Economics* pp. 771–805.
- Acemoglu, D., Robinson, J. A. & Torvik, R. (2013), ‘Why do voters dismantle checks and balances?’, *Review of Economic Studies* **80**(3), 845–875.
- Alesina, A. & Rosenthal, H. (1996), ‘A Theory of Divided Government’, *Econometrica* **64**(6), 1311–1341.
URL: <https://www.jstor.org/stable/2171833>
- Algan, Y., Guriev, S., Papaioannou, E. & Passari, E. (2017), ‘The European Trust Crisis and the Rise of Populism’, *Brookings Papers on Economic Activity* (Fall), 309–382.
URL: <https://www.jstor.org/stable/10.2307/90019460>
- Allcott, H. & Gentzkow, M. (2017), ‘Social media and fake news in the 2016 election’, *The Journal of Economic Perspectives* **31**(2), 211–235.
URL: <http://www.jstor.org/stable/44235006>
- Alsan, M., Braghieri, L., Eichmeyer, S., Kim, M. J., Stantcheva, S. & Yang, D. Y. (2023), ‘Civil liberties in times of crisis’, *American Economic Journal: Applied Economics* **15**(4), 389–421.
- Amador, M., Werning, I. & Angeletos, G.-M. (2006), ‘Commitment vs. Flexibility’, *Econometrica* **74**(2), 365–396.
- ANES (2013), ‘Anes 2012 time series study full release [dataset and documentation]’.
URL: www.electionstudies.org
- ANES (2017), ‘Anes 2016 time series study full release [dataset and documentation]’.
URL: www.electionstudies.org
- ANES (2021), ‘Anes 2020 time series study full release [dataset and documentation]’.
URL: www.electionstudies.org
- ANES (2022), ‘Anes-gss 2020 joint study [dataset and documentation]’.
URL: www.electionstudies.org
- Aridor, G., Jiménez-Durán, R., Levy, R. & Song, L. (2024), ‘The economics of social media’, *Journal of Economic Literature* **62**(4), 1422–74.
URL: <https://www.aeaweb.org/articles?id=10.1257/jel.20241743>
- Bakshy, E., Messing, S. & Adamic, L. A. (2015), ‘Exposure to ideologically diverse news and opinion on facebook’, *Science* **348**(6239), 1130–1132.
URL: <https://www.science.org/doi/abs/10.1126/science.aaa1160>
- Barbieri, F., Camacho-Collados, J., Espinosa Anke, L. & Neves, L. (2020), TweetEval: Unified benchmark and comparative evaluation for tweet classification, in ‘Findings of the Association for Computational Linguistics: EMNLP 2020’, Association for Computational Linguistics, Online, pp. 1644–1650.

- Bellodi, L., Morelli, M. & Vannoni, M. (2024), ‘A costly commitment: Populism, economic performance, and the quality of bureaucracy’, *American Journal of Political Science* **68**(1), 193–209.
- Berman, S. (2021), ‘The Causes of Populism in the West’, *Annual Review of Political Science* **24**(1), 1–18.
- Besley, T. & Persson, T. (2019), ‘Democratic Values and Institutions’, *American Economic Review: Insights* **1**(1), 59–76.
- Bonomi, G., Gennaioli, N. & Tabellini, G. (2021), ‘Identity, Beliefs, and Political Conflict’, *Quarterly Journal of Economics* pp. 2371–2411.
- Broniecki, P., Leemann, L. & Wüest, R. (2022), ‘Improved Multilevel Regression with Poststratification through Machine Learning (autoMrP)’, *Journal of Politics* **84**(1), 597–601.
- Bueno de Mesquita, E. & Friedenbergh, A. (2011), ‘Ideologues or Pragmatists?’, *Journal of the European Economic Association* **9**(5), 931–951.
- Buisseret, P. & Van Weelden, R. (2020), ‘Crashing the party? elites, outsiders, and elections’, *American Journal of Political Science* **64**(2), 356–370.
- Campante, F., Durante, R. & Sobbrío, F. (2017), ‘Politics 2.0: The multifaceted effect of broadband internet on political participation’, *Journal of the European Economic Association* **16**(4), 1094–1136.
URL: <https://doi.org/10.1093/jeaa/jvx044>
- Campbell, A., Converse, P. E., Miller, W. E. & Stokes, D. E. (1960), *The American Voter*, University of Chicago Press, Chicago.
- Canes-Wrone, B., Herron, M. C. & Shotts, K. W. (2001), ‘Leadership and pandering: A theory of executive policymaking’, *American Journal of Political Science* **45**(3), 532–550.
URL: <http://www.jstor.org/stable/2669237>
- Caselli, F. & Morelli, M. (2004), ‘Bad politicians’, *Journal of Public Economics* **88**, 759–782.
- Cha, J., Kuriwaki, S. & Snyder, J. M. (2021), ‘Candidates in American General Elections’, *Harvard Dataverse* .
- Colantone, I. & Stanig, P. (2018), ‘The Trade Origins of Economic Nationalism: Import Competition and Voting Behavior in Western Europe’, *American Journal of Political Science* **62**(4), 936–953.
- Crutzen, B. S. Y., Sisak, D. & Swank, O. H. (2024), ‘Left behind voters, anti-elitism and popular will’, *Quarterly Journal of Political Science* **19**(2), 127–156.
- Denter, P. & Ginzburg, B. (2024), ‘Troll farms’.
URL: <https://arxiv.org/abs/2411.03241>

- Di Tella, R. & Rotemberg, J. J. (2018), ‘Populism and the return of the “Paranoid Style”: Some evidence and a simple model of demand for incompetence as insurance against elite betrayal’, *Journal of Comparative Economics* **46**(4), 988–1005.
- Dornbusch, R. & Edwards, S. (1991), The Macroeconomics of Populism in Latin America, in R. Dornbusch & S. Edwards, eds, ‘The Macroeconomics of Populism in Latin America’, University of Chicago Press, Chicago.
- Duggan, J. & Martinelli, C. (2017), ‘The political economy of dynamic elections: Accountability, commitment, & responsiveness’, *Journal of Economic Literature* **55**(3), 916–984.
- Edwards, S. (2019), ‘On latin american populism, and its echoes around the world’, *Journal of Economic Perspectives* **33**(4), 76–99.
URL: <https://www.aeaweb.org/articles?id=10.1257/jep.33.4.76>
- Egorov, G. & Sonin, K. (2021), ‘Elections in Non-Democracies’, *Economic Journal* **131**(636), 1682–1716.
- Enamorado, T. & Imai, K. (2019), ‘Validating Self-Reported Turnout by Linking Public Opinion Surveys with Administrative Records’, *Public Opinion Quarterly* **83**(4), 723–748.
- Enke, B. (2020), ‘Moral values and voting’, *Journal of Political Economy* **128**(10), 3679–3729.
URL: <https://doi.org/10.1086/708857>
- Fearon, J. D. (1999), Electoral accountability and the control of politicians: Selecting good types versus sanctioning poor performance, in B. Manin, A. Przeworski & S. Stokes, eds, ‘Democracy, Accountability, and Representation’, Cambridge University Press, Cambridge.
- Ford, R. & Jennings, W. (2020), ‘The Changing Cleavage Politics of Western Europe’, *Annual Review of Political Science* **23**, 295–314.
- Foster, C. & Frieden, J. (2017), ‘Crisis of trust: Socio-economic determinants of europeans’ confidence in government’, *European Union Politics* **18**(4), 511–535.
- Fox, J. & Shotts, K. W. (2009), ‘Delegates or trustees? A theory of political accountability’, *Journal of Politics* **71**(4), 1225–1237.
- Funke, M., Schularick, M. & Trebesch, C. (2023), ‘Populist leaders and the economy’, *American Economic Review* **113**(12), 3249–88.
- Gelman, A. & Little, T. C. (1997), ‘Poststratification into many categories using hierarchical logistic regression’, *Survey Methodology* **23**(2), 127–135.
- Gennaioli, N. & Tabellini, G. (2023), Identity Politics.
URL: <https://ssrn.com/abstract=4410239>
- Gennaro, G., Lecce, G. & Morelli, M. (2024), ‘Intertemporal evidence on the strategy of populism in the united states’, *Electoral Studies* **92**, 102853.
URL: <https://www.sciencedirect.com/science/article/pii/S0261379424001112>

- Ghosh, B. S. & Tripathi, V. (2012), ‘Ideologues Beat Idealists’, *American Economic Journal: Microeconomics* **4**(2), 27–49.
- Gilardi, F., Alizadeh, M. & Kubli, M. (2023), ‘ChatGPT outperforms crowd workers for text-annotation tasks’, *Proceedings of the National Academy of Sciences* **120**(30).
- Golder, M. (2016), ‘Far right parties in europe’, *Annual Review of Political Science* **19**(Volume 19, 2016), 477–497.
URL: <https://www.annualreviews.org/content/journals/10.1146/annurev-polisci-042814-012441>
- Graham, M. H. & Svolik, M. W. (2020), ‘Democracy in America? Partisanship, Polarization, and the Robustness of Support for Democracy in the United States’, *American Political Science Review* **114**(2), 392–409.
- Gratton, G. & Lee, B. E. (2023), Drain the Swamp: A Theory of Anti-Elite Populism.
- Gratton, G. & Morelli, M. (2022), ‘Optimal Checks and Balances Under Policy Uncertainty’, *International Economic Review* **63**(2), 549–569.
- Guiso, L., Herrera, H., Morelli, M. & Sonno, T. (2024), ‘Economic insecurity and the demand for populism in europe’, *Economica* **91**(362), 588–620.
URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/ecca.12513>
- Guiso, L., Morelli, M., Sonno, T. & Herrera, H. (2021), The Financial Drivers of Populism in Europe.
URL: <http://dx.doi.org/10.2139/ssrn.3922661>
- Guriev, S. & Papaioannou, E. (2022), ‘The Political Economy of Populism’, *Journal of Economic Literature* **60**(3), 753–832.
- Guriev, Sergei, Melnikov, N. & Zhuravskaya, E. (2021), ‘3G Internet and Confidence in Government’, *Quarterly Journal of Economics* pp. 2533–2613.
- Günther, L. R. K. (2024), ‘Political representation gaps and populism’.
URL: <https://ssrn.com/abstract=4230288>
- Halberstam, Y. & Knight, B. (2016), ‘Homophily, group size, and the diffusion of political information in social networks: Evidence from twitter’, *Journal of Public Economics* **143**, 73–88.
URL: <https://www.sciencedirect.com/science/article/pii/S0047272716301001>
- Hawkins, K. A., Aguilar, R., Castanho Silva, B., Jenne, E. K., Kocijan, B. & Rovira Kaltwasser, C. (2019), Measuring Populist Discourse: The Global Populism Database.
URL: https://populism.byu.edu/App_Data/Publications/Global_Populism_Database_Paper.pdf
- Kartik, N., Van Weelden, R. & Wolton, S. (2017), ‘Electoral Ambiguity and Political Representation’, *American Journal of Political Science* **61**(4), 958–970.
- Kaufman, R. R. & Stallings, B. (1991), *2. The Political Economy of Latin American Populism*, University of Chicago Press, Chicago, pp. 15–44.
URL: <https://doi.org/10.7208/9780226158488-005>

- Krosnick, J. A. & Alwin, D. F. (1989), ‘Aging and susceptibility to attitude change’, *Journal of Personality and Social Psychology* **57**(3), 416–425.
- Kuriwaki, S. (2022), ‘Cumulative CCES Common Content’.
URL: <https://doi.org/10.7910/DVN/II2DB6>
- Levy, G. & Razin, R. (2019), ‘Echo chambers and their effects on economic and political outcomes’, *Annual Review of Economics* **11**, 303–328.
- Levy, G., Razin, R. & Young, A. (2022), ‘Misspecified Politics and the Recurrence of Populism’, *American Economic Review* **112**(3), 928–62.
- Manacorda, M., Tabellini, G. & Tesei, A. (2022), ‘Mobile internet and the rise of communitarian politics’, *CESifo Working Paper* (9955).
- Mikolov, T., Chen, K., Corrado, G. & Dean, J. (2013), ‘Efficient estimation of word representations in vector space’, *1st International Conference on Learning Representations, ICLR 2013 - Workshop Track Proceedings* pp. 1–12.
- Mudde, C. (2004), ‘The Populist Zeitgeist’, *Government and Opposition* pp. 541–563.
- Nguyen, D., Frey, V., González, S. & Brezzi, M. (2021), Survey design and technical documentation supporting the 2021 OECD Survey on Drivers of Trust in Government Institutions, Technical Report 53, Organisation for Economic Co-operation and Development (OECD).
- Norris, P. & Inglehart, R. (2019), *Cultural Backlash: Trump, Brexit, and Authoritarian Populism*, Cambridge University Press, Cambridge.
- Petrova, M., Schubert, G., Taska, B. & Yildirim, P. (2024), Automation, career values, and political preferences, Working Paper 32655, National Bureau of Economic Research.
URL: <http://www.nber.org/papers/w32655>
- Pew Research Center (2024), ‘Public trust in government: 1958–2024’. Accessed: 2025-05-02.
URL: <https://www.pewresearch.org/politics/2024/06/24/public-trust-in-government-1958-2024/>
- Rodriguez, P. L. & Spirling, A. (2022), ‘Word embeddings: What works, what doesn’t, and how to tell the difference for applied research’, *Journal of Politics* **84**(1922658), 101–115.
- Rodrik, D. (2018a), ‘Is populism necessarily bad economics?’, *AEA Papers and Proceedings* **108**, 196–99.
- Rodrik, D. (2018b), ‘Is populism necessarily bad economics?’, *AEA Papers and Proceedings* **108**, 196–99.
URL: <https://www.aeaweb.org/articles?id=10.1257/pandp.20181122>
- Sasso, G. & Morelli, M. (2021), ‘Bureaucrats under Populism’, *Journal of Public Economics* **202**, 104497.

- Shayo, M. (2009), ‘A model of social identity with an application to political economy: Nation, class, and redistribution’, *American Political Science Review* **103**(2), 147–174.
- Sunstein, C. R. (2017), *Republic: Divided Democracy in the Age of Social Media*, hardcover edn, Princeton University Press, Princeton, NJ. Afterword by Cass R. Sunstein.
URL: <https://press.princeton.edu/books/hardcover/9780691175515/republic>
- Taber, C. S. & Lodge, M. (2006), ‘Motivated skepticism in the evaluation of political beliefs’, *American Journal of Political Science* **50**(3), 755–769.
URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-5907.2006.00214.x>
- Tausanovitch, C. & Vavreck, L. (2021), ‘Democracy Fund + UCLA Nationscape Project, October 10-17, 2019 (version 20211215)’.
URL: <https://www.voterstudygroup.org/data/nationscape>
- Urbinati, N. (2019), ‘Political theory of populism’, *Annual Review of Political Science* **22**, 111–127.
- Valgarðsson, V., Jennings, W., Stoker, G., Bunting, H., Devine, D., McKay, L. & Klassen, A. (2025), ‘A crisis of political trust? global trends in institutional trust from 1958 to 2019’, *British Journal of Political Science* **55**, e15.
- Van Weelden, R. (2013), ‘Candidates, credibility, and re-election incentives’, *Review of Economic Studies* **80**(4), 1622–1651.
- Vosoughi, S., Roy, D. & Aral, S. (2018), ‘The spread of true and false news online’, *Science* **359**(6380), 1146–1151.
URL: <https://www.science.org/doi/abs/10.1126/science.aap9559>
- Warshaw, C. & Rodden, J. (2012), ‘How should We measure district-level public opinion on individual issues?’, *Journal of Politics* **74**(1), 203–219.
- Zaslove, A. & Meijers, M. (2023), ‘Populist democrats? unpacking the relationship between populist and democratic attitudes at the citizen level’, *Political Studies* .
- Zhuravskaya, E., Petrova, M. & Enikolopov, R. (2020), ‘Political effects of the internet and social media’, *Annual Review of Economics* **12**(Volume 12, 2020), 415–438.
URL: <https://www.annualreviews.org/content/journals/10.1146/annurev-economics-081919-050239>

Appendix

A	Theory Proofs	A1
B	Survey Experiment	A5
C	District-level Estimates of Distrust in Government	A8
	C1 Comparing ANES with NationScape MRP Estimates	A11
D	Obtaining Candidates' Twitter Handles	A14
E	Measuring Commitment and Populist Rhetoric in Tweets	A15
F	Topic Detection	A21
G	Robustness Tests	A25
	G1 Turnout	A29

A Theory Proofs

Proof of Lemma 1. In the primary elections, members of party G vote for the candidate proposing the electoral platform that maximizes their expected utility, taking into account the probability of winning at the general election. We therefore have to compute the probability of a candidate winning at the general election. Consider a citizen v in party G who must choose between voting or not voting at the general elections, when the two candidates run with electoral platforms $\hat{S}$ and S' . Let us assume, without loss of generality, that platform $\hat{S}$ gives to citizen v in party G a higher utility from policy between the two electoral platforms: $EU^G(\hat{S}) > EU^G(S')$. Citizens choose their voting action as if pivotal. Then citizen v in party G votes for the candidate running with $\hat{S}$ if $EU^G(\hat{S}) - c_v^G - \mu^G > EU^G(S')$, otherwise she abstains. It follows that citizen v with cost of voting below $\underline{c}_v^G := EU^G(\hat{S}) - EU^G(S') - \mu^G$ votes (for $\hat{S}$).

Suppose $EU^G(\hat{S}) > EU^G(S')$ and $EU^{G'}(\hat{S}) < EU^{G'}(S')$.²⁸ For a given μ^G , the probability that the candidate running with $\hat{S}$ wins is

$$\begin{aligned} \mathbb{P}(\underline{c}_v^G > \underline{c}_v^{-G}) &= \mathbb{P}(\mu^{-G} > EU^{-G}(e') - EU^{-G}(e) - EU^G(e) + EU^G(e') + \mu^G) = \\ &= \frac{1}{2} + \psi(-EU^{G'}(S') + EU^{G'}(\hat{S}) + EU^G(\hat{S}) - EU^G(S') - \mu^G) \end{aligned} \quad (11)$$

that is, more voters show up for the candidate running with platform $\hat{S}$ than for the candidate running with platform S' in the general elections. Therefore, the expected probability of victory of candidate g at the time of the primary elections is

$$\frac{1}{2} + \psi(-EU^{G'}(S') + EU^{G'}(\hat{S}) + EU^G(\hat{S}) - EU^G(S')), \quad (12)$$

because μ^G has mean 0.

Consider now the primary elections of party G . Consider any two candidates of party G , one proposing $\hat{S}^g$ and the other one proposing a different platform $S^{g'}$. Notice that, if $\psi \rightarrow 0$ the probability of winning the general election for each of these two candidates, contingent upon having won the primaries, is the same. Therefore, members of party G at the primary elections vote for the candidate who maximizes their expected utility, taking into account the future cost of voting at the general elections. At the time of the primary elections, members of party G have the same preferences because the only heterogeneity comes from the realization of the cost of voting c_v^G at the general elections. Since each citizen at the general elections votes “as if” they were pivotal, it immediately follows that members of party G at the primary elections vote for the candidate proposing

²⁸The case in which the electoral platform $\hat{S}$ is preferred by all citizens is trivial.

the platform $\hat{S}^g$ that maximizes the utility.²⁹ Therefore, the candidate proposing $\hat{S}^g$ wins the primary elections. Competition at the primary elections guarantees that there exists at least one candidate proposing $\hat{S}^g$.

The argument immediately extends if ψ is positive and below a given threshold $\underline{\psi}$, as long as the expected gain in the probability of winning the election by reducing the turnout of the opposite party is lower than the welfare loss of having a candidate who proposes a suboptimal platform. ■

Proof of Proposition 1: When $\bar{m}_d \leq q_d W_d^G$, then p_d^G and π_d^G are equal to zero and hence discretion is always chosen by party G . When $\bar{m}_d \in (q_d W_d^G, W_d^G]$, $\pi_d^G = 0$, hence (4) becomes

$$p_d^G = \frac{\bar{m}_d - q_d W_d^G}{\bar{m}_d} \geq \frac{(1 - \gamma_d^G)}{\gamma_d^G \lambda_d + (1 - \gamma_d^G)(1 - \lambda_d)}.$$

In this range commitment is chosen therefore when q_d is sufficiently small, since the left hand side goes to 1 when q_d goes to zero and the right hand side is less than one.

When $\bar{m}_d > W_d^G$ then $\pi_d^G > 0$, but since it does not depend on q_d the impact of changes in q_d on the probability of demand of commitment is unchanged.

We can rewrite (4) as follows:

$$\bar{m}_d(1 - \gamma_d^G) < (\bar{m}_d - q_d W_d^G)[\gamma_d^G \lambda_d^G + (1 - \gamma_d^G)(1 - \lambda_d^G)] - (\bar{m}_d - W_d^G)\lambda_d^G(2\gamma_d^G - 1)$$

which implies

$$q_d < \frac{\lambda_d^G(2\gamma_d^G - 1)}{[\gamma_d^G \lambda_d^G + (1 - \gamma_d^G)(1 - \lambda_d^G)]} \equiv \bar{q}_{d|r}^G.$$

The RHS is positive whenever $\gamma_d^G > 1/2$. If $\gamma_d^G \leq 1/2$, the commitment is on s and the last inequality becomes

$$q_d < \frac{(1 - \lambda_d^G)(1 - 2\gamma_d^G)}{[\gamma_d^G \lambda_d^G + (1 - \gamma_d^G)(1 - \lambda_d^G)]} \equiv \bar{q}_{d|s}^G.$$

In this case, the RHS is positive whenever $\gamma_d^G \leq 1/2$. ■

Proof of Remark 2. The derivative of the threshold $\bar{q}_d^G$ with respect to γ_d^G is $\frac{\lambda_d^G}{(\gamma_d^G \lambda_d^G + (1 - \gamma_d^G)(1 - \lambda_d^G))^2} > 0$, if $\gamma_d^G \geq \frac{1}{2}$. The derivative is equal to $-\frac{(1 - \lambda_d^G)}{(\gamma_d^G \lambda_d^G + (1 - \gamma_d^G)(1 - \lambda_d^G))^2} < 0$, if $\gamma_d^G < \frac{1}{2}$. The derivative of $\bar{q}_d^G$ with respect to λ_d^G is $\frac{(2\gamma_d^G - 1)(1 - \gamma_d^G)}{(\gamma_d^G \lambda_d^G + (1 - \gamma_d^G)(1 - \lambda_d^G))^2} > 0$, if $\gamma_d^G \geq \frac{1}{2}$. The derivative is equal to $-\frac{(1 - 2\gamma_d^G)}{\gamma_d^G / (\gamma_d^G \lambda_d^G + (1 - \gamma_d^G)(1 - \lambda_d^G))^2} < 0$, if $\gamma_d^G < \frac{1}{2}$. ■

²⁹Consider first a member v of party G who would not change her decision to vote or abstain depending on whether the candidate g proposes $\hat{S}^g$ vs $S^{g'}$. In this case v obviously prefers that $\hat{S}^g$ is the winning electoral platform at the primary elections. Consider then a member of party G for whom instead it could make a difference whether one or the other is selected to run in the general election in terms of turnout decision: $EU^G(\hat{S}^g)EU^G(S^{g'}) \geq c_v^G + \mu^G$ and $EU^G(S^{g'}) - EU^G(S^{-g}) < c_v^G + \mu^G$. Even for such a voter the ex-ante optimal platform candidate is obviously preferred.

Proof of Proposition 2.

Suppose that candidate g proposes commitment to r_d on dimension d and g' proposes discretion. Hence we are in the case $\gamma_d^G \geq \frac{1}{2}$. Turnout of party G positively depends on the payoff they receive from their own candidate's platform and negatively on the payoff they receive from the opponent's platform. q_d only affects the utility party G members receive when candidate g' is elected, which is equal to

$$-p_d^G [\gamma_d^G \lambda_d^G + (1 - \gamma_d^G)(1 - \lambda_d^G)] - (1 - p_d^G) [\gamma_d^G (1 - \gamma_d^{G'}) + \gamma_d^{G'} (1 - \gamma_d^G)].$$

The derivative of the latter expression with respect to p_d^G is equal to

$$\begin{aligned} & -\gamma_d^G \lambda_d^G - (1 - \gamma_d^G)(1 - \lambda_d^G) + \gamma_d^G (1 - \gamma_d^{G'}) + \gamma_d^{G'} (1 - \gamma_d^G) = \\ & (1 - \gamma_d^{G'} - \lambda_d^G) (2\gamma_d^G - 1). \end{aligned}$$

The latter expression is negative if $\gamma_d^G \geq \frac{1}{2}$ and $\lambda_d^G \geq 1 - \gamma_d^{G'}$. By the fact that $\frac{\partial p_d^G}{\partial q_d} \leq 0$, it follows that the utility of party G of having an uncommitted policy-maker decreases when q_d decreases, because the probability that an uncommitted politician is captured increases. It follows that the turnout of group G increases. Similarly, the turnout of party G' positively depends on the payoff they receive from their own candidate's platform and negatively on the payoff they receive from the opponent's platform. q_d only affects the utility party G' members receive when candidate g' is elected, which decreases with $p_d^{G'}$, hence it increases with q_d . It follows that the turnout of party G' decreases, if q_d decreases. ■

Proof of Proposition 3.

Suppose that candidate g proposes commitment to r_d on dimension d and g' proposes discretion. Hence we are in the case $\gamma_d^G \geq \frac{1}{2}$. The derivative of the turnout of party G with respect to γ_d^G is the difference in the derivatives of the utilities from policy by commitment of their own candidate, and delegation of the opponent:

$$1 - 2\lambda_d^G \pi_d^G + p_d^G (2\lambda_d^G - 1) + (1 - p_d^G) (1 - 2\gamma_d^G) \quad (13)$$

A lowerbound of this derivative can be set by equating $\pi_d^G = p_d^G$:

$$1 - 2\lambda_d^G p_d^G + p_d^G (2\lambda_d^G - 1) + (1 - p_d^G) (1 - 2\gamma_d^G) = (1 - p_d^G) (2 - 2\gamma_d^G), \quad (14)$$

which is positive. The turnout of party G' does not depend on γ_d^G , hence the probability of election of candidate g increases with γ_d^G . The derivative of the turnout of party G

with respect to λ_d^G is

$$(1 - 2\gamma_d^G)\pi_d^G + p_d^G(2\gamma_d^G - 1) = (2\gamma_d^G - 1)(p_d^G - \pi_d^G), \quad (15)$$

which is positive. The turnout of party G' does not depend on λ_d^G , hence the probability of election of candidate g increases with λ_d^G . ■

B Survey Experiment

- *Platform:* Lucid.
- *Dates:* February 13-25, 2024.
- *Consent:* Subjects read an IRB-approved consent form, then voluntarily consented to participate in a research study.
- *Refusal rate:* 0.25%.
- *Screeners:* Respondents failing the attention check and those who did not complete the survey were removed.
- *Sample size of valid respondents:* 1,825.
- *Pre-analysis plan:* An anonymous version of the pre-analysis plan is available at https://aspredicted.org/Q4T_Q9W. It will be made public once the paper is published. The pre-registered hypotheses are tested in the following locations:
 - Treated respondents display lower levels of trust: Table B.1
 - Treated respondents display stronger preferences for commitment and weaker preferences for delegation: Table 5

Structure of the survey: First, respondents were asked a set of demographic and political questions. These include their age, sex, race, religion, level of education, salary, party identification, Trump/Biden supporter, and turnout intention.

Second, respondents were randomly assigned to either treatment or control conditions with equal probability. Control respondents receive no information (i.e., pure control group), whereas treatment respondents receive the following vignette. The text is a summary of a Vox.com article that can be found at the following link <https://www.vox.com/2014/4/18/5624310/martin-gilens-testing-theories-of-american-politics-explained>

“Below you will see a piece about the influence of ordinary voters and economic elites in US politics written by Andrew Prokopin, senior politics correspondent at Vox.com. Please pay attention to the articles as you will later be asked questions about the content.

Study: Politicians listen to rich people, not you

Who really matters in our democracy — the general public, or wealthy elites? That’s the topic of a recent study by political scientists Martin Gilens of Princeton and Benjamin Page of Northwestern. The study’s gotten lots of

attention over the past year, because the authors conclude, basically, that the US is a corrupt oligarchy where ordinary voters barely matter. Or as they put it, “economic elites and organized interest groups play a substantial part in affecting public policy, but the general public has little or no independent influence.”

Average citizens only get what they want if economic elites or interest groups also want it. As more and more average citizens support an issue, they’re not any more likely to get what they want. That’s a shocking finding in a democracy. In contrast, as more and more economic elites and interest groups want a certain policy change, they do become more likely to get what they want. The preferences of economic elites and interest groups — especially economic elites — are each quite influential, when the preferences of the other two groups are held constant.”

Third, respondents were asked to report their level of political trust. To measure trust, we use the same ANES question we then use to create district-level measures of distrust as well as the same survey questions on trust in government institutions designed by the OECD (Nguyen et al. 2021). The wording of the question is the following:

- ANES: How often can you trust the federal government in Washington to do what is right? [Never, Sometimes, About half of the time, Most of the time, Always]
- OECD: On a scale of 0 to 10, where 0 is not at all and 10 is completely, how much do you trust each of the following? [Federal government, President, Congress, Courts and legal system, Political parties, News media, International organizations]

Finally, respondents were asked to record their level of agreement with four different statements capturing their preferences for commitment and delegation. The wording of the statements is the following:

Commitment

- Politicians should publicly commit to a clear set of policies before they enter office.
- Politicians should try as hard as they can to implement their policy commitments once in office.

Delegation

- Politicians should receive free mandate to adjust policies that best address future challenges.
- Politicians should have the freedom to exercise their best judgment when confronted with new information once in office.

We pre-registered two different hypotheses. First, treated respondents were expected to display lower levels of trust. Second, treated respondents were expected to display

stronger preferences for commitment and weaker preferences for delegation. While the second test directly tests our theoretical prediction on the demand of commitment, the first hypothesis serves as a preliminary check that the treatment indeed affected political trust.

As described in the pre-registration report, we estimate the effect of the distrust treatment on trust and commitment/delegation with OLS as per Equation 10. The set of pre-treatment covariates include race, sex, education, religiosity, age category, salary, party identification, whether the respondent is a Trump or Biden supporter (both dummies), and turnout intention. We report heteroskedasticity-robust standard errors.

In Table B.1, we report the treatment effects on reported levels of trust towards a wide set of political institutions. Though to different levels of precision, treated respondents report lower levels of trust across all the proposed measures. The estimates achieve statistical significance at 90% or 95% level for the dichotomous version of the ANES question (Col. 1) and for measures of trust in Congress, political parties, and international organizations. These results are quite remarkable considering how hard it is to experimentally manipulate political trust.

Table B.1: Survey Experiment: Trust Results.

	ANES		OECD Trust Questions [0,10]						
	Dummy (1)	Continuous (2)	Fed. Govt. (3)	President (4)	Congress (5)	Courts (6)	Parties (7)	News (8)	Int. Orgs. (9)
Treatment	-0.04** (0.02)	-0.03 (0.05)	-0.12 (0.12)	-0.17 (0.13)	-0.22* (0.13)	-0.14 (0.13)	-0.25** (0.12)	-0.20 (0.13)	-0.32*** (0.12)
Covariates	✓	✓	✓	✓	✓	✓	✓	✓	✓
Mean DV	0.35	2.96	5.11	5.02	4.54	5.27	4.06	4.34	4.55
R ²	0.25	0.28	0.21	0.40	0.18	0.15	0.22	0.29	0.28
Observations	1,719	1,719	1,715	1,713	1,712	1,713	1,712	1,712	1,711

Notes: OLS estimates. Heteroskedasticity-robust SE in parenthesis. Outcomes are reported levels of trust. Cols. 1-2: ANES Question: ‘How often do you trust the government in Washington to do what is right?’, for the dummy measure (Col. 1), trust = 1 if respondents answer ‘Always’ or ‘Most of the times’. For the continuous score (Col. 2), the answers turned into a 1-5 scale, where 1 signifies ‘Never’ and 5 ‘Always’. Columns 3-9 report trust measures in several institutions. The precise wording of the question is ‘On a scale of 0 to 10, where 0 is not at all and 10 is completely, how much do you trust each of the following?’. Individual-level pre-treatment covariates include: race, sex, education, religiosity, age category, salary, party identification, Trump and Biden supporter, and turnout intention. Signif. codes: ***: 0.01, **: 0.05, *: 0.1

C District-level Estimates of Distrust in Government

To build district-level estimates of distrust, we first download survey data from the ANES Data Center. ANES releases surveys every 4 years. We focus on the last three waves (i.e., 2012, 2016, 2020) so that we can match the estimates with the time coverage of the Twitter activity of candidates (2012-2021).

To measure distrust we focus on the following survey question: “How often can you trust the federal government in Washington to do what is right?”. The possible answers are reported below. We dichotomize the variable so that distrust = 1 if the respondent *never* trusts the government. This allows us to estimate the share of distrustful respondents.

- -9. Refused
- -8. Don’t know
- 1. Always
- 2. Most of the time
- 3. About half the time
- 4. Some of the time
- **5. Never**

In 2012 some respondents were asked an additional question with a slightly different wording (“How much of the time do you think you can trust the government in Washington to do what is right?”) with 4 instead of 5 possible answers (“About half of the time” was not included as a possible answer). Respondents were asked either the 5-scale or the 4-scale question. In both cases, we code distrust = 1 if the answer is *never*. Consistent with common practice, we code as 0 respondents who answered “Don’t know” or who refused to answer.

This question is particularly suitable for the test we aim to perform. First, it explicitly asks respondents about *trust*, without resorting to other proxies or other similar but distinct concepts like confidence of competence. Second, it refers to the federal government in Washington, hence respondents are prompted to think about the same entity when answering the question. In fact, if asked about trust in politicians or trust in government in general, respondents might think about their state or local government, thus undermining the comparability of the answers. Clearly, we do not need to assume no interference between respondents’ reported trust in federal versus local government. We take the response at face value and we are aware that the level of distrust towards the government in Washington can be affected by local-level factors.

As for the choice of individual-level predictors to include both in the multilevel model and the post-stratification steps, we select sex, age, and education categories, for they

are at the same time good predictors of distrust and the only variables for which we know the joint distribution at the district level from the U.S. Census. These are also the individual-level predictors commonly used in other MRP applications.

To facilitate greater pooling across districts, we include in the model several district-level variables that are plausibly correlated with distrust. We include the share of black population, the Gini index, median income, employment rate, and the share of votes for the Republican presidential candidate in the same year at the state level.

The strength of the individual level predictors is evidenced by the regression estimates reported below. We fit linear probability models following different specifications to show that the individual level variables chosen for the multilevel model are good predictors of distrust and are robust to a series of different specifications.

Table C.2: Individual Predictors of Distrust.

	Distrust			
	(1)	(2)	(3)	(4)
Constant	0.210*** (0.012)			
Sex Category [1-2]	-0.019*** (0.004)	-0.020*** (0.005)	-0.021*** (0.005)	-0.021*** (0.005)
Age [1-5]	-0.015*** (0.002)	-0.016*** (0.002)	-0.016*** (0.002)	-0.018*** (0.002)
<i>District-level Covariates</i>				
Education [1-5]	-0.011*** (0.002)	-0.010*** (0.002)	-0.013*** (0.002)	-0.013*** (0.002)
% Black Population			-0.316** (0.125)	
Gini Index			0.369 (0.441)	
Median Income			0.006 (0.012)	
Employment Rate			0.303 (0.293)	
% of Republican Votes (State level)			0.004 (0.118)	
R ²	0.007	0.036	0.044	0.093
Observations	17,547	17,547	17,547	17,547
RMSE	0.287	0.283	0.282	0.274
District FE		✓	✓	
Wave FE			✓	
District × Year FE				✓

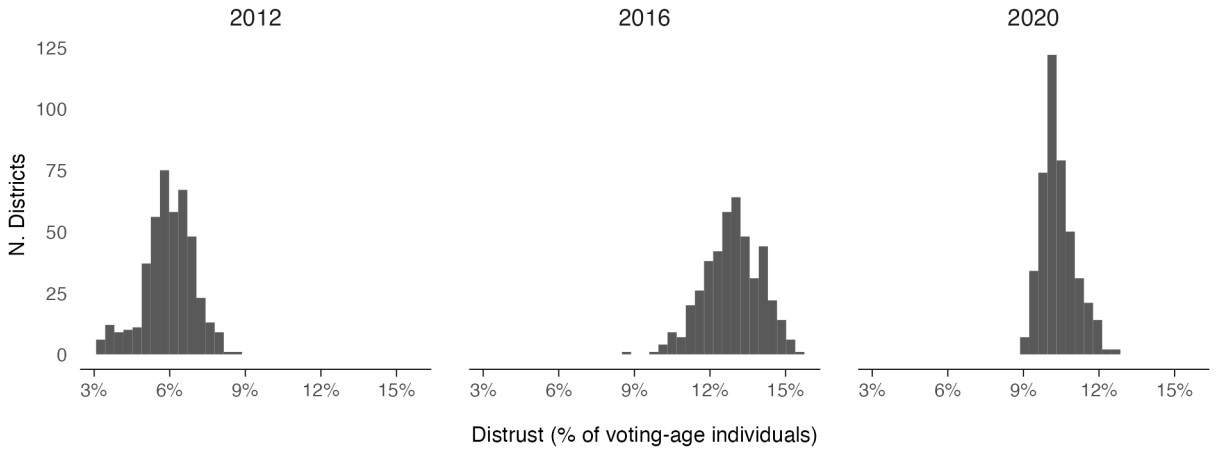
Notes: OLS estimates. SE clustered by congressional district. DV is a dummy which equals 1 when the respondent said they never trust the government in Washington to do what is right.

We then perform the MRP in two steps. First, we fit a multilevel model to the survey data including a set of random effects for J education categories, K sex categories, M age categories, and N districts. X is a matrix of predictors that vary at the district level. We can imagine U.S. citizens as consisting of these different ideal types.

$$\begin{aligned}
Pr(Distrust_i = 1) &= \Phi(\beta_0 + \alpha_{j[i]}^{education} + \alpha_{k[i]}^{sex} + \alpha_{m[i]}^{age} + \alpha_{n[i]}^{district} + \mathbf{X}_{n[i]}^T \beta) \\
\alpha_{j[i]}^{education} &\sim N(0, \sigma_{education}^2), \text{ for } j = 1, \dots, J \\
\alpha_{k[i]}^{sex} &\sim N(0, \sigma_{sex}^2), \text{ for } k = 1, \dots, K \\
\alpha_{m[i]}^{age} &\sim N(0, \sigma_{age}^2), \text{ for } m = 1, \dots, M \\
\alpha_{n[i]}^{district} &\sim N(0, \sigma_{district}^2), \text{ for } n = 1, \dots, N
\end{aligned} \tag{16}$$

In the second step, we calculate the predicted share of distrustful individuals for each ideal type based on Equation 16, which we denote $\hat{\pi}_{jkmn}$. We then use post-stratification and calculate the weighted average of $\hat{\pi}_{jkmn}$ based on the *actual* prevalence of that ideal type in the population of the district. Since the predictions are not linear in the random effects, we need census data on the joint distribution of education, sex, and age in each district. These steps are performed semi-automatically through the **AutoMrP** package in R, which uses an ensemble modeling approach that provides better small-area estimates (Broniecki et al. 2022). We produce estimates of distrust for 435 congressional districts (plus the District of Columbia) over three waves, estimated separately for each year. In Figure C.1 we report the distribution of our estimates of distrust over time.

Figure C.1: Distribution of distrust estimates over time.



Notes: Distribution of estimates of distrust in government for each congressional district over the three waves of survey data.

C1 Comparing ANES with NationScape MRP Estimates

To validate our MRP estimates, we show that the district level estimates that can be obtained from the relatively small ANES sample are comparable to those obtained from much larger samples.

We replicate the MRP method for the year 2020 with survey data from NationScape (Tausanovitch & Vavreck 2021), an online, weekly survey with approximately 10,000 interviews per week conducted in the period 2019-2021. We keep a large random sample of 50,000 responses recorded before the 2020 election day for which there is information on the congressional district of the respondent, their sex, age, and level education. NationScape does not include questions on trust in government, therefore we cannot directly compare the estimates of distrust used in the analysis. However, we can compare the estimates for two other political questions, one on the perceived state of the economy and one on political interest, which are included both in ANES and NationScape. Table C.3 reports the question prompts with the possible answers.

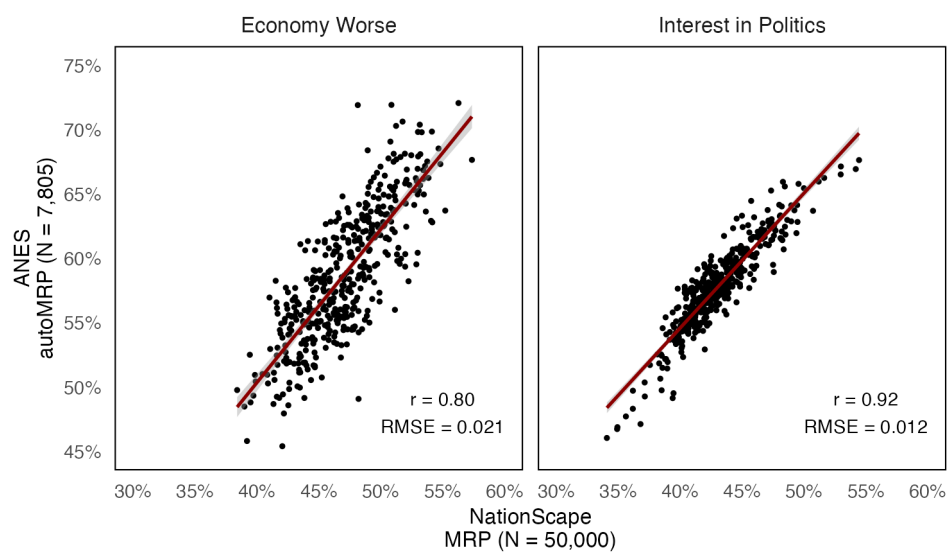
Table C.3: Survey questions, ANES and NationScape.

ANES	NationScape
Now thinking about the economy in the country as a whole, would you say that over the past year the nation's economy has gotten better, stayed about the same, or gotten worse?	Would you say that as compared to one year ago, the nation's economy is now better, about the same, or worse?
• -9. Refused • -8. Don't know • 1. Gotten better • 2. Stayed about the same • 3. Gotten worse	• 1. Better • 2. About the same • 3. Worse • . Respondent Skipped
Sample mean: .61	Sample mean: .48
How often do you pay attention to what's going on in government and politics?	Some people follow what's going on in government most of the time, whether there's an election going on or not. Others aren't that interested. Would you say you follow what's going on...
• -9. Refused • 1. Always • 2. Most of the time • 3. About half the time • 4. Some of the time • 5. Never	• 1. Most of the time • 2. Some of the time • 3. Only now and then • 4. Hardly at all • . Respondent Skipped
Sample mean: .63	Sample mean: .43
Respondents: 7,805	Respondents: 414,318

Notes: Question wording and possible answers in ANES and NationScape surveys. Answers re-coded to obtain dichotomous responses, values in bold equal to 1 and 0 otherwise.

As shown in the table, the average responses are quite different in the two surveys. However, if the size of the ANES survey is large enough, we should observe a strong correlation between the estimates derived from the large NationScape sample and the ANES sample. Figure C.2 shows there is a positive and strong correlation between the estimates. ANES estimates are overall larger than NationScape estimates, consistently with the larger sample means in the ANES sample of respondents, but the correlation is very high, equal or greater than .8 for both variables. Similarly, the standard error of a linear regression of the NationScape estimates on the ANES estimates is very low. This test suggests that the ANES sample yields estimates comparable to those obtained from a sample size $\times 7$ larger.

Figure C.2: MRP Estimates from ANES and NationScape



Notes: On the horizontal axis, estimated share of NationScape respondents who believe the economy in the past year has worsened (panel on the left) and who follow what is going on in government most of the times. On the vertical axis, estimated share of ANES respondents for comparable questions.

D Obtaining Candidates' Twitter Handles

We scraped the Twitter handle of candidates from several sources: Ballotpedia, an online repository on U.S. elections and candidates; the House of Representatives Press Gallery;³⁰ and Politwoops, an organisation that tracks deleted tweets by public officials and maintains a list of active Twitter accounts.³¹ For the candidates not matched in these sources, we program a scraper to directly obtain the Twitter account through the search engine of the official Twitter website.

³⁰Twitter handle for member of the 117th Congress available at <https://pressgallery.house.gov/member-data/members-official-twitter-handles>.

³¹Dataset available at <https://www.propublica.org/datastore/dataset/politicians-tracked-by-politwoops>.

E Measuring Commitment and Populist Rhetoric in Tweets

For any natural language processing task, we need a numerical representation of text. Here, we build on recent advancements in deep learning and use a language model pre-trained on a large corpus of tweets that is suitable for downstream NLP tasks with social media text (Barbieri et al. 2020). The model takes as input the text of the tweet – minimally preprocessed – and produces a fixed-length, dense vector which encodes its meaning and semantic properties in a numerical form. As a preprocessing step, we simply remove urls and replace account names with a general “@user”. Producing an embedding for 5.9 million tweets would consume a significant amount of computational resources and may take considerable time. Therefore, we produce vector representation of tweets posted in the three election years we studied (i.e., 2012, 2016, 2020).

Once we have an embedding for every tweet, we assemble the training dataset on which to train the classifier. For a classifier to make good predictions, the classes in which documents are labelled by the annotators (i.e., the same categories predicted by the classifier) need to be balanced, namely we should have a similar number of texts in category A and category B (for a 2-class classification task). Classifiers learn better from a balanced distribution of classes in a dataset, hence we ensure balance with the most common technique: oversampling texts in the minority category. Clearly, populist rhetoric and policy commitment tweets are the two minority classes.

To ensure class balance, we first compute the embedding for a “commitment” and “populist rhetoric” poles, given by the average embedding of 30 examples of tweets containing policy commitment and populist rhetoric generated by GPT. We gave GPT two prompts, one for populist rhetoric and one for policy commitments. For the populist rhetoric prompt, we simply asked “Write me 30 examples of tweets posted by politicians using anti-elite, populist rhetoric.” and we found it to work very well, for populist rhetoric is clear as a concept (although it might be harder to detect it in real text). Conversely, for the commitment prompt we were more specific, for commitment can be understood in multiple ways, not necessarily with regard to a specific policy or action. Hence, we prompted GPT with the following command: “Candidates often commit to specific policies on social media, outlining the proposals or actions that they intend to take. Write 30 examples of tweets containing a commitment to a specific policy or a clear promise of action.”

The texts produced by GPT are included in the replication files. In Table E.4 below we report a subset due to space limitations.

Table E.4: Texts generated by GPT.

Sentences generated by ChatGPT
<i>Policy Commitments</i>
Mental health is as important as physical health. If elected, I will push for increased access to mental health services and work to remove the stigma surrounding mental illness. #MentalHealthMatters
Small businesses are the backbone of our economy. If elected, I will advocate for policies that promote entrepreneurship and provide support to local businesses. #SupportSmallBusiness
Education is the foundation of a strong society. I pledge to prioritize funding for public schools and ensure every child has access to quality education. #EducationMatters
Our veterans deserve the utmost respect and care. I will work tirelessly to improve healthcare services and provide better support for our brave servicemen and women. #SupportOurTroops
Equal pay for equal work should be a reality. I will push for legislation that eliminates the gender pay gap and ensures fair wages for all. #EqualPay
<i>Populist Rhetoric</i>
The wealthy and well-connected have too much influence in our government. It's time to level the playing field and make sure everyone has an equal voice.
The elite think they can buy their way into power, but we know that true democracy can't be bought. Let's take back our government and make it work for all of us!
The political class has become disconnected from the realities of everyday life. We need representatives who understand and fight for the concerns of ordinary citizens.
The political class has been making decisions based on their own self-interest, but it's time to demand a government that is accountable to the people. Let's make our voices heard!
The political establishment is more interested in serving themselves than the American people. We need a government that is accountable to its citizens, not just the powerful.

Notes: Subset of the statements generated by ChatGPT.

We estimate embeddings for each of the populist-rhetoric and commitment texts generated by GPT and we compute the average for each class, which will represent the two poles against which oversampling tweets. We then estimate the cosine similarity between each vector representation of the tweets and the two poles. We extract the 1,500 tweets with the largest cosine similarity to each poles and other 1,500 tweets at random from the remaining tweets to obtain a training datasets of 3,000 tweets per classification task. We then ask GPT to annotate each of the 3,000 tweets based on the prompt reported in Table E.5. To ensure the replicability of the annotation, we set the temperature of the model — a parameter that governs the randomness and thus the creativity of the responses — to 0, which means that the responses will be almost deterministic, yielding the same response over and over again to a given prompt. We leave all other parameters to their default settings.

One key advantage of using GPT for classifying populist rhetoric and policy commitments is that, while GPT clearly suffers from the same biases of the humans who produced the text on which it has been trained, it is likely to give more stable answer

to what are generally difficult questions. Detecting policy commitments and populist rhetoric in tweets is hard, and human annotators might suffer attention decays. Moreover, because tweets can never be fully anonymous, it is likely that coders will adopt a “double-standard” classification when reading tweets (clearly) posted by Democrats or Republicans, displaying a form of confirmation bias. Relying on GPT, besides being a cheap and efficient way of annotating text, also allows to limit human arbitrariness in the classification process.

Table E.5: GPT prompt and answer example.

Prompt	Answer
I will show you a tweet posted by a candidate to the U.S. Congress. Candidates often commit to specific policies on social media, outlining the proposals or actions that they intend to take. Does the tweet contain a commitment to a specific policy or a clear promise of action? Reply with yes or no and give a short justification of your response. This is the tweet: <i>“I am committed to protecting our children, stopping human trafficking, and getting these heinous offenders off our streets. I support the #PROTECTAct.”</i>	Yes. The tweet contains a commitment to a specific policy, which is the #PROTECTAct. The candidate is stating their support for the act and their commitment to protecting children and stopping human trafficking.
I will show you a tweet posted by a candidate to the U.S. Congress. Candidates sometimes use populist rhetoric in their tweets, appealing to the concerns and grievances of the common people against a perceived elite or establishment. Tweets with populist rhetoric therefore either praise the people or attack the elites. Does the tweet contain populist rhetoric? Reply with yes or no and give a short justification of your response. This is the tweet: <i>“We need an America that puts working people first! Not parties. Not politicians. Not corporations. It’s time for someone who works for all the people. It’s time for @Nate_McMurray for Congress – NY27 – Leadership matters! SPECIAL ELECTION June 23”.</i>	Yes, the tweet contains populist rhetoric. The tweet appeals to the concerns and grievances of the common people against the perceived elite or establishment by attacking politicians and corporations and emphasizing the need for someone who works for all the people. The use of the phrase “working people” also suggests a focus on the common people.

Notes: Examples of prompts fed to GPT through the OpenAI API and the response given by the model.

To compare the validity of GPT’s annotation, we extract two random samples of 100 tweets each and ask an independent coder to replicate the annotation task. We give the coder the same prompt given to GPT (see Table E.5). The coder and GPT agree 82% (commitment) and 84% (populist rhetoric) of the times, which suggests that we can confidently use GPT to train the classifier.

On the annotated dataset, we then train 4 different classifiers, available through the `sklearn` library in Python:

1. Gradient Boosting Classifier (GB)
2. AdaBoost Classifier (ABC)
3. Random Forest (RF)

4. Linear Discriminant Analysis (LDA)

We set the model parameters to their default setting. We train the model on 75% of the tweets in the training dataset, and we leave the remaining 25% as a held-out test set on which to assess the performance of the classifier.

In Table E.6 below, we report the performance metrics of each of the four classifiers. While all classifiers work well on the held-out test set, we select the Gradient Boosting classifier, for it is the one performing best. We then train the model on the entire training dataset (this time including the test set too). Finally, we apply the classifier on the universe of the tweets posted in 2012, 2016, and 2020.

Table E.6: Performance metrics of classifiers.

Classifier	Commitment				Populist Rhetoric			
	F1	Accuracy	Recall	Precision	F1	Accuracy	Recall	Precision
GB	0.875	0.876	0.886	0.865	0.886	0.868	0.908	0.865
RF	0.858	0.857	0.878	0.839	0.880	0.860	0.910	0.852
ABC	0.868	0.869	0.878	0.859	0.835	0.812	0.842	0.828
LDA	0.850	0.848	0.878	0.824	0.833	0.808	0.851	0.816

Notes: Performance metrics for each classifier used on two separate 2-classes training datasets: one for policy commitments and one for populist rhetoric. Performance metrics calculated from comparing the model predictions to the annotated labels of held-out test sets (25% of the tweets in a 3,000-tweet training dataset).

Finally, in Table E.7 and Table E.8 we report representative tweets with the high and medium predicted probability of belonging to the commitment and populist-rhetoric classes. We display tweets with medium probability to showcase the performance of the classifier, which is able to distinguish tweets with reference to policy (medium probability) from tweets with reference to policy *commitments* (high probability). Similarly, for populist-rhetoric tweets, the classifier is able to separate tweets where there is a general praise of the people (medium probability) from those where people and elites are portrayed as antagonistic groups (high probability).

Table E.7: Examples of commitment tweets.

Tweet	Party	Predicted Probability
<i>Policy Commitments - High Probability</i>		
Medicare for All. Universal healthcare will ensure that all of us get the care we need.	D	0.976
FUND. THE. POLICE. I just introduced the Commitment to American Security Act with @RepMikeRogersAL, which will increase funding for local police for better training and 500,000 new body cameras.	R	0.943
My top priority will be ending illegal immigration. I will support the President in building a wall, and push immigration policies that put the American worker before foreign workers. If you're tired of Americans being put last, donate to my campaign! https://t.co/k2WhFkJuPM	R	0.983
Every LGBTQ North Texan deserves to have protection under our civil rights laws. One year ago, I voted to pass the #EqualityAct. Today on #IDAHOTB, we must recommit to the principle that no matter who you are or who you love, everyone gets equal protection against discrimination. https://t.co/Yt8JxCn2NW	D	0.971
<i>Policy Commitments - Medium Probability</i>		
People do not go to sleep at night worried about environmental justice. They're worried about feeding their families, having safe streets for their kids to play in and quality education. The congressman's priorities are all wrong.	R	0.502
The term "'privatizing'" has a bad connotation because gambling with the financial security and well-being of seniors is a terrible idea. https://t.co/hZOcfPFpdB	D	0.539
Hundreds of Iowans still haven't received the economic impact payments promised in the CARES Act pandemic relief package that we passed nearly six months ago. It's unacceptable, and today I wrote @stevenmnuchin1 demanding answers. #IA01 https://t.co/2fR3pMHw6H	D	0.485
We must speak openly as we rebuild the relationship between police and their communities @CNN #charlotte https://t.co/aJPJyXrMHC	R	0.555

Notes: Representative tweets with high and medium predicted probability of belonging to the commitment class.

Table E.8: Examples of populist-rhetoric tweets.

Tweet	Party	Predicted Probability
<i>Populist Rhetoric - High Probability</i>		
If only Congress @SpeakerPelosi would condemn violent funded and organized Marxist BLM/Antifa, who use sticks and stones to break bones of innocent people, yet apparently now it's words that don't hurt anyone that Congress condemns. Protect #FreeSpeech NOT domestic terrorists! https://t.co/DgOpZ1HXy2	R	0.976
President @realDonaldTrump has to go up against: -The Chinese virus -The Fake News -Pelosi & the radical left -Soros, Bloomberg & other activist billionaires -The Establishment -Hollywood -ANTIFA/BLM -Corrupt Investigators And yet he's still going to win. TRUMP 2020!	R	0.977
For every progressive policy proposal that continues to stall in Congress despite being popular, there are special interests spending millions to stop our progress. We need to get PACs and lobbyists out of our elections. https://t.co/KNlWyxdS9P	D	0.977
Americans are sick of pharmaceutical companies using their \$\$\$ to influence government policy to their own financial advantage.	D	0.984
<i>Populist Rhetoric - Medium Probability</i>		
Our young people deserve to have a future worth believing in.	D	0.603
We all deserve to thrive, not just survive. The census provides a fair say in choosing our leaders and the power to build the world around us. Complete yours at https://t.co/lmmejhZh5h #WhyCensusMatters #2020Census	D	0.633
Now they're talking about impeaching new Supreme Court Justice Amy Coney Barrett. This insanity needs to stop! https://t.co/MqTBcCGV5W	R	0.421
There's a clear difference in this election. I understand how important Texas energy development is to America's economy, job creation, national security, & funding quality public education – while Candace Valenzuela wants to decimate these jobs & end American energy independence https://t.co/SktFLUg8hc	R	0.571

Notes: Representative tweets with high and medium predicted probability of belonging to the populist-rhetoric class.

F Topic Detection

Table F.9: Topics and questions.

Inferred Topic	CCES Question	Years	N. Times
Abortion	abortion_20weeks	2020	1
	abortion_always	2020	1
	abortion_coverage	2020	1
	abortion_prohibition	2016, 2020	23
Environment	enviro_airwateracts	2016, 2020	2
	enviro_carbon	2016, 2020	21
	enviro_mpg_raise	2020	1
	enviro_renewable	2020	2
Gun Regulation	guns_assaultban	2020	2
	guns_bgchecks	2016	59
Healthcare	healthcare_aca	2012, 2016, 2020	43
	healthcare_medicare	2020	2
Immigration	immig_border	2012, 2020	45
	immig_employer	2012, 2016	6
	immig_legalize	2012, 2020	4
	immig_police	2012	2
	immig_report	2020	1
	immig_services	2012	15
	immig_wall	2020	5
Military	military_democracy	2012, 2016	21
	military_genocide	2020	1
	military_helpun	2012	1
	military_oil	2012, 2016, 2020	12
	military_protectallies	2012, 2016, 2020	33
	military_terroristcamp	2012, 2016	3
Same-sex Marriage	gaymarriage_legalize	2012	13

Notes: Survey questions and inferred topic where partisan respondents expressed highest level of agreement (lowest standard deviation). The table reports the number of times and the years when the question appear as the one displaying lowest standard deviation in any state-party-year groups of respondents.

Table F.10: Most homogeneous topics across state-party-year groups of respondents.

State	Democrats			Republicans		
	2012	2016	2020	2012	2016	2020
AL	Immigration	Gun Reg.	Immigration	Same-sex Marr.	Healthcare	Immigration
AK	Healthcare	Environ.	Abortion, Immigration	Same-sex Marr.	Military	Immigration
AZ	Military	Gun Reg.	Environ.	Military	Healthcare	Immigration
AR	Military	Gun Reg.	Military	Same-sex Marr.	Healthcare	Immigration
CA	Immigration	Gun Reg.	Military	Military	Healthcare	Immigration
CO	Military	Gun Reg.	Immigration	Military	Healthcare	Immigration
CT	Healthcare	Gun Reg.	Abortion	Military	Gun Reg.	Immigration
DE	Military	Gun Reg.	Environ.	Military	Gun Reg.	Abortion
FL	Immigration	Gun Reg.	Environ.	Military	Gun Reg.	Immigration
GA	Immigration	Gun Reg.	Abortion, Immigration	Military	Healthcare	Immigration
HI	Immigration	Gun Reg.	Abortion	Military	Gun Reg.	Military
ID	Healthcare	Gun Reg.	Environ.	Same-sex Marr.	Military	Immigration
IL	Military	Gun Reg.	Environ.	Military	Gun Reg.	Immigration
IN	Military	Gun Reg.	Abortion	Military	Gun Reg.	Immigration
IA	Military	Gun Reg.	Abortion	Military	Healthcare	Immigration
KS	Healthcare	Gun Reg.	Environ.	Military	Healthcare	Immigration
KY	Military	Gun Reg.	Abortion	Same-sex Marr.	Healthcare	Immigration
LA	Immigration	Gun Reg.	Healthcare	Military	Healthcare	Immigration
ME	Same-sex Marr.	Abortion	Immigration	Military	Healthcare	Immigration
MD	Immigration	Gun Reg.	Abortion	Military	Gun Reg.	Immigration
MA	Immigration	Gun Reg.	Environ., Gun Reg.	Military	Military	Immigration
MI	Military	Gun Reg.	Environ.	Military	Healthcare	Immigration
MN	Military	Gun Reg.	Environ.	Military	Healthcare	Immigration
MS	Immigration	Gun Reg.	Healthcare	Same-sex Marr.	Gun Reg.	Immigration
MO	Healthcare	Gun Reg.	Immigration	Military	Healthcare	Immigration
MT	Military	Abortion	Abortion	Military	Healthcare	Immigration
NE	Military	Gun Reg.	Abortion	Military	Healthcare	Immigration
NV	Military	Gun Reg.	Abortion, Environ.	Immigration	Abortion	Immigration
NH	Immigration	Gun Reg.	Environ.	Immigration	Military	Immigration
NJ	Immigration	Gun Reg.	Environ.	Military	Gun Reg.	Immigration
NM	Military	Gun Reg.	Environ.	Military	Healthcare	Immigration
NY	Immigration	Gun Reg.	Environ.	Military	Gun Reg.	Immigration
NC	Immigration	Gun Reg.	Abortion	Same-sex Marr.	Healthcare	Immigration
ND	Military	Gun Reg.	Immigration	Healthcare	Healthcare	Immigration
OH	Immigration	Gun Reg.	Environ.	Immigration	Healthcare	Immigration
OK	Military	Gun Reg.	Abortion	Same-sex Marr.	Healthcare	Immigration
OR	Healthcare	Gun Reg.	Abortion	Military	Military	Immigration
PA	Military	Gun Reg.	Abortion	Military	Gun Reg.	Immigration
RI	Healthcare	Gun Reg.	Abortion	Military	Immigration	Military
SC	Military	Gun Reg.	Abortion	Same-sex Marr.	Gun Reg.	Immigration
SD	Healthcare	Gun Reg.	Environ.	Military	Immigration	Abortion
TN	Immigration	Gun Reg.	Environ.	Same-sex Marr.	Healthcare	Immigration
TX	Immigration	Gun Reg.	Environ.	Military	Healthcare	Immigration
UT	Military	Gun Reg.	Immigration	Same-sex Marr.	Healthcare	Military
VT	Healthcare	Gun Reg.	Environ.	Military	Immigration, Military	Abortion
VA	Immigration	Gun Reg.	Environ.	Military	Healthcare	Immigration
WA	Healthcare	Gun Reg.	Abortion	Military	Immigration	Immigration
WV	Military	Gun Reg.	Abortion	Military	Military	Immigration
WI	Healthcare	Gun Reg.	Abortion	Immigration	Healthcare	Immigration
WY	Healthcare	Military	Environ., Gun Reg., Healthcare	Healthcare	Environ., Healthcare	Immigration

Building Topic-specific Dictionaries

To build topic-specific dictionaries, we extract the 20 words whose vector representation is most similar to the topic label (e.g., immigration) with `word2vec` (Mikolov et al. 2013), an unsupervised algorithm that learns fixed-length feature representations from how often words co-occur with one another, with the assumption that the meaning of a word is given by the company it keeps (Rodriguez & Spirling 2022).

First, we pre-process the 5.9 million tweets posted by Congressional candidates between 2012 and 2021. We lower case the text, we remove the Twitter accounts tagged in the text (e.g., @NAME), we keep ash-tags because they generally convey important information, we remove punctuation, and we convert all nouns to their singular form (e.g., houses $\rightarrow$ house). We then allow bi-grams to emerge based on how often two words occur next to one another, imposing a minimum count of 200. This step allows for words like “gun_control” or “health_care” to be considered single words when training the model. We finally train the `word2vec` model on the pre-processed corpus of tweets with the `gensim` Python library, estimating 200-dimensional vectors, excluding words appearing less than 10 times, and setting a window size (where to compute word co-occurrences) to 4.

Once we have a word embedding for each word used at least ten times in the corpus, we extract the 20 words with the largest cosine similarity to the topic label. We manually remove false positives to ensure that each word is used almost exclusively in the context of the topic (e.g., we remove the word “government” from the military-specific dictionary, for it can be used in many different contexts without referring to military issues) and we assemble the eight topic-specific dictionaries reported in Table F.11 below.

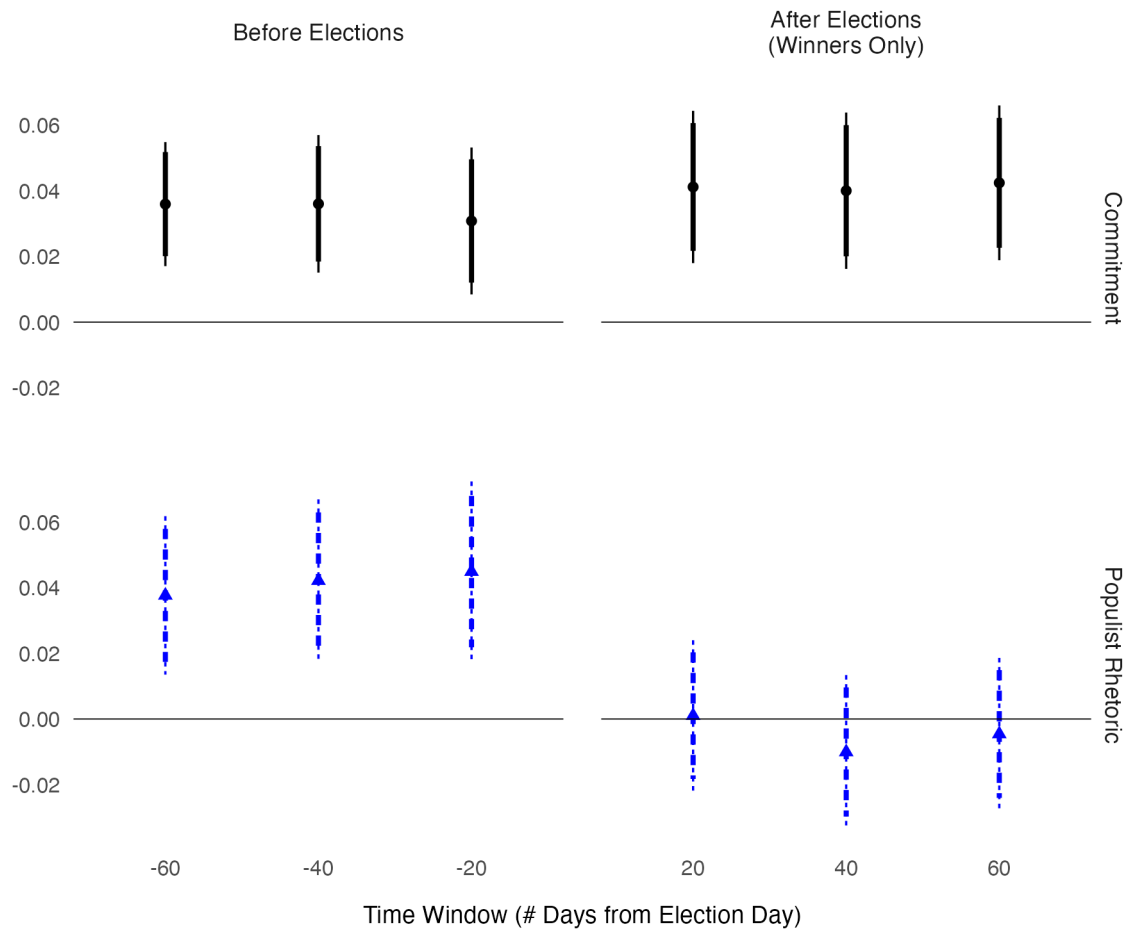
Table F.11: Topic specific dictionaries.

Topic	Dictionary
Gun Regulations	gun, firearm, handgun, assault_weapon, weapon_of_war, bump_stock, rifle, weapon, assault_rifle, domestic_abuser, silencer, shotgun, pistol, assaultweapons, weapon_ban, gunfreezones, semiautomatic, concealed_carry
Immigration	immigration, immigrant, asylum_seeker, undocumented_immigrant, refugee, asylumseekers, immigrant_and_refugee, legal_immigration, undocumented, immig, immigration_policy, legal_immigrant, illegal_immigration, migrant, deportation, imm, family_reunification, familybased, assimilation, illegalimmigration, sanctuarycities, chain_migration
Abortion	abortion, lateterm_abortion, infanticide, birth_control, family_planning, abortion_clinic, abortionist, legal_abortion, abortion_provider, unborn_baby, abortion_ban, contraceptive, planned_parenthood, reproductive_health, contraception, unborn_child, title_x, born_alive
Military	military, armed_force, USA_military, military_personnel, servicemembers, service_member, military_readiness, our_troop, uniform, armedforces, dod, warfighters, civilian, troop, navy, coast_guard, department_of_defense, national_defense, military_service, law_enforcement
Environment	environment, our_environment, ecosystem, our_planet, enviro, air_and_water, env, clean_air, airwater, climate, quality_of_life, cleanair, natural_resource, planet, biodiversity, wildlife, energy_sector, waterway, landscape
Health Care	healthcare, health_care, affordable_healthcare, hc, access_to_healthcare, quality_healthcare, heathcare, healthcare_coverage, health_coverage, health, health_insurance, reproductive_healthcare, insurance_coverage, retirement_security, hcare, child_care, universal_healthcare, preventative_care, quality_care, medicare
Same-sex Marriage	gay_marriage, gay, samesex_marriage, polygamy, lgbt, homosexual, lesbian, gaymarriage, marriage_equality, interracial, ssm, conversion_therapy, bisexual, homosexuality, samesex_couple, antigay, doma, samesexmarriage

Notes: Topic specific dictionaries derived from the `word2vec` model based on the largest cosine similarity between words and the topic label.

G Robustness Tests

Figure G.3: Effect of distrust across different time windows.



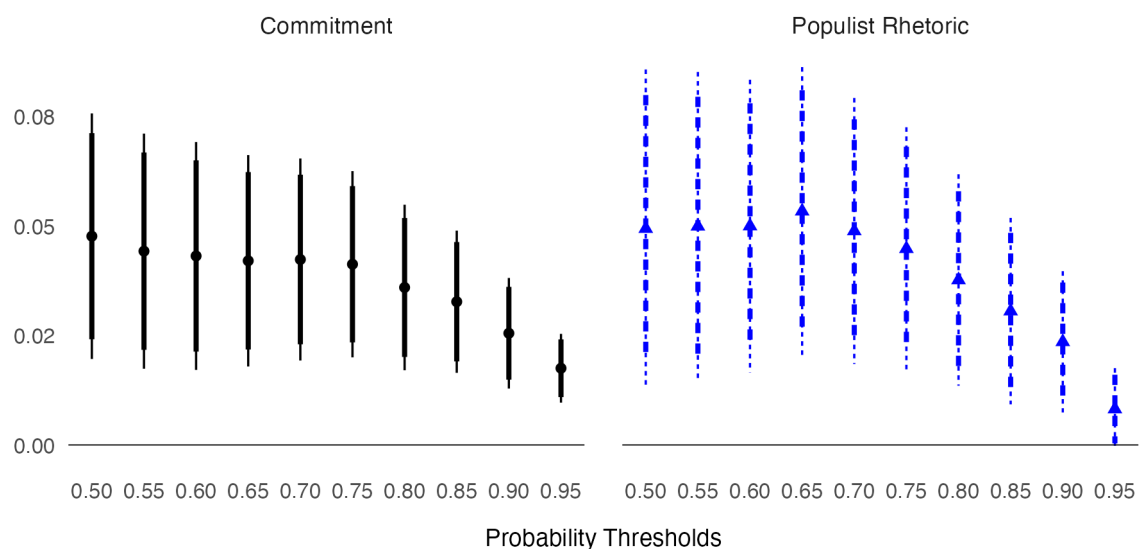
Notes: Estimated effect of distrust with 90 and 95% confidence intervals estimated across different windows of time expressed in days from election day. Specification includes candidate, district, state-by-year fixed effects. Candidate covariates: incumbent status. District covariates: gini index, share of black individuals, employment rate, median income (log). Standard errors clustered by district.

Table G.12: Distrust, supply of commitments and populist rhetoric. Continuous outcomes.

	Commitment			Populist Rhetoric		
	(1)	(2)	(3)	(4)	(5)	(6)
Distrust	0.017** (0.008)	0.021** (0.010)	0.037*** (0.012)	0.024** (0.011)	0.021 (0.013)	0.043*** (0.015)
Controls: Candidate	✓	✓	✓	✓	✓	✓
Controls: District	✓	✓	✓	✓	✓	✓
Mean DV	0.20	0.20	0.20	0.32	0.32	0.32
R ²	0.118	0.155	0.158	0.149	0.204	0.206
Observations	274,253	274,253	274,253	274,253	274,253	274,253
District FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓		✓	✓	
Candidate FE		✓	✓		✓	✓
State-Year FE			✓			✓

Notes: OLS estimates. SE clustered by district. DV is a predicted probability of commitment/populist tweets. Distrust is the standardized share of distrustful individuals in each district. Candidate covariates: party (only in columns 1 and 4, when candidate fixed effects are not included) and incumbent status. District covariates: gini index, share of black individuals, employment rate, median income (log). Signif. codes: ***: 0.01, **: 0.05, *: 0.1.

Figure G.4: Effect of distrust using different measures of commitment and populist rhetoric.



Notes: Estimated effect of distrust with 90 and 95% confidence intervals with different probability thresholds used to dichotomize the outcome (reported on the horizontal axis). Specification includes candidate, district, state-by-year fixed effects, incumbency status. Candidate covariates: incumbent status. District covariates: gini index, share of black individuals, employment rate, median income (log). Analysis performed on the sample of tweets posted in the 60 days prior to the election as in the main specification.

Table G.13: Alternative measures of distrust.

	Commitment			Populist Rhetoric		
	(1)	(2)	(3)	(4)	(5)	(6)
Distrust - Govt. Do Right	0.037*** (0.010)			0.039*** (0.013)		
Distrust - Big Interests		0.006 (0.004)			0.012* (0.006)	
Distrust - Politicians Corrupt			0.020** (0.009)			0.022* (0.012)
Controls: Candidate	✓	✓	✓	✓	✓	✓
Controls: District	✓	✓	✓	✓	✓	✓
R ²	0.069	0.069	0.069	0.121	0.121	0.121
Observations	274,253	274,253	274,253	274,253	274,253	274,253
Candidate FE	✓	✓	✓	✓	✓	✓
District FE	✓	✓	✓	✓	✓	✓
State-Year FE	✓	✓	✓	✓	✓	✓

Notes: OLS estimates. SE clustered by district. DV is a dummy measure for commitment/populist tweets. Distrust variables are standardized shares of distrustful individuals in each district. Candidate covariates: party (only in columns 1 and 4, when candidate fixed effects are not included) and incumbent status. District covariates: gini index, share of black individuals, employment rate, median income (log). Govt. Do Right: never trust the government in Washington to do what is right. Politicians Corrupt: Most/All in government are corrupt. Big Interests: Government run by a few big interests. Signif. codes: ***: 0.01, **: 0.05, *: 0.1.

G1 Turnout

To confirm the descriptive evidence from Figure 3, we estimate the effect of the triple interaction between the Trump periods (i.e., election year dummies), self-identifying as a Republican, and reporting distrust in the federal government. In particular, we estimate the “Trump” effect on the probability of turnout among distrustful Republicans with the following linear probability model:

$$\begin{aligned} Pr(\text{Turnout}_{idt} = 1) = & \eta_d + \delta_t + \beta_1 \text{Republican}_{idt} + \beta_2 \text{Distrust}_{idt} \\ & + \beta_3 \text{Republican}_{idt} \times \text{Distrust}_{idt} + \beta_4 \text{Republican}_{idt} \times \delta_t + \beta_5 \text{Distrust}_{idt} \times \delta_t \\ & + \beta_6 \text{Republican}_{idt} \times \text{Distrust}_{idt} \times \delta_t + v\mathbf{X} + \epsilon_{idt} \end{aligned} \quad (17)$$

where $Pr(\text{Turnout}_{idt} = 1)$ is self-reported voting of individual i resident in district d in presidential election t . In our preferred specification, we include district fixed effects η_d , and a set of individual-level covariates $\mathbf{X}$ (i.e., sex, age and race categories, employment status and university education). β_6 estimates the difference in the probability to turnout for Republicans with low trust between the three time periods with and without Trump. We expect this coefficient to be positive for election years 2016 and 2020, when Trump was the Republican presidential candidate, although Figure 3 already shows that distrustful Republicans were not more likely to turn out to vote in the 2020 elections compared to 2012. Standard errors are clustered by congressional district.

Table G.14 displays the effect of the triple interaction between Republican, distrust, and election year on the probability of turnout in the presidential elections for different categories of respondents. Compared to 2012, distrustful Republicans in 2016 are 13-15 percentage points more likely to turn out to vote during the Trump’s first candidacy.

Table G.14: Effect of Trump candidacy on Turnout of Distrustful Republicans.

	Pr(Turnout = 1)		
	(1)	(2)	(3)
Republican	0.053*** (0.011)	0.062*** (0.012)	0.046*** (0.011)
Distrust	-0.102*** (0.032)	-0.095*** (0.032)	-0.063** (0.031)
EY 2016	-0.102*** (0.012)	-0.100*** (0.012)	-0.120*** (0.012)
EY 2020	-0.005 (0.009)	-0.006 (0.009)	-0.042*** (0.010)
Republican $\times$ Distrust	0.019 (0.057)	0.002 (0.058)	-0.009 (0.056)
Republican $\times$ EY 2016	0.035* (0.020)	0.024 (0.020)	0.026 (0.020)
Republican $\times$ EY 2020	-0.018 (0.015)	-0.024 (0.016)	-0.018 (0.015)
Distrust $\times$ EY 2016	0.002 (0.045)	-0.002 (0.045)	-0.022 (0.044)
Distrust $\times$ EY 2020	-0.053 (0.039)	-0.052 (0.039)	-0.034 (0.038)
Republican $\times$ Distrust $\times$ EY 2016	0.134* (0.074)	0.146* (0.075)	0.150** (0.072)
Republican $\times$ Distrust $\times$ EY 2020	0.043 (0.073)	0.051 (0.073)	0.032 (0.072)
Individual Controls			✓
R ²	0.038	0.073	0.137
Observations	15,086	15,086	14,602
State FE	✓		
Congressional District FE		✓	✓

Notes: OLS estimates (linear probability model). SE clustered by congressional district. Outcome is probability to vote in the presidential elections. Individual-level covariates include: sex, age categories, university education, employment status. Reference election year is 2012. Distrust = 1 if respondent never trusts the government in Washington to do what is right. Signif. codes: ***: 0.01, **: 0.05, *: 0.1.