

The Shift to Commitment Politics and Populism: Theory and Evidence*

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

November 20, 2023

Abstract

The decline in voters' trust in government and the rise of populism are two concerning features of contemporary politics. In this paper, we present a model of commitment politics that elucidates the interplay between distrust and populism. Candidates supply policy commitments to mitigate voters' distrust in government, shrinking politicians' levels of discretion typical of representative democracies. Alongside commitments, candidates rationally choose the main strategies associated with populism, namely anti-elite and pro-people rhetoric. We match novel data on voters' distrust towards the U.S. federal government with the Twitter activity of more than 2,000 candidates over five congressional elections and show that distrust is strongly associated with candidates' supply of commitments and populist rhetoric, which are also effective strategies at mobilizing distrustful voters. We also show theoretically that the shift to commitment politics determines greater aversion to checks and balances, and hence even illiberal populism can emerge.

Keywords: Populism, Commitment, Anti-Elite Rhetoric, Trust, Turnout, Agencies of Restraint

JEL codes: D72, D78, P16

*We wish to thank Jack Blumenau, Paul Bose, Maria Carreri, Carlo Cusimano, Georgy Egorov, Ben Enke, Gabriele Gratton, Matt Jackson, Giovanna Invernizzi, Ken Shepsle, Ken Shotts, Christopher Pfeufer, Carlo Schwartz, Konstantin Sonin, Richard van Weelden, and Stephane Wolton for helpful comments and interesting discussions on earlier drafts. We thank Domenico Tripodi, Marvin Pappalettera, and Francesco Migliore for excellent research assistance. Comments by seminar participants at EUI, Zurich, Leibniz Center, CESifo, Bari 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. The usual disclaimer applies.

[†]Bocconi University

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

[§]University of Padua and University of Manchester

[¶]Free University of Bozen-Bolzano

1 Introduction

Populism is on the rise globally and manifests itself in various forms. Scholars across the social sciences have shown that populism is primarily a rhetorical strategy, with populist leaders using anti-elite discourse to appeal to the grievances of the “pure people” against a “corrupt elite”. Populists capitalize on people’s fears and perceived threats, ranging from immigration to automation and the more general risk of economic insecurity. Populist leaders often advocate for illiberal reforms aimed at building an unmediated relationship between the people and their leader, “draining the swamp”, attacking democratic institutions, ostensibly giving the country back to the people. Ultimately, populism has been considered detrimental for the economy and at odds with the core principles of liberal democracies. However, social scientists are left with no theory that condenses these aspects in an integrated account of populism, and the connections between populism – on the one hand – and policy, rhetoric, and democracy – on the other hand – often seem to escape. In this paper, we provide novel theory and evidence of populism centered on commitment politics, which shed new light on how these aspects are tied to one another.

For years, advanced democracies have witnessed a steady decline in citizens’ trust in government. In the U.S., for instance, the share of individuals who say they trust the government to do what is right dropped from 73% in 1958 to 20% in 2022.¹ Social scientists have identified several sources of such growing distrust, encompassing two broad categories: economic and cultural threats. The costs imposed on some groups of citizens by rising unemployment (Algan et al. 2017), globalization (see e.g., Autor et al. 2020, Rodrik 2018, Colantone & Stanig 2018), automation (see e.g., Acemoglu & Restrepo 2020), and the 2008 financial crisis (Algan et al. 2017, Guiso et al. 2021) all increased economic insecurity and distrust towards political institutions. A parallel strand of work has focused instead on the cultural causes of distrust. Trust in political elites’ abilities to protect group-based identities plummeted, triggering value change in subsets of the population. (see e.g., Laitin 2018, Norris & Inglehart 2019, Berman 2021).²

In this landscape of generalized grievances and resentment, populist parties and leaders found fertile ground. In their 70-country-wide study, Funke et al. (2023) show that, in 2018, 16 countries were ruled by a populist (25% of their sample), and similar patterns are reported by the PopuList team for the vote share of populist parties in Europe (Rooduijn et al. 2019). The key trait of these parties and leaders, compared to traditional political actors, is a clear proposed

¹See <https://www.pewresearch.org/politics/2022/06/06/public-trust-in-government-1958-2022/>.

²Both the economic and cultural mechanisms are well summarized in a survey article by Guriev & Papaioannou (2022).

distinction between the “pure” people and the corrupt elites, deemed to be responsible for all society’s ills. This people-vs-elites feature is the cornerstone of every definition of populism used in the social sciences (Mudde & Rovira Kaltwasser 2018). Several scholars consider populism a thin-centered ideology which portrays society as divided into the “pure people” and the “corrupt elite” (the so-called “ideational” approach) (Mudde 2004), whereas others view populism as a strategy employed by politicians in an attempt at “winning and exercising power”, by creating unmediated ties between the leader and voters (Weyland 2001, 2017). Other definitions used to study populism – especially in Latin America – still focus on candidates’ incentives to propose policies that promote the interest of common citizens and signal distance from traditional elites (Dornbusch & Edwards 1991, Acemoglu, Egorov & Sonin 2013). What all these definitions have in common is a clear, moral separation between the people and the elites. As a result, the use of anti-elite and pro-people rhetoric is a standard measure of populism in more recent empirical work too (Hawkins et al. 2019, Gennaro et al. 2021, Di Cocco & Monechi 2021, Dai & Kustov 2022).

In this paper, we show that commitment politics is what links the erosion of voters’ trust in government and the rise of populism. Reduced trust in political institutions, together with increased demand for protection, increases the electoral chances of candidates who commit to specific policies, pre-emptying the discretion typically granted to politicians in representative government. The demand for policy commitments – often in the form of protection policies – is to be understood broadly, as a response to the most pressing economic or cultural threats. We also show that, once politicians shift to a commitment strategy, all the other typical features of populist behavior follow as rational complementary strategies, namely anti-elite and pro-people rhetoric, fake-news production, and aversion to checks and balances.

First, we present a simple principal-agent baseline model, where a representative voter (principal) selects a politician (agent) who chooses a policy. At the time in which a politician has to be chosen by the principal, the optimal policy is uncertain and depends on the realization of a state of the world. The voter selects either a politician who commits to implement an ex-ante determined policy (a committed agent) or one who is delegated to choose the ex-post optimal policy for the voter (a trustee). In the latter case, whether the politician will choose or not the ex-post optimal policy depends on the probability that *the elite* distorts her choice. The choice of commitment is more likely the greater the expected strength and distortion caused by the elite. A lower trust that politicians can resist the elite capture makes the principal (voter) prefer a committed agent who panders to the voter’s ex-ante desiderata. In order to

keep the model simple, we assume that a commitment can be maintained.³ Pandering to the ex-ante voter’s beliefs has been considered by itself an important feature of populism (see e.g., 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.

After presenting the simplest prediction of the principal-agent framework, we extend the analysis to an electoral competition setting that closely represents that of the United States. In primary elections, two principals (parties) select two politicians, and such agents then compete in a general election. In this case, the two primaries can lead to different choices of candidates to run in the general elections, depending on the different levels of distrust, beliefs, and preferences between the two parties. Given a vector of parameters, it is possible that, in the general elections, one party has selected a candidate supplying a greater level of commitment than the other. We show that the more committed candidate has stronger incentives to use anti-elite and pro-people rhetoric to increase the relative turnout of her own supporters. We also show that decreased trust increases turnout for the more committed candidate and depresses the turnout for a full-delegation/trustee candidate. In terms of the likelihood of selection of a committed candidate by a party, we show that it is more likely the less uncertain its members are on the optimal policy ex-ante. By the same token, it is in the interest of a committed candidate to unite her electorate on the policy she is committed to pursuing. This can induce polarization whenever candidates commit to different policies.

We provide novel empirical evidence supporting the key predictions of the theory on the connection between distrust, shift to commitment politics, and populist rhetoric. In the context of the U.S. congressional elections, we create district-level estimates of voters’ distrust in the federal government, which we match with the universe of tweets posted by more than 2,000 candidates during the five congressional elections between 2012 and 2020. We use natural language processing and supervised machine learning techniques to detect when a tweet contains a policy commitment or populist rhetoric. We find that voters’ distrust is strongly associated with candidates’ supply of commitments and populist rhetoric, especially when the candidates tweet about topics where partisan voters display high levels of consensus and convictions on what the right policy is. We show that, compared to the 2012 presidential elections, distrustful Republicans were more likely to cast a vote during the 2016 elections when Trump was the Republican candidate, lending support to the prediction whereby the supply of commitment

³Note that even in the absence of a commitment technology, the principal can effortlessly monitor an agent who made a clear promise to deliver a specific policy, while, in the case of full delegation, the principal has to learn (with a costly effort) the state of the world to understand whether the delegated agent made the principal or elite’s interests when choosing the policy.

and populist rhetoric is effective at mobilizing distrustful voters.

After establishing, theoretically and empirically, that the shift to commitment politics is the logical mechanism that connects the erosion of trust phenomenon with the populism wave, we show in an extension that the shift to commitment politics also determines a demand for reduced checks and balances, hence rationalizing existing evidence on voters' willingness to trade off democratic principles for policy preferences (Graham & Svulik 2020, Alsan et al. 2023, Zaslove & Meijers 2023) as well as anecdotal regularities on the anti-democratic steps taken by populist governments. We illustrate this by introducing in the theory an "agency of restraint", a player that can veto a policy chosen by the policy maker – which could be free media, independent bureaucracy, or independent judiciary. In such an augmented model, we show that, under general assumptions, voters who choose a committed agent typically also desire to weaken or eliminate such an agency of restraint. In contrast, voters who prefer full delegation typically prefer to keep an agency that produces checks and balances. Similarly, we argue that an extension of the model could easily deliver the prediction that the committed candidate benefits from making information acquisition costlier for voters, a benefit that can, for instance, be attained by favoring the spread of fake news. Thus, the shift to commitment politics not only correlates with a clear shift towards populist rhetoric in electoral campaigns, but it also determines a potential demand for illiberal reforms, a pattern that has received sustained attention among political commentators and democratic theorists (for a recent review, see Urbinati 2019).

Once the connections between distrust, commitment, and populism are drawn, we can better understand the policy consequences of populist governments. Our predictions are consistent with the findings in Sasso & Morelli (2021) and Bellodi et al. (2023), who show that when populists are in power, they replace competent bureaucrats with loyal bureaucrats to minimize the risk of resistance to the populist commitment agenda. Moreover, there is mounting anecdotal evidence that populist governments from across the political spectrum tend to weaken (besides bureaucracy) other agencies of restraints, from the media to the judiciary. Evidence from Hungary shows, for instance, that checks and balances have been reduced during Orban's governments.⁴ Similarly, Mexico's President Obrador recently approved a reform that strips the National Electoral Institute's independence ahead of the 2024 presidential elections.⁵ In Poland, the Law and Justice Party (PiS) substantially reduced judicial independence and media freedom after a successful electoral campaign dense with policy commitments.⁶ The findings in

⁴The constitutional court can no longer block executive policy decisions. See <https://www.bbc.com/news/world-europe-21740743>.

⁵See <https://foreignpolicy.com/2023/03/23/lopez-obrador-electoral-reforms-mexico-democracy-ine/>.

⁶See <https://www.bbc.com/news/world-europe-65809525>.

Funke et al. (2023) – who identify a negative long-term effect of populist leaders on GDP per capita – are broadly consistent with the expectation that the pursuit of unconditional, ex-ante determined commitments is likely to be very costly for the economy. Our paper provides a new, integrated framework for studying populism that can answer theoretical questions about candidates’ incentives to adopt a populist behavior while rationalizing several empirical regularities about populists in power.

2 Related Literature

This paper contributes to several strands of the literature in political economy and political science. The distinction between the committed agent versus trustee model of political agency has first been introduced by Fox & Shotts (2009) in the context of optimal accountability of incumbents. Along the same dimension, Kartik et al. (2017) show that in electoral competition the equilibrium degree of discretion and ambiguity in candidates’ policy platforms depends on the divergence in policy preferences that exists between voters and candidates.⁷ Di Tella & Rotemberg (2018) show that distrust is also related to betrayal aversion, which strengthens the desire for commitments. While some accounts focus on the credibility of candidates’ pledges (Van Weelden 2013), for simplicity our model assumes full credibility of commitments. Our results would be qualitatively the same when allowing for partial credibility.

Several papers provide theories of populism that assume cognitive differences among voters, which we do not invoke in our model. Levy et al. (2022), for instance, depict populist policies as simplistic ones desired by unsophisticated voters, who sometimes win the elections because of an intense dislike for the status quo, whereas Crutzen et al. (2020) show that, if voters are divided into an informed minority and an uninformed majority, parties tend to cater more to the better-informed elite, increasing common people’s disaffection for traditional parties and incentives to enter for populist parties. Egorov & Sonin (2021) present a model where an informed minority (the elite) can advise the uninformed majority on candidates’ competence, when one candidate is biased towards the elite and the other one is unbiased, and look at the conditions under which the uninformed majority follows the elite’s advice. Unlike this body of work, in our model there are no assumptions on cognitive differences across groups, and we focus instead on the role of trust, beliefs, and preferences.

Our paper also speaks to recent work that emphasizes changes in social identification, highlighting how the national versus global identity has become the most relevant cleavage, even

⁷On commitment versus flexibility, see also Amador et al. (2006). A similar distinction is present in Ghosh & Tripathi (2012) and Bueno de Mesquita & Friedenbergh (2011), but in their context, the committed agent is an “ideologue”. In our paper ideology is not a constitutive feature of the populist’s strategy.

more relevant than the standard economic left-right ideological dimension (see e.g., Shayo 2009, Ford & Jennings 2020, Besley & Persson 2019, Bonomi et al. 2021, Gennaioli & Tabellini 2023). Our take on social identification phenomenon is that, once it becomes rational for parties and politicians to choose a commitment strategy and voters’ distrust stems from a perceived cultural or immigration threat, such parties and politicians may become, as a rational consequence, strategic suppliers of messages about the protection of national and communal values.

Finally, in the extension section on the consequences of the shift to commitment politics for liberal democracy in general, we contribute 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 happy 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.

3 Distrust and the Demand of Commitment

In this section, we present a baseline principal-agent framework that highlights the sources of “demand” of commitment.

A principal G (a party or a voter) delegates at time 1 an agent (a politician) g to choose a policy $q \in \{\ell, h, z\} \equiv Q$ at time 2 in case g is the decision maker. The ex-post optimal policy for G depends on the state of the world at time $t = 2$, which is unknown at time 1 and is observed by the agent g with probability 1 at time 2.⁸ The principal chooses $S \subseteq Q$, i.e. the degree of discretion left to the agent. In other words, S is the subset of available policies from which the agent is delegated to choose at time 2 after observing the realization of the state of the world. When payoffs are introduced below, we describe formally the incentives that induce the agent to choose a policy inside S . If $S = Q$, then there is full delegation and the agent can choose any policy. On the opposite extreme, if $|S| = 1$, the set S is a singleton, and the agent is forced to commit to implement the policy ex-ante chosen by the principal.

The tension between these strategies, full commitment, partial commitment or full del-

⁸We could also assume that the political agent observes the state of the world with probability lower than one that depends on her competence or ability and we would obtain qualitatively similar results.

egation, comes from the fact that the agent can be influenced by an “elite”.⁹ The elite may distort the agent’s choice. Full delegation allows an unbiased 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, on the contrary, prevents the elite from distorting the agent’s choice, but it does not allow to adjust the policy ex-post. We adopt the simplest setting to highlight this tension.

Assume that the principal has a belief $\lambda > 1/3$ that the elite’s preferred policy is z and, with equal probability $\frac{1-\lambda}{2} < \lambda$, that the elite’s preferred policy is one of the other two policies, l or h . In contrast, we assume that the principal’s belief that policy z is ex-post optimal, γ_z , is lower than the probability she attaches that reform h or l can be ex-post optimal, denoted respectively by γ_h and γ_l . This assumption reflects a situation where the principal feels that policy z – for instance a policy preserving the status quo – is unlikely to be optimal, and she is more likely to prefer policy h or l , that may represent policy change. Conversely, $\lambda > \frac{1}{3}$ captures the higher probability that the elite may favor a policy z that preserves the status quo.

We denote by $q^* \in Q$ the ex-post optimal policy when uncertainty over the realization of the state of the world is resolved. In this section we assume – without loss of generality for the principal-agent baseline – that between h and l the ex-ante optimal policy is h : $\gamma_h > \gamma_l$.¹⁰

The respective payoffs for the principal and the agent are assumed to be as follows.

- The principal obtains a utility normalized to 0 if the implemented policy is q^* . If the implemented policy is different from q^* , then she gets -1 .
- The agent’s payoff is $R - \epsilon(\mathbf{1}_{q \neq q^*})$, with $\epsilon < R$, if she chooses a policy within the set S ; whereas if she chooses a policy outside S , then the reward R does not accrue.¹¹

The principal assigns a positive probability p that the elite influences the agent in her choice within S .¹² In other words, p reflects the principal’s distrust that the political class acts in the interests of its constituents. On the other hand, we assume that R is large enough to make sure that the agent never chooses a policy outside the set S .

⁹The elite is never an explicit player in our model, and for different situations the elite could mean the establishment, economic powerful actors, or any combination of external interest groups that could influence policy makers. As it will become clearer, we do not need to be specific about the identity of the elite. For this reason, and for simplicity, we do not treat the elite as an active player.

¹⁰In the next section, when we study the competition between two principals (parties), we allow for heterogeneity in ex-ante preferences and/or beliefs across parties.

¹¹The most obvious interpretation of R is the value of being reconfirmed in the future.

¹²We do not model the bargaining game between the elite and the politician but we assume that with probability p the elite manages to influence the agent to choose the elite’s preferred policy if it is within S . If we had the elite as a player of the bargaining game, like in Morelli et al. (2021), p could be interpreted as the probability that the elite bribes the agent with an amount in (ϵ, R) .

A contract between the principal and the agent specifies the set S of policies the agent can choose from. We can now study which contract a principal offers to the agent. If the principal offers full delegation, $S = Q$, the expected payoff is equal to

$$-p \left[\lambda(1 - \gamma_z) + \frac{1 - \lambda}{2}(1 - \gamma_h) + \frac{1 - \lambda}{2}(1 - \gamma_l) \right] = -p \left[\lambda(1 - \gamma_z) + \frac{1 - \lambda}{2}(1 + \gamma_z) \right].$$

If the principal offers partial commitment, $S = \{h, l\}$, the expected payoff is equal to

$$-p \left[\frac{1 - \lambda}{2}(1 - \gamma_h) + \frac{1 - \lambda}{2}(1 - \gamma_l) \right] - \gamma_z = -p \left[\frac{1 - \lambda}{2}(1 + \gamma_z) \right] - \gamma_z.$$

Thus, partial commitment yields higher expected utility than full delegation if

$$p > \frac{\gamma_z}{\lambda(1 - \gamma_z)} \equiv p_1. \quad (1)$$

If the principal offers a full commitment contract, $S = \{h\}$, the expected payoff is equal to $-\gamma_l - \gamma_z$. Thus, full commitment is preferred ex ante to full delegation if

$$p > \frac{(1 - \gamma_h)}{\lambda(1 - \gamma_z) + (1 - \lambda)(1 + \gamma_z)/2} \equiv p_2 \quad (2)$$

This implies that the principal chooses full delegation to the agent (trustee) if p is less than the minimum of the two thresholds p_1 and p_2 in (1) and (2). Given that p captures *distrust* in the simplest possible way, this already suffices to see that whenever trust goes sufficiently down, then political agency choices shift towards some type of commitment, partial or full.

Proposition 1. *As long as either inequality (1) or (2) (or both) are satisfied, full delegation is never chosen as agency contract. Partial commitment is chosen for intermediate levels of distrust and full commitment for higher levels of distrust.*

Note that full commitment is preferred to partial commitment when

$$-p \left[\frac{1 - \lambda}{2}(1 + \gamma_z) \right] - \gamma_z \leq -\gamma_l - \gamma_z,$$

or

$$p \geq \frac{2\gamma_l}{(1 - \lambda)(1 + \gamma_z)}.$$

Note also that when the principal is quite sure about what she prefers ex-ante, namely γ_h is very high, then the RHS of (2) is very low, and hence the full commitment choice is typically optimal when, on top of low trust, there is also a clear ex-ante policy preference.

The following comparative statics corollary connects the thresholds for the different contracts and the key parameters of the model.

Corollary 1. *p_1 and p_2 are increasing in γ_z and decreasing in λ , within their restricted domains.*

The meaning of this corollary can be illustrated by considering the joint effect of a lower γ_z and a higher λ . This joint effect represents a higher distance between the principal's evaluation of a policy that preserves the status quo (z) and the likely preference of the elite for z . When this distance increases, the two thresholds p_1 and p_2 decrease, which means that it is more likely that the principal prefers commitment over delegation.

Consistent with this corollary, and as we show in the following section, a committed agent will try to convince voters that λ is high and γ_z is low. This type of campaign attempting to increase the perceived distance between the preferences of the principal and those of the elite is the model's representation of anti-elite rhetoric, which is a constitutive feature of most definitions of populism.

Discussion on the Main Assumptions

The model assumes that R is high enough to make sure that the agent always chooses a policy inside S , and that the principal has no cost in withholding R (i.e., firing the agent) in case of contract violation.

These assumptions are consistent with both theoretical and empirical work. The first assumption relates to the value of office and the presence of accountability, while the second is typical, for example, of retrospective voting models. For instance, it has been shown that voters punish politicians who deviate from their electoral commitments without confirming them in office (see the survey by Duggan & Martinelli 2017). Similarly, there is evidence on reputation costs (see e.g. Alesina 1988, Levy 2007) and cost of lying (Callander & Wilkie 2007, Carrillo & Dewatripont 2008), both pointing to politicians' reluctance to deviate from electoral promises (for experimental evidence, see Corazzini et al. 2014). Even within the citizen-candidate literature the selection of a candidate is motivated precisely by the policies that she is credible to pursue. Politicians will run electoral campaigns with promises that are coherent with their preferences for policies, which will be indeed implemented once elected.

In Section 6.1, we model a scenario where commitments are not always implemented by allowing the possibility that the media, the judiciary, or the bureaucracy can block such commitments. While the presence of these agencies of restraint may reduce the space of parameters under which commitment is chosen, we actually show that when the principal prefers commitment over delegation, she also prefers weaker checks and balances. Finally, the empirical

analysis in Section 5 shows that distrust is strongly associated with greater supply of commitments, regardless of how credible such commitments are in the eyes of voters.

4 Supply of Commitment, Turnout, and Populism

While in the previous section we have shown that distrust determines a demand of commitment by voters, in this section we need to show that, under corresponding conditions, the equilibrium behavior of politicians includes both commitment and anti-elite rhetoric, which is considered in the literature the most prominent strategy of populist politicians. We thus introduce electoral competition in the model and focus on the mobilization strategies determining the probability of winning for candidates potentially supplying different levels of commitment. To preview the results, we show that any campaign strategy aimed at increasing voters' anti-elite sentiment and conviction about the ex-ante desirability of a policy commitment determines a higher probability of winning for the most committed candidate.

4.1 Elections

We extend the model to allow for two-party electoral competition in order to characterize the equilibrium behavior of candidates and to emphasize what determines their probability of electoral success.

Let there be two parties, $G = A, B$ each with mass equal to one. Each of them elects a candidate in a primary election (time $t = 0$) and then the selected candidates $g = a, b$ compete in a general election at time $t = 1$. There are at least two candidates for each party in the primary and each candidate j of party G proposes an electoral platform $S^j \subseteq Q$. For each party G , the ex-post optimal policy q^{G*} depends on the state of the world at time $t = 2$, which is unknown at the time of the elections, but observed by the elected politician before making the decision. The winner w of the general election chooses a policy $q \in S^w$.

As for the agent's payoff in Section 3, the candidate's payoff is $R - \epsilon(\mathbf{1}_{q \neq q^{G*}})$ with $\epsilon < R$ if she chooses a policy within the set S^w ; whereas if she chooses a policy outside S^w , then the reward R does not accrue. The payoff of a member of party g is, as for the principal, normalized to 0 if the implemented policy is q^{G*} , and -1 otherwise.

We denote by γ_q^G the probability that the representative member of party G assigns to the event that policy q can be ex-post optimal. As in the previous section, we assume that for each party G , $\gamma_z^G < \gamma_t^G$, $t \in \{h, l\}$. We denote by $\lambda^G > 1/3$ the probability assigned by members of party G that the policy z will be the one preferred by the elite at time 2, while the probability that the elite prefers h or l is, for each of these two policies, $\frac{1-\lambda^G}{2}$. Finally, members of party

G assign a positive probability p^G that the choices of politicians once in office are influenced by the elite. We assume that in each party G there is a continuum (unit size) of party members that all share the parameters γ_q^G , λ^G and p^G , while being heterogeneous in terms of cost of voting.

The cost of voting can be ignored for primaries, given that all voters of the same party have the same beliefs and preferences, but it matters instead in the general elections. To determine endogenous turnout, we make an “as if pivotal” assumption (Alesina & Rosenthal 1996), i.e. we assume that each citizen turns out to vote 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 $\mu^G \in [-\frac{1}{2\psi}, \frac{1}{2\psi}]$ 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 no one votes, election results are determined by the toss of a fair coin.

In order to limit the attention to an interior and unique solution, we make the following assumption.

Assumption 1: $\psi \leq \bar{\psi}$, with $\bar{\psi} > 0$ sufficiently low.

4.2 Equilibrium Turnout Effects

We solve the model by backward induction. Consider the general election. 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 $EU^G(S_K^j)$ denote the expected policy utility for a voter of party G if a candidate with platform S_K^j from party $K = A, B$ is elected. Given that at the general elections stage there is only one candidate for each party, we can now ignore the superscript j .

If $EU^G(S_G) > EU^G(S_{-G})$, voter v of party G votes for her preferred candidate if $EU^G(S_G) - EU^G(S_{-G}) \geq c_v^G + \mu^G$, and otherwise she abstains.

We now denote by \hat{S}^g the electoral platform that gives to the members party G the highest expected policy utility. The following lemma helps characterize the results of the primary elections.

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

The intuition of the proof, which we report in the appendix, is the following. Even in cases

in which the most desirable candidate and platform conditional on winning (\hat{S}^g) is not 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 about the strategy of the opponent makes \hat{S}^g preferable for voters of party G .¹³

The lemma implies that the choice of an electoral platform within each party G is determined by inequalities similar to those presented for the demand side, with the difference that p^G and γ_q^G are party specific. Consistently, in the primary elections, voters may demand a full or partial commitment and may have different willingness to trade the possibility of getting an ex-post optimal policy for the reduction of the probability that the elite can distort politicians' policy ex-post. For example, if p^A is high and higher than p^B , then it is possible that A 's primaries see the victory of a candidate with higher commitment than the candidate winning B 's primaries. Thus, the unique equilibrium can display general election competition between candidates with different levels of commitment.¹⁴

The following proposition highlights that citizens' distrust towards politicians affects turnout of parties in the general elections, but in different ways depending on the electoral platform offered by the candidates.

Proposition 2. *If both candidates' equilibrium platform displays full commitment, then a variation in the distrust of any party does not have any effect on voters' turnout. However, if candidate g offers full commitment and the opponent does not, the following holds:*

- the higher is p^G , the higher is turnout of party G ;
- the higher is p^{-G} , the lower is turnout of party $-G$.

This result establishes that distrust can have opposite effects on voters' turnout in the general elections when the two candidates supply different levels of commitment. Intuitively, if at the time of the elections there is an additional reduction in trust (p^G and p^{-G} go up), the voters who chose a more committed candidate see their utility differentials between their preferred candidate and the opponent increase and hence they are more likely to turn out to vote. The opposite is true for voters who had chosen a candidate supplying less commitment.

4.3 Anti-Elite Rhetoric

So far we have kept the parameters of the model, γ_q^G , λ^G and p^G as given. We now allow candidates to add to their platform choice the complementary choices of campaign messages

¹³This lemma would continue to hold even with slightly different assumptions from Assumption 1. 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.

¹⁴Uniqueness of equilibrium is guaranteed by lemma 1.

and rhetoric aimed at updating voters' beliefs about the ex-post optimal policy and the policy preferred by the elite.

Let us recall from Corollary 1 that commitment is more likely than delegation when anti-elite sentiments among voters are strong (high λ and low γ_z). We now stress that there is a corresponding supply side of the same coin. A committed candidate will have an incentive to fuel anti-elite sentiment further.

Without attempting any microfoundation of how rhetoric can successfully modify such beliefs, in the model the closest representation of a rhetorical strategy aimed at increasing anti-elite sentiment (the so-called anti-elite rhetoric) entails an attempt to reduce γ_z^G and to increase λ^G (e.g., such a strategy could be summarized in a sentence like “The elites want to keep things as they are and this would be bad for the people.”).¹⁵

Proposition 3. *Whenever a party G chooses commitment, full or partial, and faces a fully delegated opponent, then it is in the interest of their candidate g to use anti-elite rhetoric in the electoral campaign, pushing γ_z^G down and λ^G up.*

It is important to note that, even though we do not explicitly model the competition in primary elections, a candidate who chooses a commitment platform in the primaries and faces a candidate who offers full delegation has the same vested interest to employ anti-elite rhetoric right from the primaries.¹⁶

4.4 Policy Convictions and Belief Polarization

An additional campaign strategy at candidates' disposal is to influence voters' ex-ante beliefs about the ex-post optimal policy (e.g., γ_h^G). To explain what we mean by “increasing conviction”, let party G 's voters have $\gamma_h^G > \gamma_l^G$. Then, policy convictions increase iff γ_h^G goes further up. Recall from Equation 2, where we characterized the conditions under which the principal prefers full commitment over full delegation, that the threshold p_2 above which full commitment is preferred is decreasing in the ex-ante belief that a policy is ex-post optimal (γ_h). Intuitively, the more certain voters are that one specific policy is ex-post optimal (i.e., strong policy convictions), the greater will be the demand for commitment on that policy. Therefore, higher voters' convictions within parties translate into higher chances of parties shifting to committed candidates. On the supply-side, this implies that a committed candidate, in order to mobilize

¹⁵We could also allow campaign strategies to increase p^G (i.e., the probability of elite capture) without any substantive change to our results. However, we intentionally put less emphasis on such strategies because trust – as all cultural parameters – is stickier than voters' beliefs about optimal policies and elites' interests.

¹⁶In order to have candidates with different platforms in primaries, we would have to add realistic heterogeneity of candidates in terms of trustworthiness, ability, credibility, etc. We do not do this for the sake of simplicity.

their voters, will run an electoral campaign trying to further increase voters’ convictions that the chosen policy is ex-post optimal.

The following proposition summarizes these observations.

Proposition 4. *A commitment strategy is more likely to be chosen in the presence of strong policy convictions. Moreover, committed candidates can always improve their probability of winning by increasing the conviction of their voters over the committed policy.*

Now consider a situation in which the two parties have ex-ante different preferred policies, $\gamma_h^A > \gamma_l^A$ and $\gamma_l^B > \gamma_h^B$. In this context, an attempt by one or both candidates to increase the policy convictions of her supporters (increasing γ_h^A and/or γ_l^B) determines higher “belief polarization”, defined as in Haghtalab et al. (2021) as divergence in beliefs.¹⁷

Commitment politics can therefore have a twofold connection with belief polarization. First, strong policy convictions within parties increase the probability of commitment candidates. Second, commitment candidates will reinforce convictions within parties to strengthen voters’ preferences for commitment and ultimately be selected by their voters. We should therefore expect policy commitments to be about high conviction policies.

5 Evaluation of the Predictions

In this section we provide evidence in favor of the main predictions contained in our propositions.

- (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, as discussed in Proposition 4.
- (iii) Distrust increases turnout by supporters of parties running with a commitment candidate, as per Proposition 2.

We test these predictions in the context of U.S. congressional 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 precisely mirror important features of the model. Finally, candidates’

¹⁷Belief polarization is potentially connected to high intensity of preferences and affective polarization. Even though this connection is outside the scope of this paper, it is intuitive that two camps with opposite policy convictions can lead to tensions and increased inter-group animosity.

campaigns for the House take place primarily on social media, and most candidates rely on social networks to communicate with their constituents in a high-frequency and highly personalized fashion. Specifically, 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 mobilize voters, attack their opponents, and commit to policy.

To preview the results, we show that (i) candidates running in districts with higher distrust towards the federal government use more commitments and anti-elite/populist rhetoric in their tweets, (ii) a tweet that pertains to topics with high within-party policy convictions is more likely to contain commitments, and that (iii) turnout in the general elections of voters of a party that chose a commitment candidate in the primaries increases when distrust is large.

5.1 Data & Measurement Strategies

In this section 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.

5.1.1 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, as is the case for respondents in a sample representative of the national population. 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

Table 1: ANES Surveys, Descriptive Statistics.

Year	N. Respondents	% Distrust
2012	5,663	5.9
2016	4,079	12.8
2020	7,093	8.7

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

to do what is right?” and re-code the variable in order to have dichotomous answers. 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. First, it explicitly asks respondents about *trust*, without resorting to other proxies or distinct concepts like confidence of competence.¹⁸ Second, by referring to the government in Washington, respondents are prompted to think about the same entity when answering the question. In Table 1 we report the sample size and the share of respondents who report never trusting the government in Washington across each wave.¹⁹

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).²⁰ First, we fit a multilevel model to the survey data to predict the level of distrust in government for specific categories of respondents (i.e., combining sex-education-age variables into unique categories of “ideal types” distributed across the sub-national units). 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 of the Census Bureau.²¹

In Figure 1, we plot the estimates for the year 2020. Each same-size hexagon represents a

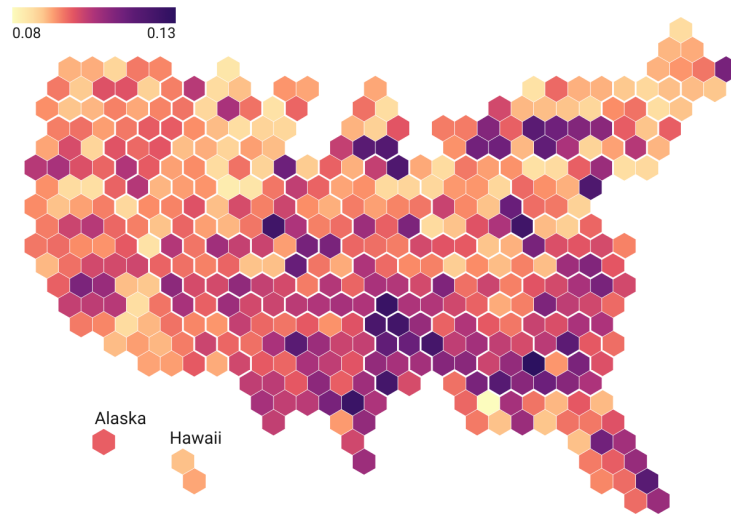
¹⁸In the appendix (Table F.11) we show that the results are similar when using alternative survey questions: “How many of the people running the government are corrupt?” (Distrust = 1 for “All”/“Most” responses), “Would you say the government is pretty much run by a few big interests looking out for themselves or that it is run for the benefit of all the people?” (Distrust = 1 for “Run by a few big interests” responses).

¹⁹The 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 B1 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).

²⁰We describe the estimation in greater detail in Section B in the appendix.

²¹With the proposed method we are not able to produce estimates of distrust for Republican and Democratic voters separately, for we do not have census data on party identification of individuals living in congressional districts.

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.

congressional district and darker shades indicate higher distrust. The average share of estimated distrustful individuals across the districts is 9.7%, compared to 6% in 2012, and 13% in 2016. Alabama, Louisiana, and Mississippi are the states where average distrust is highest in 2020, all with a share of distrustful individuals above 10%. There is also significant within-district and over-time 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 50% larger.

5.1.2 Supply of Commitment and Populist Rhetoric

To measure the supply of commitments and populist rhetoric we look at the Twitter activity of U.S. congressional candidates. 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 C 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. We are relatively certain that for 11% of the 1,283 candidates for which we do not find an account, the reason is that indeed they do not have an active Twitter account. Given the extensive scraping performed, it is highly likely that the remaining 1,145 candidates (who ran only once) do not currently have an active Twitter account. We downloaded 5.9 million tweets (in English and excluding re-tweets) from the Twitter API for

the period 2012-2021 posted by the 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 build on recent advancements in natural language processing (NLP). 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. First, we produce a numerical representation of tweets which will serve as a matrix of predictors. Second, we build and annotate a training dataset and train the classifier. Third, we assess the performance of the classifier on held-out test sets. Fourth, we use the classifier to make out-of-sample predictions on the universe of tweets collected.²²

To produce numerical representations of tweets we use embeddings, fixed-length vectors capturing the semantic properties of text and allowing us to infer similarity between tweets (i.e., tweets with similar vector representations will have a similar meaning). We compute the vectors with a RoBERTa-base model pre-trained on a large corpus of tweets on top of the original RoBERTa-base checkpoint, which improves the performance of several NLP tasks for Twitter data (Barbieri et al. 2020).²³

We then construct two training datasets, one for policy commitments and one for populist rhetoric. Tweets containing commitments and populist rhetoric are relatively rare and two random samples of tweets would yield high class imbalance (i.e., very few instances of commitment and populist rhetoric tweets). We take several steps to ensure a sufficient balance between classes – which is key for the quality of the classifier (see Section D in the appendix) – to obtain a total of 3,000 tweets for each of the two training datasets.

We annotate the tweets with the assistance of ChatGPT. Among the many capabilities of this large language model, Gilardi et al. (2023) have shown that ChatGPT outperforms crowdworkers for standard annotation tasks. Moreover, ChatGPT allows to annotate large training datasets in a highly cost-effective fashion.²⁴ We pass the $3,000 \times 2$ tweets to the model via the OpenAI API, using the `gpt-3.5-turbo` model. The model takes as input a prompt and returns an answer as in an ordinary chat. In Table 2 we show the prompts we used to ask ChatGPT to annotate the tweets and an example of a response. As a further check on the reliability of ChatGPT as an annotation tool, we compare the labels of ChatGPT to those decided by a

²²We describe each one of these steps in detail in Section D in the appendix.

²³We use the `twitter-roberta-base` model accessed through Huggingface <https://huggingface.co/cardiffnlp/twitter-roberta-base>.

²⁴Letting ChatGPT classify 6 million tweets would however be highly impractical and expensive, but it can be a very reliable assistant for smaller-scale tasks.

Table 2: ChatGPT 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 ChatGPT through the OpenAI API and the response given by the model.

human annotator trained specifically for this task. From the two training datasets, we extract 100 random tweets and asked an independent coder to decide whether the tweets contained commitments or populist rhetoric, based on the same prompts given to ChatGPT. Despite the difficulty of the task, the human annotator and ChatGPT agreed on 82% and 84% of the cases for the commitment and populist-rhetoric tasks, respectively.

Finally, we train several machine learning classifiers on the vector representation of 75% of the tweets in the training dataset, 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.²⁵ If the tweet has a large predicted probability of belonging to the commitment or populist rhetoric classes (i.e., probability larger than .8) is assigned a score equal to 1, and 0 otherwise. We select a high arbitrary threshold to make sure that, for instance, simple references to policies or vague policy promises are not classified as policy commitments, and that general praises of “the American people” are not classified as populist. 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

²⁵The classification exercises for commitment and populist rhetoric display very good performance metrics, with accuracy and F1 scores above .85.

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.²⁶

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!

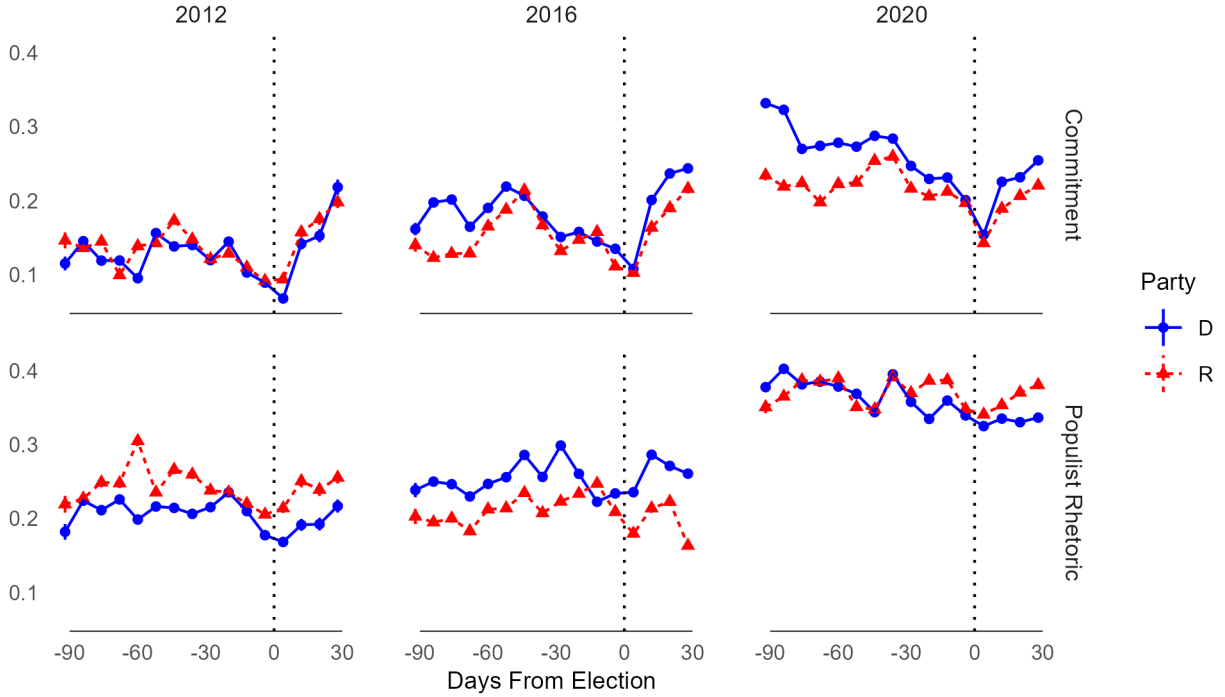
Greg Musselwhite (R-FL) – 2020

In 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 D.5 and Table D.6).

In Figure 2, we show the dynamic share of commitment and populist-rhetoric tweets posted by Democratic and Republican candidates during the 90 days before and the 30 days after the elections. Three patterns emerge from the data. First, the use of policy commitments on Twitter is similar for both Republican and Democratic candidates. Except for the beginning of the 2020 campaign, when Democrats' share of commitments is larger than Republicans', there is no marked difference in candidates' supply of commitments. For the 2012 and 2016 elections, the supply of commitments peaks in the 60-30 days before election day and then follows a decreasing trend until the day of the election, whereas in 2020 the supply of commitment was already high 90 days before the elections. Interestingly, 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 when they will ultimately have the chance to translate those commitments into policy. Second, with respect to populist rhetoric, Republicans and Democrats posted a larger share of populist-rhetoric tweets in the 2012 and 2016 elections, respectively. However, the share of both parties' populist tweets converge in the 10-20 days before the election. Perhaps unsurprisingly, given the heated 2020 campaign, the share of populist-rhetoric tweets among candidates from both parties is almost twice as large in the 2020 elections compared to 2012 and

²⁶As a robustness test, in the analysis we use the continuous predicted probability as outcome variable and find similar results (see Table F.10 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 too (see Figure F.4 in the appendix).

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 September, October, and November (up to election day) of 2012, 2016, and 2020.

2016. Finally, an interesting pattern in the data is that, in the post-election period, candidates from the party that lost the presidency post more tweets containing populist rhetoric.

5.1.3 Policy Convictions

Recall that our theory suggests that the likelihood of commitment is larger if policy convictions are high. To test this prediction we need to identify tweets about policy on which voters' opinions are highly homogeneous.

To identify such policies in the least arbitrary way possible, 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 individuals identifying with the Republican or Democratic party in any given state and year.²⁷ The CCES study regularly asks a representative sample of U.S. individuals 55 questions about policy preferences (2006-2021). Because the CCES includes only a limited set of policies, we select the question with the lowest standard deviation in the responses across party-year-state combinations, which we easily map to seven different topics/policies based on the wording of the questions. The seven topics are

²⁷We group respondents at the state-level to ensure large samples of respondents and geographical variation in policy preferences.

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. The precise wording of the question is “*On the issue of gun regulation, do you support or oppose each of the following proposals? Background checks for all sales, including at gun shows and over the Internet.*” The topic we infer from the question is “gun regulation”. We then repeat this exercise for both parties over time and across states. In the appendix we report the topics with the largest convictions by party-year-state combinations (Table E.8) as well as the mapping of survey questions into topics (Table E.7).

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 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 E.9 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.

5.2 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 and Twitter activity. In particular, we estimate the following linear probability model

$$y_{icdt} = \alpha_c + \eta_d + \delta_t + \beta \text{Distrust}_{dt} + \epsilon_{icdt} \quad (3)$$

where y_{icdt} is the probability of a commitment or populist rhetoric tweet, α_c , η_d , and δ_t are candidate, district, and election-year fixed effects, In line with Propositions 1 and 3, we expect $\hat{\beta} > 0$. In additional specifications, we also estimate state-by-year fixed effects, to account for time-changing factors at the state level. Because β is identified by comparing over-time changes

Table 3: Distrust, supply of commitments and populist rhetoric.

	Commitment			Populist Rhetoric		
	(1)	(2)	(3)	(4)	(5)	(6)
Distrust	0.012*** (0.004)	0.016** (0.007)	0.023*** (0.007)	0.020*** (0.008)	0.024*** (0.009)	0.026*** (0.009)
Controls: Party	✓			✓		
Controls: Incumbent	✓	✓	✓	✓	✓	✓
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 standardized share of distrustful individuals in each district. Signif. codes: ***: 0.01, **: 0.05, *: 0.1

in the Twitter activity among candidates who experience a change in the level of distrust in the district where they run, candidates running only once (or twice but not in two of the 2012, 2016, 2020 elections) receive a 0-weight and do not enter the “effective” sample used in the regression (Aronow & Samii 2016). To mitigate the implicit loss of observations that contribute towards the estimation, we also present results without candidate fixed effects and simply conditioning on the party of each candidate. In our preferred specification we also include incumbency status as a covariate. Standard errors are clustered by district.²⁸

To better characterize the use of Twitter as a communication device during electoral campaigns, we perform the analysis on tweets posted during the two months prior to 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. As a robustness test, Figure F.3 shows results for different windows of time.

In Table 3, we report the main results from estimating Equation 3. The coefficients represent estimated changes in the probability of commitment and populist-rhetoric tweets resulting from an increase in distrust by +1 standard deviation. A 1 SD increase in the share of individuals who never trust the government in Washington to do what is right is associated with an increase in the probability of a commitment tweet by 1.7-2.3 percentage points and of a populist-rhetoric tweet by about 2.4-2.6 percentage points. The effects are precisely estimated

²⁸The results are almost identical if we cluster the standard errors by candidate.

Table 4: High conviction topics and supply of commitments.

	Commitment				
	(1)	(2)	(3)	(4)	(5)
High Conviction Topic	0.172*** (0.014)	0.168*** (0.013)	0.168*** (0.013)	0.168*** (0.013)	0.168*** (0.013)
Distrust				0.023*** (0.007)	0.023*** (0.007)
Controls: Party	✓				
Controls: Incumbent	✓		✓	✓	✓
R ²	0.052	0.069	0.070	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. 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 standardized share of distrustful individuals in each district. Signif. codes: ***: 0.01, **: 0.05, *: 0.1

and sizable if compared to the sample mean (i.e., 7.3 % of tweets are classified as commitments and 13.6 as populist-rhetoric).²⁹

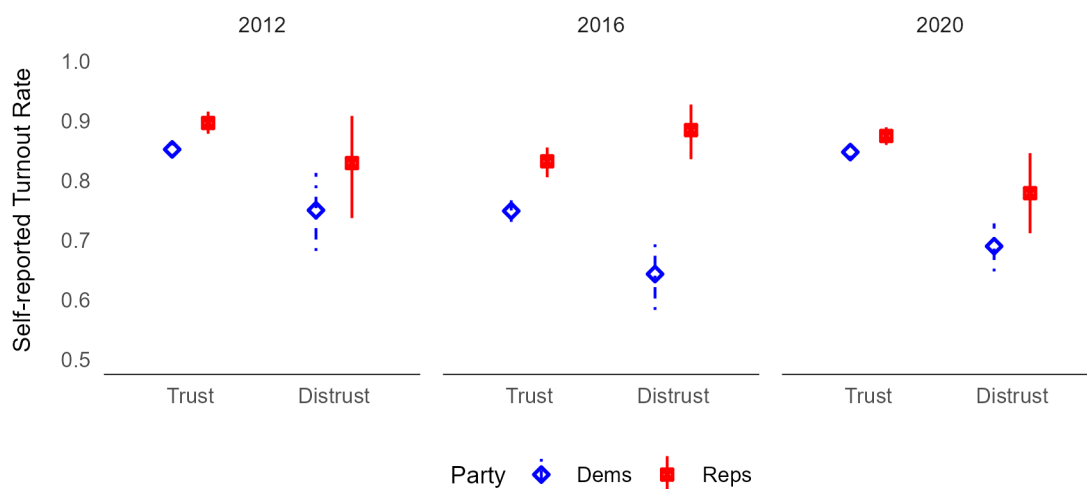
Besides the positive effect of distrust on the probability of commitment and populist rhetoric tweets, we expect commitments to be more likely when the candidate tweets about topics on which voters have strong convictions. We therefore expect the average effect of high conviction topic – i.e., holding distrust fixed – to be positive. We test this prediction by adding high conviction topic to Equation 3. As shown in Table 4, we detect 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 holding distrust fixed.

5.3 Evidence on Turnout

Proposition 2 shows that if distrust increases among voters, so do the chances of election of a committed candidate, while the opposite is true for a candidate without commitments. We give empirical support to this prediction by looking at the effect of the candidacy of Donald Trump, an uncontroversial example of a committed and populist candidate (e.g., building a

²⁹In Figure F.3 in the appendix we show how the results are robust to using different time windows, from 60 to 20 days before and 60 to 20 days after the precise election day. Interestingly, we find that the effect on commitment persists in the post-election period, when winning candidates re-state their commitments to voters, whereas the effect of distrust on populist rhetoric shrinks towards zero after election day.

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).

wall, America first, and trade war with China), on turnout of distrustful Republicans. As 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 70% more commitment tweets and 7% more populist rhetoric tweets compared to Clinton.³⁰

For this test, we use ANES data on self-reported party identification, distrust, and turnout across the three waves/presidential elections used to estimate distrust at the district level, one without Trump (2012) and two when Trump was running as the Republican candidate. ANES self-reported turnout – as is generally the case for self-reported measures of turnout (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%.³¹

In Figure 3, we display the self-reported turnout rate by party, election year, and level of trust in government. Consistent with established literature on trust and political participation (Citrin 1974), turnout is generally lower for more distrustful voters. However, in 2016 the relationship flips for republicans with low trust. 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.

³⁰Donal Trump's tweets downloaded from <https://github.com/MarkHershey/CompleteTrumpTweetsArchive/tree/master>.

³¹Data from <https://www.electproject.org/>, accessed on 29 March 2023.

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

$$\begin{aligned}
Pr(\text{Turnout}_{idt} = 1) = & \eta_d + \delta_t + \beta_1 \text{Republican}_{idt} + \beta_2 \text{Distrust}_{idt} \\
& + \beta_3 \text{Republican}_{idt} \times \text{Distrust}_{idt} + \beta_4 \text{Republican}_{idt} \times \delta_t + \beta_5 \text{Distrust}_{idt} \times \delta_t \\
& + \beta_6 \text{Republican}_{idt} \times \text{Distrust}_{idt} \times \delta_t + \nu \mathbf{X} + \epsilon_{idt}
\end{aligned} \tag{4}$$

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 5 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.³² Compared to 2012, distrustful Republicans in 2016 are 13-15 percentage points more likely to turn out to vote, during the first Trump candidacy. The effect remains positive but is noisier for the 2020 elections. There 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 points to an interesting avenue for future research, where temporal dynamics and the probability of delivering on the promised commitments enter voters’ incentives to turn out to vote in the presidential elections.

³²For full regression table, see Table F.12 in the appendix.

Table 5: Effect of Trump candidacy on Turnout of Distrustful Republicans.

	Pr(Turnout = 1)		
	(1)	(2)	(3)
Republican × Distrust	0.019 (0.057)	0.002 (0.058)	-0.009 (0.056)
Republican × Distrust × EY 2016	0.134* (0.074)	0.146* (0.075)	0.150** (0.072)
Republican × Distrust × 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

6 Commitment and the Illiberal Consequences of Populism

In this section we show how the baseline model can be easily generalized to obtain a clear connection with three other important characteristics of populism: aversion to checks and balances, lower competence, and fake-news production.

6.1 Commitment Politics and the Elimination of Checks and Balances

There is ample evidence showing how, once elected, several populist leaders sought to capture (Orbán in Hungary), dismantle (Fujimori in Peru), sabotage (Trump in the US), and reform (Blocher in Switzerland) the state bureaucracy (Bauer & Becker 2020). Sasso & Morelli (2021) and Bellodi et al. (2023) provide theoretical and causal findings in support of populists’ preference for less competent bureaucrats, who are less likely to resist the populist agenda. Besides undermining the bureaucracy, populists have often attacked free media and the judiciary, delegitimizing such institutions on the basis of their unelected nature and their interference in government. These illiberal consequences of populism seem to match the preferences of segments of voters who support populist parties. Survey evidence shows that, even though “populist” voters are supportive of democracy in theory, they are highly dissatisfied with how democracy works in practice (Rovira Kaltwasser & Van Hauwaert 2020) and are highly critical of the mediating role of political parties and judicial constraints on the executive (Zaslave & Meijers 2023).

To see the connection between the shift to commitment politics and the erosion of checks and balances, we present a simple extension of our baseline principal-agent model. Any “agency of restraint”³³ can be modeled as a player A that can intervene only after the policy-maker has chosen the policy (at time 2 in the original timing). In particular, assuming that z represents the status quo and h and l two possible reforms, we can assume that A can only act as a veto player in case the policy-maker chooses h or l . When the policy is blocked, the status quo z remains. However, whenever the agency wants to block a reform, it succeeds with probability k . The principal assigns a positive probability ρ that the elite manages to influence A .

Consider the case in which the policy maker commits to policy h . The expected utility for a principal is

$$U_{comm}^{cb} = -\gamma_l - \gamma_z \left(\left(\frac{1-\lambda}{2} \right) \rho + \left(\lambda + \frac{1-\lambda}{2} \right) (1-k) \right) - \gamma_h \lambda k \rho. \quad (5)$$

Let us understand the three terms of the above expected utility under commitment to h when there are checks and balances (cb in short in the superscript). First, when l is optimal, commitment to h always leads to a loss, whether or not the agency intervenes. Second, when the status quo z is optimal, there is a loss when (i) the elite prefers h (probability $\frac{1-\lambda}{2}$) and manages to influence the agency of restraint to not block the reform (probability ρ) and (ii) the elite has no interest to influence the agency of restraint because its preferred policy is either z or l (probability $\lambda + \frac{1-\lambda}{2}$), but the agency does not manage to block the reform (probability $1-k$). Third, when h is optimal, there is a loss when at the same time it is true that the elite prefers z (probability λ) and manages to influence the agency to block the reform (probability ρ), and the agency is successful at blocking the reform (probability k).

The derivative of this utility with respect to k tells us if a principal with a committed agent prefers to have more or less checks and balances

$$\frac{\partial U_{comm}^{cb}}{\partial k} = \gamma_z \left(\lambda + \frac{1-\lambda}{2} \right) - \gamma_h \lambda \rho.$$

Thus, if $\gamma_h > \gamma_z \left(\frac{1+\lambda}{2\rho\lambda} \right)$ (i.e., the principal believes that the probability that h is optimal is sufficiently larger than the probability that z is optimal), the derivative is negative and the principal prefers to reduce the power of the agency: the higher are ρ and λ , the lower is γ_z , the more likely the principal prefers to reduce the agency’s power. It follows that populist rhetoric, even if it is limited to affect λ and γ_z , and without explicitly attacking the agency (increase in

³³The term “agency of restraint” can be used interchangeably with reference to free media, independent civil servants, or independent judiciary, for all of these institutions impose checks on executive decisions and balance executive power. For simplicity, in this section we lump them all into a single construct.

ρ), induces voters to prefer a reduction of the agency's power.

The expected utility of the principal under full delegation and without checks and balances is

$$U_{del} = -p \left((\gamma_h + \gamma_l) \frac{1 + \lambda}{2} + \gamma_z (1 - \lambda) \right). \quad (6)$$

The principal suffers a utility loss when the elite manages to distort the agent's decision (probability p) and the optimal policy is not the elite's preferred policy.

In presence of the agency, we assume that the elite coordinates its efforts to obtain the preferred policy trying to influence the agency when it fails to influence the decision maker. In particular, when the optimal policy is either h or l and the elite's preferred policy is z , the elite distorts the agent's decision with probability p , but in case it fails (probability $1 - p$), it can still try to influence the agency and block the reform (with probability ρk). In this case, the presence of the agency negatively affects the principal's payoff because it represents a further opportunity for the elite to influence policy. When, instead, the optimal policy is z and the elite's preferred policy is either l or h , the agency with probability $(1 - \rho)k$ can block the bad reform. In this case, the agency limits the bias that the elite determines for the decision maker.

The expected utility of the principal under full delegation and with checks and balances is

$$U_{del}^{cb} = -(\gamma_h + \gamma_l) \left(p \frac{1 + \lambda}{2} + \lambda(1 - p)\rho k \right) - \gamma_z (1 - \lambda) p ((1 - \rho)(1 - k) + \rho).$$

Similarly to the case of commitment, we can look at the derivative of the utility from full delegation with respect to k

$$\frac{\partial U_{del}^{cb}}{\partial k} = -(1 - \gamma_z)\lambda(1 - p)\rho + \gamma_z(1 - \lambda)(1 - \rho)p,$$

which is positive if

$$\gamma_z \geq \frac{\lambda(1 - p)\rho}{\lambda(1 - p)\rho + (1 - \lambda)(1 - \rho)p}.$$

Notice that this threshold goes to zero for p that tends to 1, while the threshold for which more checks and balances increase the utility from full commitment remains positive. Hence, when there is sufficiently large distrust towards politicians, there is a non-empty set of parameters $(\rho, \lambda, \gamma_z)$ such that a principal would prefer more checks and balances in case the agent has full delegation, while they would prefer fewer checks and balances in case of full commitment.

6.2 Incompetence, Fake News, and Information Incentives

In our simple baseline model, the delegated politician sees the realization of the state of the world with probability 1, while the principal never sees the state of the world. Relaxing these two simplifying assumptions yields two additional important consequences of the shift to commitment politics for voters' demand of competence and candidates' supply of misinformation, both consistent with the populism literature and anecdotal evidence on the campaign strategies of populist leaders.³⁴

First, the relaxation of perfect understanding of the state of the world by agents yields a clear difference in terms of the desirability of competence. Once a principal has chosen a commitment, she does not care about competence as much as a principal who chose delegation. Second, if the principal can make a costly effort to figure out the state of the world (to make it costlier for the agent to choose the wrong policy), she will do so more under delegation, while she has no interest in doing so under commitment. Therefore, lower information by voters can be a consequence rather than a cause of populism.

Finally, consider a situation in which there is electoral competition between a committed candidate and a fully delegated candidate. If the fully delegated candidate wins the election, voters in equilibrium may want to exert costly effort to figure out the state of the world. As a result, the committed candidate may benefit from making it costlier for voters such process of information acquisition. A common strategy for undermining information acquisition by voters is to spread fake news. This element of the committed candidate's campaign reduces the expected utility of electing a fully delegated candidate hence demobilizing their constituents.

7 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 a preference for candidates making 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. 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. We have presented the demand side of this shift with a simple principal-agent model and the

³⁴These additional and complementary components of the populist strategy are also addressed formally in Morelli et al. (2021).

corresponding changes on the supply side with an electoral competition model with endogenous turnout. In addition, we have tested the key predictions of our theory with novel evidence on U.S. Congressional candidates and campaigns. Finally, we have drawn the connection between the shift to commitment politics and many other features of illiberal populism discussed by scholars in various fields.

While the principal-agent results on the connection between distrust and the demand of commitment are general and not limited to specific institutional configurations, the balance of incentives on the supply side between the different platforms may depend on electoral systems and other contextual details. Majoritarian electoral systems, for instance, might be an important scope condition of our findings. Clearly, an extension of the empirical analysis to countries with different electoral systems would be harder, especially in the absence of clear geographic representation. However, there is plenty of anecdotal evidence that confirms the importance of the shift to commitment politics even in countries with mixed electoral systems. Italy, a non-majoritarian democracy where candidates do not need to win a specific geographic constituency, offers several examples of commitment candidates across the political spectrum. Berlusconi's contract with the Italians and the Five-Star Movement's commitments to citizenship income and a "binding mandate" – which emblematically represents a commitment to stick to the original commitments – are consistent with the idea that commitment politics grows where distrust flourishes.

In this paper, we have shown the logic and implications of the growing demand and supply of policy commitments holding constant their credibility, and indeed the empirical evidence represents a conservative test of our theory, given the fact that, in reality, people do not believe every commitment and promise to be credible. An interesting avenue for future research could incorporate dynamics between the credibility and the fulfillment of such commitments, as well as candidates' ex-ante (ex-post) attempts to increase the perception of credibility (fulfillment) of commitments. Finally, 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 can also explain the crisis of representative democracy in terms of preferences for reduced checks and balances and for illiberal reforms more generally. Hence, the dynamic consequences and conditions for the persistence versus reversibility of the shift to commitment politics and its consequences is an important path for future research.

References

- Acemoglu, D., Egorov, G. & Sonin, K. (2013), ‘A Political Theory of Populism’, *Quarterly Journal of Economics* pp. 771–805.
- Acemoglu, D. & Restrepo, P. (2020), ‘Robots and jobs: Evidence from US Labor Markets’, *Journal of Political Economy* **128**(6), 2188–2244.
- 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. (1988), ‘Credibility and Policy Convergence in a Two-Party System with Rational Voters’, *American Economic Review* **78**(4), 796–805.
URL: <https://econpapers.repec.org/RePEc:aea:aecrev:v:78:y:1988:i:4:p:796-805>
- 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>
- 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.
URL: <https://www.aeaweb.org/articles?id=10.1257/app.20210736>
- Amador, M., Werning, I. & Angeletos, G.-M. (2006), ‘Commitment vs. Flexibility’, *Econometrica* **74**(2), 365–396.
- Aronow, P. M. & Samii, C. (2016), ‘Does Regression Produce Representative Estimates of Causal Effects?’, *American Journal of Political Science* **60**(1), 250–267.
- Autor, D., Dorn, D., Hanson, G. & Majlesi, K. (2020), ‘Importing political polarization? The electoral consequences of rising trade exposure’, *American Economic Review* **110**(10), 3139–3183.
- 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.
URL: <https://aclanthology.org/2020.findings-emnlp.148>
- Bauer, M. W. & Becker, S. (2020), ‘Democratic Backsliding, Populism, and Public Administration’, *Perspectives on Public Management and Governance* **3**(1), 19–31.
- Bellodi, L., Morelli, M. & Vannoni, M. (2023), ‘A Costly Commitment: Populism, Government Performance, and the Quality of Bureaucracy’, *American Journal of Political Science* pp. 1–17.
- 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 Post-stratification through Machine Learning (autoMrP)’, *Journal of Politics* **84**(1), 597–601.
- Bueno de Mesquita, E. & Friedenber, A. (2011), ‘Ideologues or Pragmatists?’, *Journal of the European Economic Association* **9**(5), 931–951.
- Callander, S. & Wilkie, S. (2007), ‘Lies, damned lies, and political campaigns’, *Games and Economic Behavior* **60**(2), 262–286.
- Carrillo, J. D. & Dewatripont, M. (2008), ‘Promises, Promises, . . .’, *Economic Journal* **118**, 1453–1473.
- Cha, J., Kuriwaki, S. & Snyder, J. M. (2021), ‘Candidates in American General Elections’, *Harvard Dataverse* .
- Citrin, J. (1974), ‘Comment: The Political Relevance of Trust in Government’, *American Political Science Review* **68**(3), 973–988.
URL: <https://www.jstor.org/stable/1959141>
- 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.
- Corazzini, L., Kube, S., Maréchal, M. A. & Nicolò, A. (2014), ‘Elections and deceptions: An experimental study on the behavioral effects of democracy’, *American Journal of Political Science* **58**(3), 579–592.
- Crutzen, B. S. Y., Sisak, D. & Swank, O. H. (2020), Left Behind Voters, Anti-Elitism and Popular Will.
- Dai, Y. & Kustov, A. (2022), ‘When Do Politicians Use Populist Rhetoric? Populism as a Campaign Gamble’, *Political Communication* **39**(3), 383–404.
URL: <https://doi.org/10.1080/10584609.2022.2025505>
- Di Cocco, J. & Monechi, B. (2021), ‘How Populist are Parties? Measuring Degrees of Populism in Party Manifestos Using Supervised Machine Learning’, *Political Analysis* pp. 1–17.
- 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.
URL: <https://doi.org/10.1016/j.jce.2018.03.001>
- 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.
- 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.

- Ford, R. & Jennings, W. (2020), 'The Changing Cleavage Politics of Western Europe', *Annual Review of Political Science* **23**, 295–314.
- 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.
URL: <https://www.aeaweb.org/articles?id=10.1257/aer.20202045&from=f>
- 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. (2021), Mobilization and the Strategy of Populism Theory and Evidence from the United States.
URL: https://cepr.org/active/publications/discussion_papers/dp.php?dpno=15686
- 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).
URL: <https://www.pnas.org/doi/abs/10.1073/pnas.2305016120>
- Graham, M. H. & Svobik, 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., 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.
- Haghtalab, N., Jackson, M. O. & Procaccia, A. D. (2021), 'Belief polarization in a complex world: A learning theory perspective', *Proceedings of the National Academy of Sciences* **118**(19), e2010144118.
URL: <https://www.pnas.org/doi/abs/10.1073/pnas.2010144118>
- 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.
- Kuriwaki, S. (2022), 'Cumulative CCES Common Content'.
URL: <https://doi.org/10.7910/DVN/II2DB6>

- Laitin, D. D. (2018), ‘Populism: the Muslim Connection’, *Keynote Address at the CESifo Conference “Globalization and Populism - Past and Present”*.
URL: <http://laitinresearch.stanford.edu/wp-content/uploads/2013/10/PopulismCESkeynoteDL.pdf>
- Levy, G. (2007), ‘Decision making in committees: Transparency, reputation, and voting rules’, *American Economic Review* **97**(1), 150–168.
- Levy, G., Razin, R. & Young, A. (2022), ‘Misspecified Politics and the Recurrence of Populism’, *American Economic Review* **112**(3), 928–62.
URL: https://personal.lse.ac.uk/YoungA/Populism_Jan27.pdf
- 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.
- Morelli, M., Nicolò, A. & Roberti, P. (2021), A Commitment Theory of Populism.
URL: <https://cepr.org/publications/dp16051>
- Mudde, C. (2004), ‘The Populist Zeitgeist’, *Government and Opposition* pp. 541–563.
- Mudde, C. & Rovira Kaltwasser, C. (2018), ‘Studying Populism in Comparative Perspective: Reflections on the Contemporary and Future Research Agenda’, *Comparative Political Studies* **51**(13), 1667–1693.
URL: <https://doi.org/10.1177/0010414018789490>
- Norris, P. & Inglehart, R. (2019), *Cultural Backlash: Trump, Brexit, and Authoritarian Populism*, Cambridge University Press, Cambridge.
- 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. (2018), ‘Populism and the economics of globalization’, *Journal of International Business Policy* **1**(1-2), 12–33.
URL: <https://doi.org/10.1057/s42214-018-0001-4>
- Rooduijn, M., Van Kessel, S., Froio, C., Pirro, A., De Lange, S., Halikiopoulou, D., Lewis, P., Mudde, C. & Taggart, P. (2019), ‘The PopuList: An Overview of Populist, Far Right, Far Left and Eurosceptic Parties in Europe’.
URL: www.popu-list.org
- Rovira Kaltwasser, C. & Van Hauwaert, S. M. (2020), ‘The populist citizen: Empirical evidence from europe and latin america’, *European Political Science Review* **12**(1), 1–18.
- Sasso, G. & Morelli, M. (2021), ‘Bureaucrats under Populism’, *Journal of Public Economics* **202**, 104497.
URL: <https://doi.org/10.1016/j.jpubeco.2021.104497>
- 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.
- Studies, A. N. E. (2013), ‘Anes 2012 time series study full release [dataset and documentation]’.
URL: www.electionstudies.org
- Studies, A. N. E. (2017), ‘Anes 2016 time series study full release [dataset and documentation]’.
URL: www.electionstudies.org
- Studies, A. N. E. (2021), ‘Anes 2020 time series study full release [dataset and documentation]’.
URL: www.electionstudies.org

- Tausanovitch, C. & Vavreck, L. (2021), ‘Democracy Fund + UCLA Nationscape Project, October 10-17, 2019 (version 20211215)’.
URL: <https://www.voterstudygroup.org/data/nationscape>
- Urbanati, N. (2019), ‘Political theory of populism’, *Annual Review of Political Science* **22**, 111–127.
- Van Weelden, R. (2013), ‘Candidates, credibility, and re-election incentives’, *Review of Economic Studies* **80**(4), 1622–1651.
- Warshaw, C. & Rodden, J. (2012), ‘How should We measure district-level public opinion on individual issues?’, *Journal of Politics* **74**(1), 203–219.
- Weyland, K. (2001), ‘Clarifying a Contested Concept: Populism in the Study of Latin American Politics’, *Comparative Politics* **34**(1), 1–22.
URL: <https://www.jstor.org/stable/422412>
- Weyland, K. (2017), Populism : A Political-Strategic Approach, in C. R. Kaltwasser, P. Taggart, P. Ochoa Espejo & P. Ostiguy, eds, ‘The Oxford Handbook of Populism’, Oxford University Press.
- Zaslove, A. & Meijers, M. (2023), ‘Populist democrats? unpacking the relationship between populist and democratic attitudes at the citizen level’, *Political Studies* .
URL: <https://doi.org/10.1177/00323217231173800>

Appendix

A	Theory Proofs	A1
B	District-level Estimates of Distrust in Government	A4
	B1 Comparing ANES with NationScape MRP Estimates	A7
C	Obtaining Candidates' Twitter Handles	A9
D	Measuring Commitment and Populist Rhetoric in Tweets	A10
E	Topic Detection	A15
F	Robustness Tests	A19

A Theory Proofs

Proof of Corollary 1. The signs of the derivative of p_1 with respect to γ_z and λ follow from inspection of the expression for p_1 . Consider the derivative of p_2 with respect to γ_z . The derivative of the denominator is $-\lambda + (1 - \lambda)/2$, which is negative. The derivative of the numerator crucially depends on additional assumptions of how an increase in γ_z affects γ_l and γ_h . It can be assumed that either γ_h becomes lower or is not affected by an increase in γ_z . In either scenario the derivative of p_2 with respect to γ_z is positive. The derivative of denominator of p_2 with respect to λ is $1 - \gamma_z - 1/2 - \gamma_z/2 = 1/2 - \frac{3}{2}\gamma_z$ which is larger than zero if $\gamma_z \leq \frac{1}{3}$. ■

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 winning of a candidate 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 $EU^{-G}(\hat{S}) < EU^{-G}(S')$.³⁵ For a given μ^G , the probability that candidate running with \hat{S} wins is

$$\begin{aligned} \mathbb{P}(\underline{c}_v^G > \underline{c}_v^{-G}) &= \mathbb{P}\left(\mu^{-G} > EU^{-G}(e') - EU^{-G}(e) - EU^G(e) + EU^G(e') + \mu^G\right) = \\ &= \frac{1}{2} + \psi\left(-EU^{-G}(S') + EU^{-G}(\hat{S}) + EU^G(\hat{S}) - EU^G(S') - \mu^G\right) \end{aligned} \quad (7)$$

that is, more voters show up for 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

$$\frac{1}{2} + \psi\left(-EU^{-G}(S') + EU^{-G}(\hat{S}) + EU^G(\hat{S}) - EU^G(S')\right), \quad (8)$$

because μ^G has mean 0.

Consider now the primary elections of party G . Consider any two candidates of party G ,

³⁵The case in which the electoral platform \hat{S} is preferred by all citizens is trivial.

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 policy 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 the platform \hat{S}^g that maximizes the policy utility.³⁶ 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. First, suppose member v of party G votes or abstain at the general elections both in case g , the winner of the primary elections of party G , proposes e^{g^*} and in case g proposes \hat{e}^g . In this case v prefers that e^{g^*} is the winning electoral platform at the primary elections because it gives higher policy utility. Second, suppose citizen v votes if candidate g proposes platform e^{g^*} , while she does not vote if g proposes platform \hat{e}^g , i.e. $EU^G(e^{g^*}) - EU^G(e^{-g}) \geq c_v^G + \mu^G$ and $EU^G(\hat{e}^g) - EU^G(e^{-g}) < c_v^G + \mu^G$. Member v 's utility is higher if candidate proposes e^{g^*} than in case g proposes \hat{e}^g . Therefore, the candidate proposing e^{g^*} wins the primary elections. Competition at the primary elections guarantees that there exists at least one candidate proposing e^{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 2.

Suppose candidate g at the general elections has platform $\hat{S}^G = h$. Turnout of party G depends positively on the payoff they receive by their own candidate's platform and negatively on the payoff they receive by the opponent's platform. If both candidates propose full com-

³⁶Consider 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.

mitment, the payoffs from the two platforms do not depend on p^G so there is no effect of this parameter on turnout. If $-g$ proposes partial commitment, then turnout of party G increases with p^G because (i) the payoff of platform \hat{S}^g does not depend on p^G , (ii) the payoff of partial commitment platform depends negatively on p^G . For the same reason the proposition holds when the opponent proposes full delegation. Finally, the same reasoning applied to the turnout of party $-G$ explains the comparative statics with respect to p^{-G} . ■

Proof of Proposition 3. The utility from partial commitment decreases with γ_z^G and increases with λ . The utility from full commitment decreases with γ_z^G . Finally, the utility from delegation increases with γ_z^G because $\lambda \geq \frac{1}{3}$ and decreases with λ because $\gamma_z^G \leq \frac{1}{3}$. Given that the turnout of a candidate with electoral program S against a candidate with electoral program S' positively depends on the difference $EU^G(S) - EU^G(S')$, the proof follows immediately. ■

Proof of Proposition 4. Suppose, without loss of generality, that party G 's optimal policy is h . Notice that p_1 weakly decreases with γ_h^G , if an increase in γ_h^G negatively affects γ_z^G . Similarly p_2 decreases with γ_h^G , with a primary effect going through the numerator, and a secondary effect through a possible decrease in γ_z^G in the denominator. Thus, by Proposition 1 and Lemma 1 the first part of Proposition 4 follows. The second part of the Proposition follows by inspection of the utility from full commitment. ■

B 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 Twitter activity of candidates, which was minimal in the 2008 elections.

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.

- -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. We include also random effects for the state.

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 B.1: Individual Predictors of Distrust.

	Distrust			
	(1)	(2)	(3)	(4)
Sex Category [1-2]	-0.018*** (0.004)	-0.020*** (0.005)	-0.021*** (0.004)	-0.021*** (0.005)
Age [1-5]	-0.014*** (0.002)	-0.014*** (0.002)	-0.014*** (0.002)	-0.016*** (0.002)
Education [1-5]	-0.009*** (0.002)	-0.008*** (0.002)	-0.011*** (0.002)	-0.011*** (0.002)
<i>District-level Covariates</i>				
% Black Population			-0.319** (0.125)	
Gini Index			0.402 (0.450)	
Median Income			0.013 (0.012)	
Employment Rate			0.340 (0.301)	
% of Republican Votes (State level)			0.057 (0.119)	
R ²	0.006	0.034	0.043	0.095
Observations	16,835	16,835	16,835	16,835
RMSE	0.282	0.278	0.276	0.269
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.

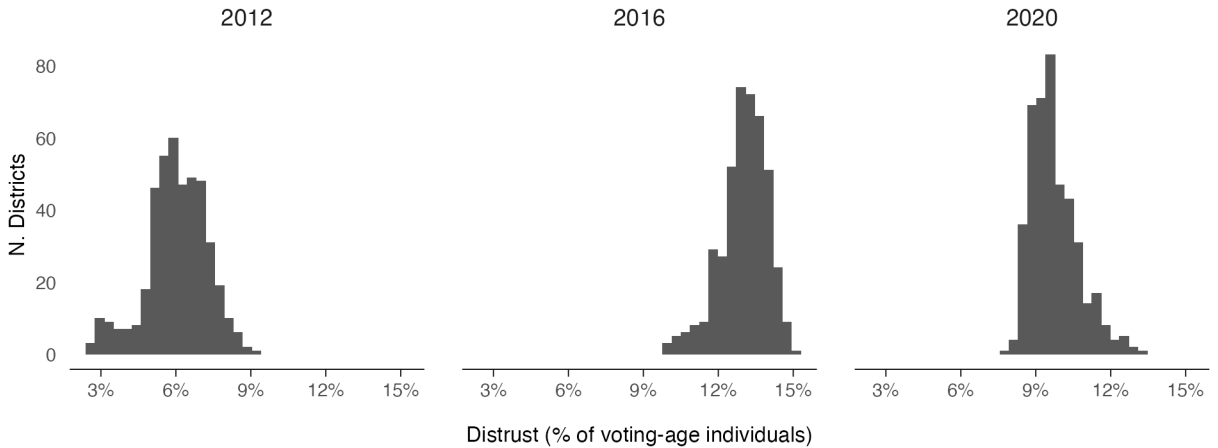
We then perform the MRP in two steps. First, we fit a multilevel model to the survey data including a set of random effects for J education categories, K sex categories, M age categories,

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.

$$\begin{aligned}
Pr(Distrust_i = 1) &= \Phi(\beta_0 + \alpha_{j[i]}^{education} + \alpha_{k[i]}^{sex} + \alpha_{m[i]}^{age} + \alpha_{n[i]}^{district} + \alpha_{z[i]}^{state} + \mathbf{X}_{n[i]}^T \beta) \\
\alpha_{j[i]}^{education} &\sim N(0, \sigma_{education}^2), \text{ for } j = 1, \dots, J \\
\alpha_{k[i]}^{sex} &\sim N(0, \sigma_{sex}^2), \text{ for } k = 1, \dots, K \\
\alpha_{m[i]}^{age} &\sim N(0, \sigma_{age}^2), \text{ for } m = 1, \dots, M \\
\alpha_{n[i]}^{district} &\sim N(0, \sigma_{district}^2), \text{ for } n = 1, \dots, N
\end{aligned} \tag{9}$$

In the second step, we calculate the predicted share of distrustful individuals for each ideal type based on Equation 9, 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 B.1 we report the distribution of our estimates of distrust over time.

Figure B.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.

B1 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, 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 B.2 report the question prompts with the possible answers.

Table B.2: Survey questions, ANES and NationScape.

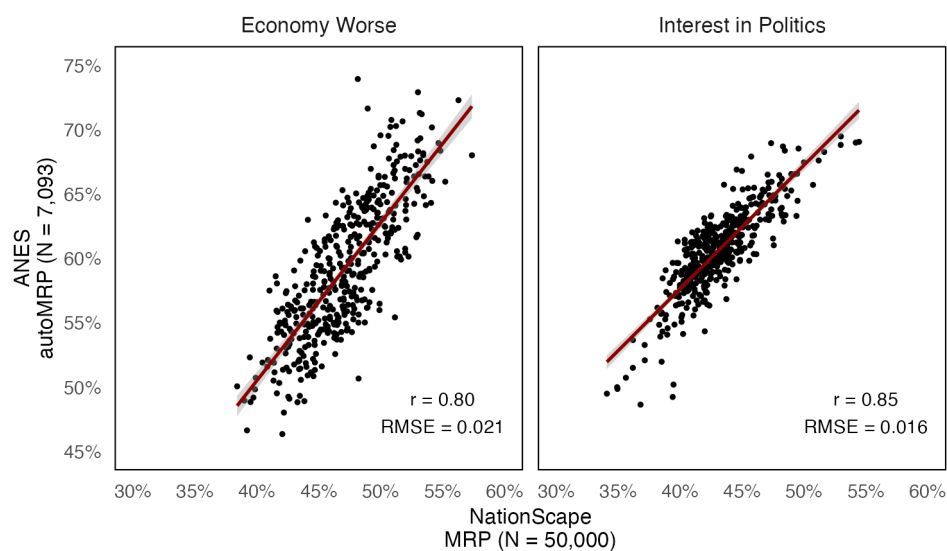
ANES	NationScape
<p>Now thinking about the economy in the country as a whole, would you say that over the past year the nation's economy has gotten better, stayed about the same, or gotten worse?</p> <ul style="list-style-type: none"> • -9. Refused • -8. Don't know • 1. Gotten better • 2. Stayed about the same • 3. Gotten worse <p>Sample mean: .61</p>	<p>Would you say that as compared to one year ago, the nation's economy is now better, about the same, or worse?</p> <ul style="list-style-type: none"> • 1. Better • 2. About the same • 3. Worse • . Respondent Skipped <p>Sample mean: .48</p>
<p>How often do you pay attention to what's going on in government and politics?</p> <ul style="list-style-type: none"> • -9. Refused • 1. Always • 2. Most of the time • 3. About half the time • 4. Some of the time • 5. Never <p>Sample mean: .63</p> <p>Respondents: 7,093</p>	<p>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...</p> <ul style="list-style-type: none"> • 1. Most of the time • 2. Some of the time • 3. Only now and then • 4. Hardly at all • . Respondent Skipped <p>Sample mean: .43</p> <p>Respondents: 414,318</p>

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. How-

ever, 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 B.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 B.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.

C 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;³⁷ and Politwoops, an organisation which tracks deleted tweets by public officials and maintain a list of active Twitter accounts.³⁸ 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.

³⁷Twitter handle for member of the 117th Congress available at <https://pressgallery.house.gov/member-data/members-official-twitter-handles>.

³⁸Dataset available at <https://www.propublica.org/datastore/dataset/politicians-tracked-by-politwoops>.

D 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 the average embedding of 30 examples of tweets containing policy commitment and populist rhetoric generated by ChatGPT. We gave ChatGPT 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 ChatGPT 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 ChatGPT are included in the replication files. in Table D.3 below we report a subset due to space limitations.

Table D.3: Texts generated by ChatGPT.

Sentences generated by ChatGPT
<i>Policy Commitments</i>
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
Combatting opioid addiction is a priority. I will support prevention programs, expand access to treatment, and work to address the root causes of this crisis. #EndOpioidAbuse
The student debt crisis is burdening our young people. I will advocate for student loan reform and explore solutions to make higher education more affordable. #StudentDebtRelief
I stand with our LGBTQ+ community. I will advocate for LGBTQ+ rights, support anti-discrimination legislation, and work towards achieving full equality. #LoveIsLove
As your future representative, I promise to fight for affordable healthcare for all. No one should go bankrupt due to medical bills. #HealthcareForAll
<i>Populist Rhetoric</i>
The political elite have failed us, but we're not going to give up. It's time to demand a government that represents the people, not just the powerful few.
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 thinks they know what's best for us, but they're out of touch with the struggles of everyday Americans. We need leaders who will put the people first.
The establishment wants to keep the status quo, but we know that real change only comes when the people demand it. Let's show them that we won't be silenced!
The elite have been making decisions that benefit themselves and their corporate backers, but it's time to put an end to their reign of power. We need leaders who will put the interests of the people first.
<i>Notes:</i> Subset of the statements generated by ChatGPT.

We estimate embeddings for each of the populist-rhetoric and commitment texts generated by ChatGPT 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 ChatGPT to annotate each of the 3,000 tweets based on the prompt reported in Table 2. 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 very 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 ChatGPT for classifying populist rhetoric and policy commitments is that, while ChatGPT 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 ChatGPT, besides being a cheap and efficient way of annotating text, also allows to limit human arbitrariness in the classification process.

To compare the validity of ChatGPT’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 ChatGPT (see Table 2). The coder and ChatGPT are in agreement 82% (commitment) 84% (populist rhetoric) of the times, which suggests that we can confidently use ChatGPT 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 the 75% of the tweets in the training dataset, and we leave the remaining 25% as held-out test set on which to assess the performance of the classifier.

In Table D.4 below, we report the performance metrics of each of the four classifiers. While all classifiers work well on 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 D.4: 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	0.859	0.860	0.867	0.851	0.879	0.859	0.910	0.850
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 D.5 and Table D.6 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, from (medium probability) from those where people and elites are portrayed as antagonistic groups.

Table D.5: 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 D.6: 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/KNiWyxds9P	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
<i>Notes:</i> Representative tweets with high and medium predicted probability of belonging to the populist-rhetoric class.		

E Topic Detection

Table E.7: 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
Environ.	enviro_airwateracts	2016, 2020	2
	enviro_carbon	2016, 2020	21
	enviro_mpg_raise	2020	1
	enviro_renewable	2020	2
Gun Reg.	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 Marr.	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 E.8: 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

Notes: Inferred topic from question displaying lowest standard deviation in responses.

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 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 → 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 positive 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) to and we assemble the eight topic-specific dictionaries reported in Table E.9 below.

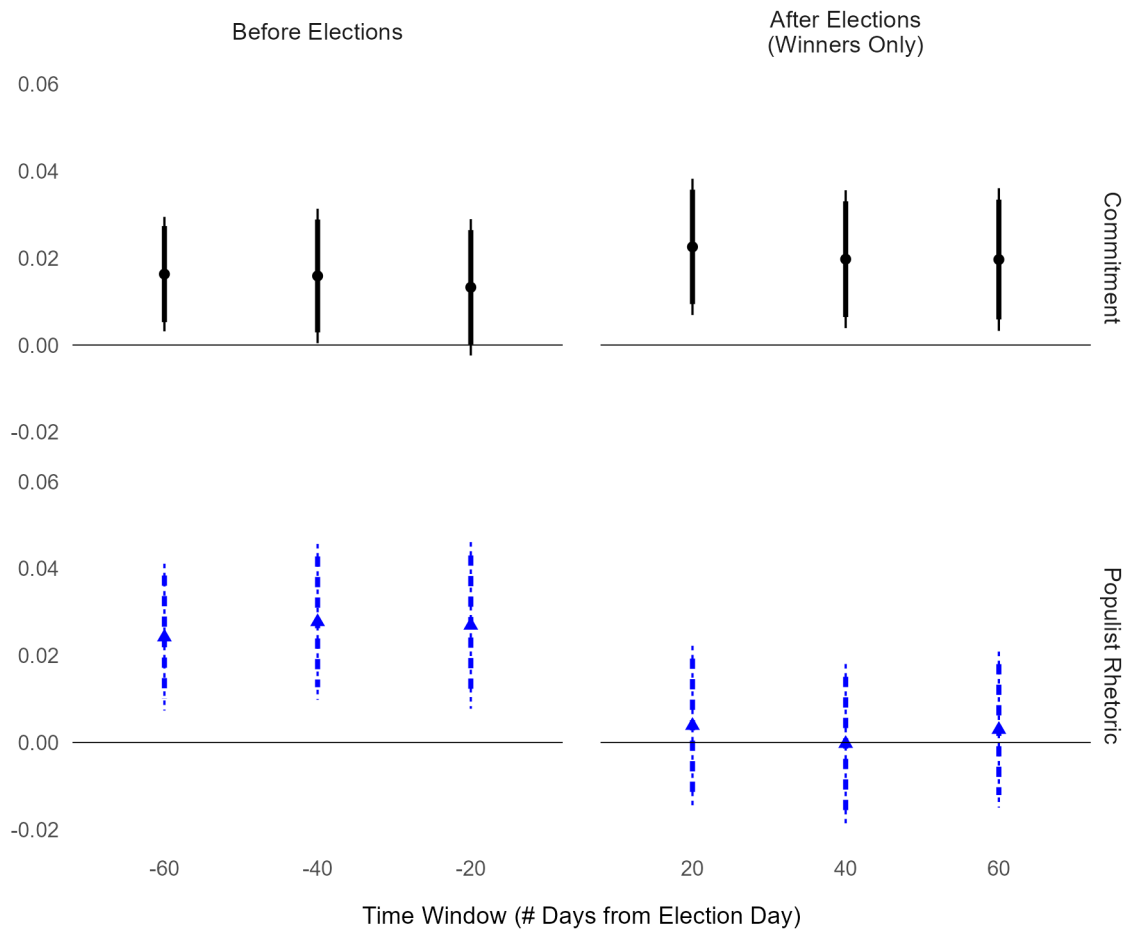
Table E.9: Topic specific dictionaries.

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, sanctuariescities, 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.

F Robustness Tests

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



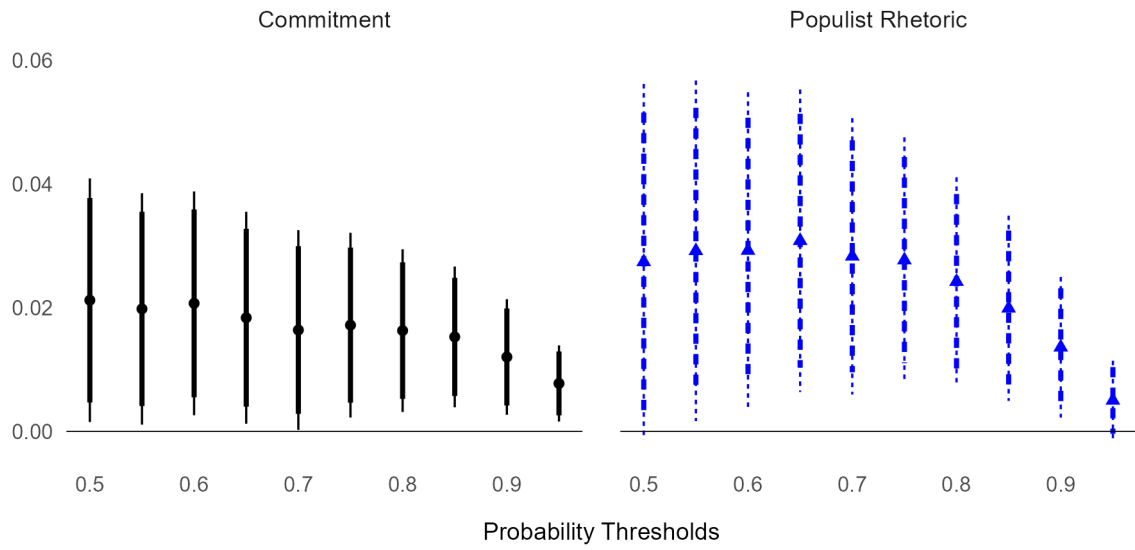
Notes: Estimated effect of distrust with 90 and 95% confidence intervals estimated across different windows of time expressed in days from election day. Specification includes candidate, year, district fixed effects and incumbency status as a covariate. Standard errors clustered by district.

Table F.10: Distrust, supply of commitments and populist rhetoric.

	Commitment			Populist Rhetoric		
	(1)	(2)	(3)	(4)	(5)	(6)
Distrust	0.021*** (0.006)	0.015* (0.009)	0.023*** (0.009)	0.026** (0.010)	0.023** (0.011)	0.031*** (0.011)
Controls: Party	✓			✓		
Controls: Incumbent	✓		✓	✓		✓
R ²	0.121	0.154	0.157	0.154	0.204	0.206
Observations	274,253	274,253	274,253	274,253	274,253	274,253
District FE	✓	✓	✓	✓	✓	✓
State-Year FE	✓		✓	✓		✓
Candidate FE		✓	✓		✓	✓
Year FE		✓			✓	

Notes: OLS estimates. SE clustered by district. DV is a predicted probability of commitment/populist tweets. Distrust is standardized share of distrustful individuals in each district. Signif. codes: ***: 0.01, **: 0.05, *: 0.1

Figure F.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, year, district fixed effects and incumbency status as a covariate. Standard errors clustered by district. Analysis performed on sample of tweets posted in the 60 days prior to election as in main specification.

Table F.11: Alternative measures of distrust.

	Commitment			Populist Rhetoric		
	(1)	(2)	(3)	(4)	(5)	(6)
Distrust - Govt. Do Right	0.017** (0.007)			0.025*** (0.009)		
Distrust - Big Interests		0.001 (0.003)			0.009* (0.005)	
Distrust - Politicians Corrupt			0.013* (0.007)			0.018** (0.009)
Controls: Incumbent	✓	✓	✓	✓	✓	✓
R ²	0.067	0.067	0.067	0.120	0.120	0.120
Observations	274,253	274,253	274,253	274,253	274,253	274,253
Candidate FE	✓	✓	✓	✓	✓	✓
District FE	✓	✓	✓	✓	✓	✓
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. 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

Table F.12: 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 × Distrust	0.019 (0.057)	0.002 (0.058)	-0.009 (0.056)
Republican × EY 2016	0.035* (0.020)	0.024 (0.020)	0.026 (0.020)
Republican × EY 2020	-0.018 (0.015)	-0.024 (0.016)	-0.018 (0.015)
Distrust × EY 2016	0.002 (0.045)	-0.002 (0.045)	-0.022 (0.044)
Distrust × EY 2020	-0.053 (0.039)	-0.052 (0.039)	-0.034 (0.038)
Republican × Distrust × EY 2016	0.134* (0.074)	0.146* (0.075)	0.150** (0.072)
Republican × Distrust × 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