Ideological Competition

Federica Izzo* Gregory J. Martin† Steven Callander‡

February 20, 2021

Abstract

We propose a model of political competition not over policy programs, but over ideologies: models of the world that organize voters' experiences and guide the inferences they draw from observed outcomes. Policy-motivated political parties develop ideologies, and voters choose the ideology that best explains their observations. Preferences over policies are then induced by the adopted ideology. Parties thus care about winning the ideological battle, as it confers an advantage in the electoral arena. We show that in equilibrium political parties always propose different models of the world. This divergence extends to all features of the environment, not just policy dimensions. A lower degree of policy extremism in the past increases the divergence on the policy dimension, thus leading to higher ideological polarization.

*University of California, San Diego. Email: fizzo@ucsd.edu.

†Stanford Graduate School of Business. Email: gjmartin@stanford.edu.

‡Stanford Graduate School of Business. Email: sjc@stanford.edu.

How do citizens form policy preferences? The question has been a central one in political science since the discipline's behavioral turn in the mid-20th century. Indeed, democratic citizenship confronts the mass public with a difficult problem: how to evaluate policy alternatives with which they may lack familiarity, experience, or basic understanding.

A half-century of research effort has produced a rich and varied set of answers. Scholars have variously argued that citizens have policy preferences that are temporally unstable, if they can be said to exist at all (Converse, 1964), formed on the spot when prompted (Zaller, 1992), formed by following elite cues (Brady and Sniderman, 1985; Lenz, 2013), or considered in reaction when politics intrudes into life unbidden (Klar and Krupnikov, 2016).

The classical formal treatments of political competition under complete information (Downs, 1957) on the other hand, sidestep the question of the origin of policy preferences, instead taking them as exogenous and fixed model primitives. This approach has proved fruitful for deriving predictions about the relationship of outcomes to the distribution of preferences in the population. But it considerably narrows the scope of political inquiry, and renders opaque much of the day-to-day work that political practitioners invest in campaigning, organizing, and honing rhetorical arguments.

In this paper we aim at bridging this gap, by offering a formal model of policy preference formation embedded in a political competition framework. Our model is a hierarchical one, in which citizens first choose ideologies — systems of belief about the underlying process that generates social outcomes — and then apply those ideologies to decide which policy inputs they prefer. Ideology thus provides a narrative that allows citizens to "navigate and orient themselves in the sea of politics" (Sartori, 1969). By providing a theory of how the world works, ideologies enable citizens to form expectations about the consequences of the various policy choices, which in turn inform their policy preferences. By their nature, such ideologically *induced* preferences are contingent and flexible, and subject to the kinds of temporal shifts and responses to elite signals that the behaviorists have documented.

This conception of ideology as distinct from and prior to policy preferences delivers a very different view of the nature of political competition among party leaders, which contrasts sharply with canonical formal treatments. Where the traditional approach views parties as strategically choosing platforms to assemble a winning electoral coalition, our model highlights the role of

parties in developing political narratives, with the aim of manipulating and directing the masses.[1] Importantly, the parties' ability to persuade voters does not come from an information asymmetry, but rather from their role as ideological entrepreneurs: parties have monopoly power over the development of ideologies.

Our model thus reverses the logic of standard formal models of elections. In our world, the parties' strategic problem is not solely to choose a policy platform that appeals to exogenous citizen preferences. Rather, political parties try to generate a favorable electoral environment by inducing citizens to adopt an ideology that translates into preferences aligned with the party's own (exogenous) policy program. The parties compete over ideology, not because they care about ideology per se but because ideology shapes the subsequent policy competition. A party that can win the ideological battle, so to speak, attains a favorable position in the policy competition. Ideology will thus be an equilibrium outcome of our model, rather than one of its primitives. The model allows us to address the following questions. To what extent is parties' ability to manipulate voters constrained in equilibrium? What are the determinants of ideological polarization? What features of the environment become "politicized," in the sense of being components of political ideologies? And finally, what is the relationship between rhetorical complexity and polarization?

MOTIVATING EXAMPLES

The hierarchical relationship between ideology and policy is not only an abstraction, but manifests in the actual behavior of political elites in democracies around the world. We note here two examples of such behavior in order to illustrate the key concepts that will appear in the model.

Thatcher In 1987, then-Prime Minister and leader of the UK Conservative party Margaret Thatcher famously stated her dis-belief in the existence of society: "[Too many people] are casting their problems on society and who is society? There is no such thing! There are individual men and women and there are families.[2]"

[1] This is akin to what Gramsci (1929) refers to as cultural hegemony: the elite construct ideological world-views to impose their belief systems to the people, and thus achieve and maintain control even without the use of coercive means.

[2] Thatcher (1993), pp. 627. Thatcher's statement echoed, in more prosaic terms, F.A. Hayek's attack on the very notion of the social as a "semantic fraud;" see discussion in Brown (2019), pp. 30.

The statement is remembered today for exemplifying Thatcher's unequivocal and uncompromising style of political communication. But it contains no reference to any particular policy or element of the Tory platform. It instead invites the audience to adopt a model of the causal forces that underlie the variation in outcomes they observe in the world. In Thatcher's telling, if one wants to understand why some people are rich and others poor, why some achieve their goals and others fail, one can safely ignore anything having to do with so-called social conditions. Individuals' skills, talents, character, and so on, are sufficient to explain and make sense of this variation. While there are no explicit references to policy in this statement, the policy implications of adopting it are profound. Policies such as jobs-training programs (to bolster individual skills), income tax cuts (to let talented individuals and families keep more of their earnings), and privatizations (because social ownership is pointless if there is no society) are all made legible by the "political rationality" (Lemke, 2001) it encapsulates.

Sanders Almost thirty years later, Vermont Senator Bernie Sanders launched a quixotic campaign for the Democratic presidential nomination, challenging the hegemonic centrist wing of the party, embodied in Hillary Clinton, from the left. Sanders' rhetoric offered a moral vision almost a mirror image of the one offered by Thatcher: "I believe. . . we are in this together. These are not just words. The truth is on some level when you hurt, when your children hurt, I hurt. And when my kids hurt, you hurt."[3]

Again, there is no policy content here; indeed, Sanders' campaign drew fire from some observers for being light on policy detail. Instead, the statement makes what amounts to an empirical claim: that our well-being (or the absence of "hurt," in Sanders' formulation) is not driven solely by our own income or consumption but by the well-being of everyone we interact with. This theory of the world provides the underpinnings for the kind of universalistic benefit and safety-net programs (Medicare for All, universal free public college, and so on) with which Sanders and the broader left wing of the Democratic party would become associated.

The key point in both examples is that, often, political rhetoric offers no *new* policy-relevant information. Neither statement clarifies the speaker's stance, describes features of any policy instrument, or presents novel facts to the audience. Instead, they organize or suggest how the

[3]Sanders (2016).

audience should be thinking about information it already has. The new organization suggested by the message, if adopted, logically leads to certain conclusions about which policies are preferable.

One way to think about such messages is as mere window-dressing: "signals" that reveal politicians' (verifiable) private information or demonstrate their willingness to absorb costs. Our model offers an alternative perspective, whereby such "fact-free" messages can be inherently persuasive, even absent any asymmetry of information. Politicians spend a great deal of time developing and honing their rhetorical messages, and our model takes seriously the idea that this distinctly *ideological* form of persuasion can be influential on voters' preferences.

OUTLINE OF THE MODEL

The model has three essential elements. First, we conceive of political parties as ideological entrepreneurs rather than policy entrepreneurs. Parties compete in an ideological contest by developing models of the world, with the goal of inducing citizens to adopt a favorable system of beliefs and, subsequently, influencing their policy preferences and electoral choices.

Second, our citizens are initially "ideologically innocent" (Kinder and Kalmoe, 2017), in the sense that they initially have no ideology of their own, and cannot or will not do the work of developing a theory of the world. In other words, parties have monopoly power over the construction of new ideologies. Citizens can evaluate an ideology if one is offered to them, but they cannot construct their own.

While this technological advantage gives parties the chance to influence citizens' (induced) preferences, this power is not unconstrained. The third key element of our model is a common-knowledge history of outcomes observed in the past, against which citizens will test the plausibility of parties' models. Our citizens adopt the model of the world that best resonates with their own lived experience and the past experience they absorb from the media and their schooling. This shared historical memory constrains and disciplines parties' ability to persuade.

Formally, our model describes a two-stage game, whose players are two policy-motivated parties and a representative citizen or voter. The voter cares exclusively about outcomes and has no intrinsic preference over the policy choice, but is uncertain about the mapping between policy and outcomes. Her policy preferences will therefore be induced by her beliefs over this mapping.

Specifically, we model outcomes as a linear function of the implemented policy, a set of non-policy covariates, and an idiosyncratic shock. At the beginning of the game, all players observe a public record of policy/covariate-outcome tuples. Each party proposes an ideology, i.e. a model of the world, to the voter. An ideology in our setting is just a vector of coefficients specifying the relationship between policy, covariates, and outcomes.

After privately observing the realization of today's policy outcome, the voter adopts the proposed ideology that best explains the observed history. Formally, the voter chooses the model under which the observed history of outcomes has the highest likelihood. In the second stage of the game, an election is held. In this election, each party is associated with a (fixed) policy position — in other words, the parties cannot credibly commit to a policy platform.

The voter's preferences over the two parties are generated by the ideology adopted in the first stage, which allows her to calculate expectations over outcomes under the different policy alternatives that the parties offer. Further, her evaluation of the parties is subject to a valence shock that realizes at the moment of her electoral choice. This structure implies that each party's objective in the first stage is to make the voter's adopted belief about the effect of the policy as large or as small as possible.

SUMMARY OF RESULTS

Our analysis shows that, while competition and the history of commonly observed outcomes limit their ability to induce voters to adopt a favorable ideology, parties are never fully constrained. In fact, the true model of the world never emerges in equilibrium, implying that parties are always able to manipulate voters' beliefs and preferences to some degree. Only in the limit where the obvserved history becomes infinitely long do parties converge on the same ideological message.

Second, we focus on the determinants of ideological polarization in equilibrium. In our context, ideological polarization is defined as the extent to which the models of the world proposed by the two parties differ in their predictions about how policy choices impact voters' welfare. We show that, surprisingly, equilibrium ideological polarization is not a function of the observed history of outcomes, but only of the history of policy. Specifically, ideological polarization decreases as past policies becomes more extreme, and vice versa.

The intuition is that as past policies become more extreme, outcomes become more informative for the voter. Historical data generated under more extreme polices thus represent a stronger constraint for political parties, and polarization decreases as both parties move towards the best-fitting model in the data. Changes in past outcomes, on the other hand, are absorbed into the parties' equilibrium ideologies. Outcome realizations that are favorable for one party translate into *both* parties shifting in the same direction, preserving both the level of polarization and balance in the probability of "winning" the ideological contest, but generating an advantage for the favored party in the electoral contest.

In contrast, the impact of today's policy extremism is in the opposite direction. As the current policy becomes more extreme, parties can gamble on extreme realizations of the outcome to improve their models' overall fit. Thus, parties can "afford" more radical ideologies, and polarization increases.

Finally, we analyze what happens when parties are allowed to include non-policy variables (over which they have no control or preferences) into their models of the world. We show that equilibrium ideologies in this multidimensional setting are always total: all available dimensions, including those that are not inherently political, are drawn in to parties' models and become sites of ideological conflict. Hence, even if the policy choice can be described by a single dimension, the political domain is inherently multidimensional. This resonates with work on affective polarization, such as Mason's (2018) study of the politicization of identity characteristics with no inherent connection to policy disagreements.

Our results illuminate the logic behind such social polarization, and allow us to make predictions about which identity attributes are likely to become woven into ideological disagreements, and when. We show that non-policy variables are incorporated in the parties' political narratives when their historical correlation with the policy dimension is sufficiently high. This allows parties to use the non-policy variables to improve the overall fit of their model, and thus to afford a higher degree of polarization on the policy dimension. The sign of the (sample) correlation with the policy variable determines the way that non-policy dimensions are sewed into each party's political narratives.

In our model parties have no control, or preferences, over non-policy dimensions. Thus, if a strong enough historical correlation with the policy variable emerges, it is purely by chance. The

results thus highlight the importance of path dependence and random events in the genesis of political ideologies. Political parties, acting as ideological entrepreneurs, will exploit emergent correlations to construct complex models of the world that support their favored policy position. This process generates a relationship between complexity and ideological polarization: the higher the number of politically salient dimensions (i.e., dimensions included in the proposed model of the world by at least one of the parties), the higher is polarization on the policy dimension in equilibrium.

Relationship to Existing Literatures

In the formal theory literature, ideology and policy preferences are typically collapsed into a single characteristic. This is not a requirement of formal theory per se, but of the often embedded assumption that voters are fully rational information processors and decision makers. Under this assumption, all information from history, from voters' experience, from the various political actors, and so on can be processed and collapsed into a final policy preference. We argue that ideology plays a distinct role in this process.

Our approach accords with the foundational work of Simon (1955) and the behavioralist tradition. That literature has emphasized the difficulty voters have in understanding politics and forming preferences over political objects. The mapping between policies and outcomes is fundamentally uncertain (Callander, 2011), and even voters who are well-informed about policies that have been tried in the past face a challenge of using this information to develop preferences about what policies to implement the future. The voter's difficulty is an instance of a more general insight that learning from data is a hard problem (Aragones et al., 2005). In practice, voters use short-cuts and processing tools to make sense of their environment, with the formation of ideology being one key step in this process. We take this reality seriously and build it into a formal model of politics. We use the tools of formal theory while making the underlying behavioral specification more realistic.

More specifically, our approach follows the behavioralist literature that emphasizes the role leaders and elites play in shaping voter's beliefs and preferences, in helping them make sense of the political world. We follow that literature in conceiving of preference formation as elite driven

and we provide a specific channel—the formation of ideologies—through which this occurs. This is, however, only a starting point in formalizing the preference formation process and gives rise to its own questions. Which elites do voters follow? What are the constraints, if any, on where leaders can lead followers? It is clear that voters are not completely blind. They do not follow elites regardless of what they say or do. Ample evidence exists that citizens process information and events themselves and that while they are influenced by elites, perhaps heavily, they are not sheep (Bullock, 2011).

The model we build offers answers to these questions. Voters decide which elite to follow through political competition, in which party leaders compete to influence voters and shape their beliefs and world view. This expanded perspective conceives of political competition more broadly, as an ongoing contest for political allegiance. This contrasts with the traditional, more narrow, view of parties as offering policy platforms once in each election cycle during the campaign season.

The ability of elites to shape voter beliefs and preferences is constrained by this competition. It is also constrained by reality. Elites cannot simply convince voters of anything. Voters check what they are offered against what they understand about the world and use this knowledge to evaluate the ideologies, the perspectives, on offer. This understanding and evaluation may be crude and imperfect, but it is sufficient to check the persuasive ability of elites to a non-trivial degree. Voters are led, to be sure, but only willingly and with open eyes.[4] The notion of ideology as shaping voters' interpretation of the observed reality distinguishes our work from Hafer and Landa (2007). Hafer and Landa (2007) also see ideology and beliefs as closely connected, but conceptualize ideology as a person's innate propensity to be persuaded by a left-wing or a right-wing argument.

The importance of reality as a check on elites is consistent with evidence from domains beyond politics. In sociology, the famous work by Goffman (1974) emphasized that people view an issue or a product through a particular lens and introduced the idea that advertisers aim to shape that lens rather than advertise features of a product directly. Goffman (1974) focused on the role of analogy in how advertising shapes the lens through which consumers view a product. We show

[4]Our formulation is consistent with dual-process theories of belief and preference change from psychology (Petty and Cacioppo, 1986), placing a sequential structure on that process. See Bullock (2011) and the cites therein for discussion of how voters combine their own information with elites cues.

how this idea applies to politics and emphasize the role ideology plays in mediating how voters view policy issues.[5]

Our work connects with a recent literature in behavioral economics that explores the role of models in decision making. In this literature, economic agents, as voters do in our model, use data to select a model through which to view the world. The behavioral limitations of agents in these models differ from those that affect voters in our model and we also differ in our focus on the competition among elites in shaping how citizens view the world (see, for example, Mullainathan et al. (2008) and Ortoleva (2012)).

We are also distinct from the strand of that literature that connects to political economy. In Levy et al. (2020) the distinction is between people with either complex or simple views of the world and not between elites and the masses. Eliaz and Spiegler (2020) focus on causal chains that form narratives about politics. Bénabou and Tirole (2006) explore citizens' view of how just the world is, developing a model of motivated reasoning in which voters distort or ignore information that doesn't fit with their preferred view of the world.

The Baseline One-Dimensional Model

The model describes a one-period, two-stage game. The players are two policy-motivated parties R and L, and a representative voter V. At the beginning of the game, the players observe a public record of T policies (z) and associated outcomes (y), indexed by $t \in \{-T, ..., -1\}$. As the notation suggests, we think of these as a history of past policy-outcome pairs, but this interpretation is not essential; the record might instead represent cross-sectional information from policies and outcomes in different countries or states. In addition, the voter privately observes the outcome of the (exogenous) policy that is in place in the first stage (z_0).

Voters in the model have no intrinsic preferences over z. Instead, they care exclusively about outcomes, with utility that is monotone in y. Preferences over policies will be fully determined by voters' beliefs about the relationship of policy to outcome.

[5]This was the inception of the field of "frame analysis," as Goffman's book was titled. In political science framing has traditionally been applied in a more narrow sense (see Chong and Druckman (2007) for a review). Our application is consistent with Goffman's original and broader meaning of framing as a lens or "model" through which the audience interprets data (Mullainathan et al., 2008, p.578).

We model the data generating process in a simple way: outcomes y_t are a linear function of the chosen policy z_t and a noise term ε_t.[6]

$$y_t = \beta_0^z z_t + \varepsilon_t \tag{1}$$

Where ε_t is i.i.d. from a normal distribution with mean 0 and variance σ^2. β_0^z here is a scalar representing the (expected) change in outcomes resulting from a one-unit change in the policy variable z. While the baseline model is one-dimensional, i.e., assumes that outcomes are only a function of a single (policy) variable, an extension presented below generalizes equation (1) to include a (possibly multi-dimensional) set of non-policy covariates.

The ideological contest. The first stage of the model is the ideological contest. The two parties move simultaneously, each proposing an ideology, i.e., a model of the world. In this one-dimensional setup, a model of the world is simply a scalar indicating the slope of the policy-mapping function. We denote the ideology proposed by party i as β_i^z. The voter faces some uncertainty about the data generating process underlying outcomes. Formally, we assume that the voter does not know the slope of the mapping function, i.e., the coefficient β_0^z.

We make two restrictions on the way that voters can use information from the past history in evaluating the parties' proposals. First, we do not specify the voter's complete beliefs over the true model of the world (i.e., the true β_0^z). Rather than the usual assumption of Bayesian voters, ours are frequentists. They will evaluate point models in terms of their likelihood, but will not hold a complete probability distribution over the full space of possible models.

Second, we do not allow the voter to freely choose her own ideology from $\mathcal{R}$. Instead, she must adopt one of the two ideologies proposed by the parties: the choice is a discrete rather than continuous one. This limitation is motivated by the classic activist's observation that political consciousness rarely forms spontaneously, but requires leadership to be expressed,[7] and the more modern social-scientific observation that most citizens invest little effort in paying attention to

[6]The policy dimension here is abstract and need not map one-to-one onto a real-world policy instrument. The important feature for our purposes is that it captures a dimension on which the parties disagree and over which they will have control once in power.

[7]E.g., Marx in *The Eighteenth Brumaire of Louis Bonaparte* analogized the French agrarian peasantry to a "sack of potatoes," atomized and unaware of their shared class position and material interests.

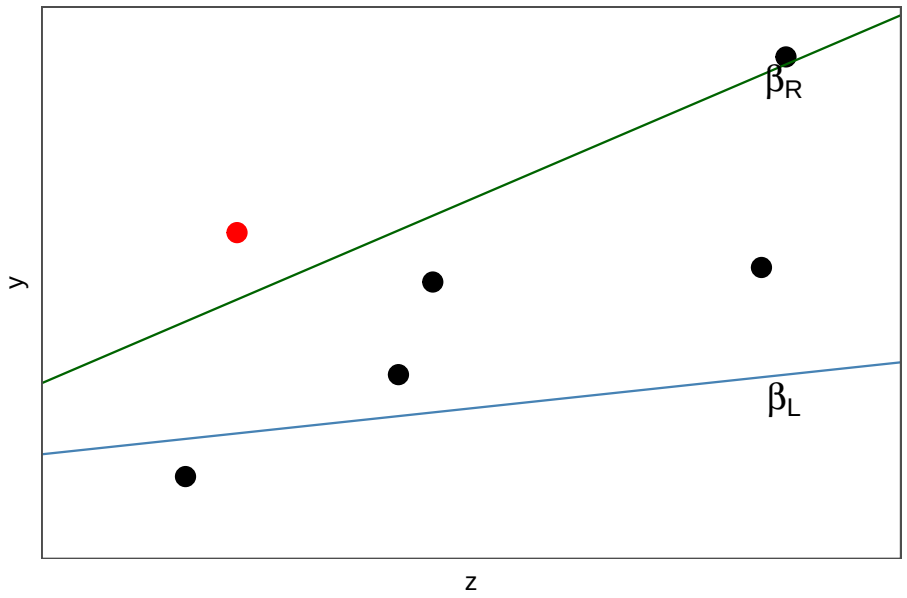

Figure 1: A visual depiction of the ideological contest in one dimension. Voter and both parties observe the policy-outcome pairs represented by the black dots; the voter privately observes an additional outcome realization at z_0, the policy in place at time 0 (the red dot). Parties L and R offer ideologies, in one dimension representable by scalars β_R^z and β_L^z (the slope of the line connecting policy and outcome).

politics or in developing their own political ideas. We defer further discussion of this crucial assumption, and of its possible interpretations and micro-foundations, to a separate section.

This combination implies that the voter will adopt the ideology, among those offered by the parties, which has the highest likelihood given the data she observes. Given the assumption that the noise term is normally distributed, this implies that the voter adopts the model of the world proposed by R if and only if

$$-\sum_{t=-T}^{0}(z_t\beta_R^z - y_t)^2 > -\sum_{t=-T}^{0}(z_t\beta_L^z - y_t)^2$$

Notice that the history of outcomes considered by the voter includes the past record of T outcomes and $T+1$ policy choices that are publicly observed, plus the outcome y_0 which is privately observed by the voter.

The ideological contest stage is described visually in Figure 1. In one dimension, the history takes the form of a scatterplot, and ideologies take the form of scalars (representing the slope of the line connecting policy and outcome).

The election. The second stage of the model is the electoral one. The voter casts her vote for one of the two parties. The winner then implements a policy $z_1 \in \mathbb{R}$. We assume that parties have no credible commitment ability, and thus once in power will implement their bliss point.[8]

Payoffs. The parties are purely policy motivated, with quadratic loss utility: $U_i = -(z_1 - z^i)^2$. z^i denotes party i's preferred policy. Assume $z^R > z^L$.

The voter cares only about outcomes. In particular, we interpret outcomes y_t as a measure of the voter's welfare. Thus, her utility is increasing in and, for simplicity but without loss of generality, simply equal to y_t. Given her ideology β_w^z, and the implemented policy z_1 the voter's expected policy utility is then $\beta_w^z z_1$.

In addition, the voter's evaluation of the right-wing party is subject to an idiosyncratic shock $\xi \sim U[-\frac{1}{2\psi}, \frac{1}{2\psi}]$.

To sum up, the game proceeds as follows:

1. The players observe a public record of $T < \infty$ policy-outcome pairs, as well as the (exogenous) policy z_0
2. The two parties propose their models of the world $\beta_R^z, \beta_L^z \in Z \subset \mathbb{R}$
3. The voter privately observes outcome realization y_0
4. The voter adopts the model with best fit to observed history $\beta_w^z \in \{\beta_R^z, \beta_L^z\}$
5. The idiosyncratic preference shock ξ realizes
6. The voter elects party R or L
7. The winner implements policy $z_1 \in \mathbb{R}$

For simplicity, we will assume that the parties know the true value of the coefficient β_0^z, although this is not essential to our story nor our results. Since we do not allow the voter to draw inferences about β_0^z from the parties' proposals, our solution concept is Subgame Perfect Nash Equilibrium.

[8]The assumption of no commitment simplifies the analysis and allows us to focus attention on the ideological stage, but is not crucial. The essential assumption is that the ideological contest and the policy contest are sequential.

Comments on the Model

The main objective of our paper is to analyze the role of political parties as ideological entrepreneurs. Thus, our voter is purely a consumer of ideology. She does not have the ability to produce her own ideology, and must select one from those that are on offer in the political marketplace. In the model, we therefore assume that the voter must adopt one of the ideologies proposed by the two parties.

There are several ways to interpret this assumption. We can imagine a voter who faces a resource constraint: developing a model of the world is costly, whether in terms of time, attention or information collection. Thus, the voter delegates this activity to political parties. Alternatively, we can think about a voter who is wary of cognitive dissonance costs. The concept of cognitive dissonance (Festinger, 1957) refers to individuals' internal need for consistency, in particular consistency between beliefs and actions. Here, regardless of what ideology she adopts, the voter is constrained to cast her ballot for one of the available parties. Being forced to support a party whose model of the world clashes with her own would be dissonant for the voter. In order to avoid such psychological discomfort, the voter restricts her choice set to the ideologies proposed by the parties.

Furthermore, in modeling political parties as ideological entrepreneurs, we do not conceptualize their advantage vis-a-vis the voters as an informational one. Indeed, the assumption that parties know the true value of β_0^z is adopted for simplicity, but is not essential for our results. Thus, parties' persuasive power does not come from an asymmetry of information, but rather from a technological advantage. Parties alone have the resources and motivation to develop and disseminate theories of the world. In this perspective, political parties are not trying to persuade the voters by giving them new information that they do not have access to. Rather, they are offering an interpretation of reality that the voter may not have considered before. Thus, we do not model this phenomenon as a signaling game. In other words, we do not allow the voter to draw inferences about the true model of the world from the party's proposals. However, our voter is not altogether passive. Conditional on these constraints, she uses the information from the past record of outcomes in the most efficient way, by adopting the ideology with the best fit

to observed data. While our voter certainly faces cognitive constraints, she is not entirely naïve, nor is she systematically biased in the way that she processes information.

Analysis

We now describe the formal analysis of the game and the derivation of comparative statics. All proofs are in the Online Appendix.

As usual, we proceed by backwards induction, beginning with the election winner's policy choice. Our parties do not have credible commitment ability. Straightforwardly, this implies that the election winner will simply implement its own preferred policy.

The election

Moving one step backward, consider now the voter's problem at the electoral stage. Having adopted ideology β_w^z, and anticipating the equilibrium policy choice of the two parties, the voter can compute the expected outcome under R and L. Thus, her expected policy utility from electing the right-wing party R is $\beta_w^z z^R$. Similarly, her expected utility from electing L is simply $\beta_w^z z^L$. Further, recall that the voter's evaluation of the right-wing party is subject to the idiosyncratic shock ξ. Thus, in equilibrium the voter chooses to elect the right-wing party if and only if

$$\beta_w^z z^R + \xi > \beta_w^z z^L \tag{2}$$

From the parties' perspective, the election is probabilistic: even after conditioning on the voter's ideology, the election outcome in fact depends on the realization of the shock ξ. Nonetheless, winning the ideological battle gives an electoral advantage.

Proposition 1. *The probability that the right-wing party wins the election is linearly increasing in β_w^z. Specifically, the probability that R wins the election is given by:*

$$\mathbb{P}_R = \frac{1}{2} + \psi \beta_w^z (z^R - z^L) \tag{3}$$

Recall that $z^R > z^L$. A direct implication of Proposition 1 is that the right-wing (left-wing) party's expected utility is linearly increasing (decreasing) in β_w^z. The larger the (absolute) value of β_w^z, the larger the difference in the voter's expected utility from the two parties. This implies that the right-wing party wants the voter to adopt an extreme right-wing ideology, as this insures against unfavorable realizations of the shock ξ. A symmetric reasoning holds for the left-wing party L. Thus, both parties have a preference for inducing extreme (and directionally opposed) ideologies.

The ideological contest

Let us now consider the strategic problem the parties face in the ideological contest. As discussed above, each party wants the voter to adopt an extreme and favorable ideology. However, parties compete with each other over ideological influence, and must consider how proposing an extreme ideology influences their chances of persuading the voter, which depend on the expected fit of the proposed model to the history of outcomes. Crucially, while the parties observe the public record of past policies and outcomes, they face uncertainty about the realization of the outcome of the policy z_0 in place at the beginning of the game. For any pair of two proposed models, the parties are therefore unsure of exactly which model will best explain the outcomes observed by the voter. They must trade off the probability of convincing the voter against the amount of extremism induced in the voter's beliefs.

Proposition 2 is a direct consequence of this uncertainty and the resulting tradeoff.

Proposition 2. *An equilibrium exists for any $T \geq 1$. The (unique) equilibrium can be characterized by:*

1. *The two parties never propose the same model of the world: ideological polarization always emerges in equilibrium. $\beta_R^{z^*} > \beta_L^{z^*}$.*

2. *The parties always win the ideological contest with equal probability, and propose models centered around the expected OLS estimate for $\hat{\beta}^z$:*

$$\frac{\beta_R^{z^*} + \beta_L^{z^*}}{2} = \beta_{OLS}^z = \frac{(\beta_0^z z_0^2 + \sum_{t=-T}^{-1} z_t y_t)}{\sum_{t=-T}^{0} z_t^2} \tag{4}$$

3. Ideological polarization between the parties is given by:

$$\beta_R^{z^*} - \beta_L^{z^*} = \frac{\sigma}{\phi(0)} \frac{|z_0|}{\sum_{t=-T}^{0} z_t^2} \tag{5}$$

We discuss each part of the main characterization proposition in turn.

No convergence

Part 1 of the proposition is reminiscent of the classical divergence theorem from probabilistic voting models (Calvert, 1985). Indeed, the proof follows an identical logic. Conjecture an equilibrium in which the parties present the same model of the world $\widehat{\beta^z}$. Suppose the right-wing party deviates to $\widetilde{\beta^z} > \widehat{\beta^z}$. Suppose that $z_1 > 0$ (an analogous reasoning applies to the symmetric case). If the realization of the outcome y_0 is sufficiently large, R will win the ideological contest with model $\widetilde{\beta^z}$, which she prefers to the conjectured $\widehat{\beta^z}$. In this case, the deviation strictly increases R's expected payoff. If instead the realization of y_0 is low, $\widehat{\beta^z}$ will have a better fit to the history of outcomes observed by the voter, and L will win the ideological contest. In this case, however, the deviation is payoff-irrelevant for the right-wing party, since it does not change the ideology the voter adopts in equilibrium. Given the parties' uncertainty about the realization of y_0, the deviation therefore always strictly increases R's expected payoff, and the conjectured equilibrium does not exist. Thus, ideological polarization always emerges.

Symmetry

Substantively, part 2 of Proposition 2 indicates that a favorable history of outcomes does not generate an advantage in the ideological contest for political parties. Naïve intuition may have suggested that a favorable history (e.g., for party R, a history showing that right-wing policies tend to generate high outcomes) would improve a party's ability to persuade the voter to adopt its proposed model of the world. Instead, the analysis shows that the parties are always equally likely to win the ideological contest in equilibrium. In our world, the two parties face symmetric strategic problems. Given Proposition 1, the right-wing party wants to induce the voter to adopt an extreme right-wing ideology, while the left-wing party would benefit from an extreme

left-wing one. Formally, R maximizes the same function that L minimizes. As a consequence, the equilibrium will take a symmetric form, with the parties ex-ante equally likely to win the ideological contest and persuade the voter to adopt their model of the world. In equilibrium, the ideologies proposed by the two parties are therefore equidistant from the model that maximizes the expected fit to the history of outcomes observed by the voter. This is the expected OLS estimate for the coefficient β^z (which we denote as β^z_{OLS}), where the unobserved outcome realization y_0 is replaced by the expectation $\beta^z_0\ z_0$.

However, the above results also highlight that a favorable history of outcomes does translate into an *electoral* advantage for political parties. As β^z_{OLS} increases, *both* equilibrium ideologies move to the right. In turn, this improves the right-wing party's prospects of winning the election, thereby increasing its expected payoff.

Finally, a corollary of the above is that some ideological manipulation always occurs in equilibrium:

Corollary 1. *For any $T < \infty$, $\beta^{z^*}_R, \beta^{z^*}_L \neq \beta^z_0$ with probability 1.*

While both parties' equilibrium ideologies are a function of β^z_0, because the expression in 4 includes the (random) outcome realizations, the true model of the world emerges only in the limit where $T \to \infty$. For any finite history, therefore, neither party offers the true model with probability 1.

Comparative Statics on Polarization

Part 3 of Proposition 2 describes the determinants of ideological polarization, here defined as the difference between $\beta^{z^*}_R$ and $\beta^{z^*}_L$, in equilibrium. Substantively, this difference captures the extent to which the models of the world proposed by the two parties disagree on the consequences of the policy dimension for the voter's welfare.

Surprisingly, the history of outcomes has no impact on ideological polarization in equilibrium.

Corollary 2. *Equilibrium ideological polarization is not a function of the history of outcomes.*

The intuition derives from the symmetry of the parties' strategic situations. In particular, moving farther away from β^z_{OLS} has an identical (negative) impact on the parties' probability of

winning the ideological contest. As a consequence, parties respond in identical ways to perturbations to the history of outcomes, both moving in the same direction as the estimate β^z_{OLS}. Thus, while each party's equilibrium ideology is a function of observed outcomes, polarization is not.

In contrast, shifts in policies always have an impact on ideological polarization.

Corollary 3. *Equilibrium ideological polarization:*

- *Decreases as past policies become more extreme (i.e., move away from 0);*
- *Increases as today's policy (z_0) becomes more extreme.*

As past policies become more extreme, historical data become more informative for the voter. If policies in the past history are close to 0, proposing a more radical ideology (i.e., a model of the world farther away from the expected best fitting one) has a small impact on the expected model fit, since the predicted outcomes under any model will be close to 0 as well. As past policies become more extreme, instead, moving the proposed model of the world away from β^z_{OLS} reduces the fit to the observed history of outcomes and thus the likelihood of winning the ideological contest. Thus, radical ideologies become more costly as past policies become more extreme, and ideological polarization decreases.

Consider instead the impact of today's policy extremism. Recall that parties face uncertainty over the exact realization of the outcome of today's policy, which is privately observed by the voter. As z_0 becomes more extreme, favorable realizations of the outcome y_0 have a larger (positive) impact on overall model fit. Thus, as today's policy moves away from 0, the parties can gamble on extreme outcomes, and can afford to propose more radical ideologies. Ideological polarization, therefore, increases.

Finally, ideological polarization increases with the uncertainty in the true data-generating process.

Corollary 4. *Equilibrium ideological polarization increases in the variance of the distribution of the error (ε), σ^2.*

This Proposition is again reminiscent of results from probabilistic models of elections (Calvert, 1985; Wittman, 1983). As the variance in the distribution of errors increases, the realization of y_0 becomes more uncertain from the parties' perspective. Thus, the outcome of the ideological

contest becomes more arbitrary, and parties are willing to propose ideologies that are farther away from the expected best fitting model.

The Multidimensional Model

So far, we have assumed that parties are constrained to include only the policy dimension z in their models of the world. We saw that even in this simple setting, a number of interesting comparative statics emerge. Nonetheless, real-world ideologies often weave in non-policy dimensions of the social world:[9] the Thatcher quote in the introduction, for example, categorically denied any effect of "society" on outcomes and asserted instead the primacy of individual and family-level attributes. We aim to capture this richness here, and to understand what happens to equilibrium polarization as models of the world become more complex and expand to encompass more features of the environment.

In this section we allow parties to include in their models, in addition to the policy variable z, an additional vector-valued set of covariates $\mathbf{x}$. The key distinction between policy and covariates is that the policy dimension z is the one over which parties are differentiated and on which they have control. Non-policy covariates may nonetheless be relevant because of the possibility that they affect outcomes. For example, we may think of some components of $\mathbf{x}$ as immutable ethnic, cultural or demographic characteristics of the population. Alternatively, covariates may represent exogenous features of the environment, such as the state of the global economy, or geopolitical factors. These elements cannot be altered by policy means (at least in the short run), but parties may nonetheless find it useful to include them in equilibrium ideologies if they help to explain variation in outcomes.

In this section, we ask if and when parties choose to include non-policy variables in their models, and whether conflict on these dimensions emerges in equilibrium, i.e. whether the proposed models differ on non-policy dimensions. Furthermore, we analyze whether there is an association between model complexity and ideological polarization in equilibrium.

[9] And sometimes also the natural world; e.g. the Lamarckian theory of evolution remained official party orthodoxy in the USSR until 1965.

The generalization from the one-dimensional setting is that the data-generating process is now taken to be $y_t = \beta_0' w_t + \varepsilon_t$, where

$$w_t \equiv \begin{bmatrix} z_t \\ \mathbf{x}_t \end{bmatrix},$$

and $\mathbf{x}_t$ is a k-dimensional vector of attributes. Without loss, we assume that the mean of each dimension of $\mathbf{x}$ is zero. It will be useful in what follows to define the stacked $(T \times (k+1))$ matrix of history W, where each row of W is an observation of w_t for $t \in \{-T, \ldots, -1\}$, and similarly the "extended" history W_+ equal to W with an additional row $w_0 = (z_0, \mathbf{x}_0)$.

Each party i will propose a vector of coefficients β_i of dimension $k+1$. Note that the model need not encompass all available dimensions, as proposing a coefficient $\beta_i^j = 0$ is equivalent to excluding dimension j from the model of the world. As in the baseline setup, the voter will select the ideology that best explains, in the likelihood sense, the history of outcomes she observes. Importantly, we do not assume that the true coefficients $\beta_0^j \neq 0$, i.e. we allow for the possibility that some dimensions of $\mathbf{x}$'s true correlation with outcomes y are 0. Initially we will take the set of covariates included in the common-knowledge history to be exogenous; later we will endogenize the choice of what to make available.

Analysis

Proposition 3 establishes the multidimensional equivalents of each part of Proposition 2. As a result, the main comparative static results from the baseline single-dimensional setup (Corollaries 1-4) will continue to hold in the multidimensional setting.[10]

Proposition 3. *If W has full rank and $w_0 \neq 0$, then a Nash equilibrium exists. In any such equilibrium:*

1. *The two parties' models of the world differ on all dimensions. $\beta_R^{j^*} \neq \beta_L^{j^*}\ \forall\ j$.*

[10]The corollaries in the previous section all hold locally, close to the mean value of the $\mathbf{x}$'s. The implicit expression in (6) reduces to an explicit one in one dimension, such that the corollaries can be stated unconditionally. In the general case, we rely on the implicit function theorem to establish local comparative statics close to the mean of the data.

2. *The parties always win the ideological contest with equal probability, and propose models whose z-component centers around the expected OLS estimate for $\hat{\beta}^z$:*

$$\frac{\beta_R^z + \beta_L^z}{2} = \left((W_+'W_+)^{-1}W_+'y_+\right)^z,$$
$$y_+ \equiv \begin{bmatrix} y \\ w_0'\beta_0 \end{bmatrix}$$

3. *Ideological polarization on all dimensions between the parties is given (implicitly) by the solution to the vector equation:*

$$\beta_R - \beta_L = \frac{\sigma}{\phi(0)} \frac{|w_0'(\beta_R - \beta_L)|}{\beta_R^z - \beta_L^z} \left(W_+'W_+\right)^{-1} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{6}$$

In addition to preserving the results established in one dimension, the multidimensional environment delivers some new comparative static results of interest on which we now expand.

Existence conditions

Proposition 3 requires that the matrix of common-knowledge history W must have full rank. Substantively, this requires that there is sufficiently independent variation in all dimensions j of the history. If there is not, parties can find a direction of change in β_i which holds the likelihood exactly constant, and exploit this to push the coefficient on z they offer to $\pm\infty$ without loss in the chance of winning the ideological contest; the contest thus exhibits infinite ideological polarization and no equilibrium in finite strategies. A necessary condition for full rank to obtain is that $T > k$, and thus the complexity of the environment relative to the length of voters' memory will be a key determinant of equilibrium polarization. We return to this point below, in Corollary 6.

Complete politicization

Part 1 of Proposition 3 shows that equilibrium ideologies are always total: *all* available non-policy dimensions are always included (in the sense of having nonzero coefficient in β_i) by at least one of the parties, and the models offered by the parties differ on all dimensions. Because the size of the history is finite, the within-sample correlation between each x^j and z is always nonzero.[11] Therefore, the (expected) OLS estimate for β^z is a function of whether x^j is included in the model, and because equilibrium proposals move with the expected OLS, at least one party will have an incentive to include it.

The intuition for divergence is that parties can use the $\mathbf{x}$ dimensions to improve the overall fit of their model, and take advantage of this better fit to push for more extremism on the policy dimension. Focus on the incentives facing the right-wing party. The party trades-off the desire to induce the voter to adopt an extreme right-wing ideology on the z dimension, with the need to win the ideological contest. Suppose that x^j and z are positively correlated, with an analogous argument holding for the mirroring case. Then, the cost of ideological extremism on the z dimension can be partially offset by proposing a low coefficient on x^j.[12] Given the positive correlation with z, this tends to move predicted outcomes closer to the observed history, thereby improving the overall fit of the model. Directionally opposite incentives emerge for the left-wing party. As a consequence, the parties' models of the world always differ on the x^j dimension. Specifically, the sign of the element of the covariance matrix corresponding to the pair (z, x^j) determines the role that x^j plays in the parties' narratives:

Corollary 5. *Suppose the element of the within-sample inverse covariance matrix* $(W_+'W_+)^{-1}$ *corresponding to the pair* (z, x^j) *is positive. Then, in an open neighborhood around the line* $w_0 = [z_0, 0, 0, \ldots, 0]$*,* $\beta_R^{j^*} > \beta_L^{j^*}$*. The converse also holds.*

In the simple bivariate setting (with scalar x), the (z, x^j) element of the inverse covariance matrix has the opposite sign as the correlation between the two variables. So, an x that is positively correlated with z will be emphasized more by the left-wing party than the right ($\beta_R^{j^*} < \beta_L^{j^*}$) and vice versa.

[11] Outside of a set of outcomes with probability zero.

[12] "Low" here means low relative to the OLS estimate on this dimension, not low in some absolute sense.

Comparative statics on non-policy dimensions

Following the logic above, parties use additional dimensions to compensate for more extremism on the policy dimension. A consequence is that adding additional dimensions to the history allows the possibility of greater observed polarization on the policy dimension. It turns out that in equilibrium, there is an association between model complexity and ideological polarization, which is conditional on the realization of the last-period policy and covariates:

Corollary 6. *Let* $\tilde{W}_+ \equiv \begin{bmatrix} W_+ & x^{k+1} \end{bmatrix}$*, and let* $w_0 = [z_0, 0, 0, \ldots, x_0^{k+1}]$*. The ratio of equilibrium ideological polarization on the* z *dimension under* $\tilde{W}_+$ *to that under* W_+ *is:*

$$\frac{\tilde{\beta}_R^z - \tilde{\beta}_L^z}{\beta_R^z - \beta_L^z} = \begin{cases} 1, & x_0^{k+1} = 0 \\ r > 1, & \textit{sign}(z_0) = \textit{sign}(x_0^{k+1}(W_+'W_+)_{z,k+1}^{-1}) \\ r > 0, & \textit{sign}(z_0) \neq \textit{sign}(x_0^{k+1}(W_+'W_+)_{z,k+1}^{-1}) \end{cases}$$

Substantively, Corollary 6 says that when today's realization of policy and covariates aligns with the pattern observed in the historical data, observed polarization is higher than it would be without the additional covariate. When there is a "surprise," in the sense that the correlation today is in the opposite direction of the historical pattern, polarization may be lower. If the realization of the covariate today is exactly at its mean (0), there is no effect on polarization of adding the additional covariate. Additional covariates thus increase the variability of observed ideological polarization. If realizations are iid then a "surprise" is less likely than not, and thus on average we expect the addition of more complexity in parties' models to increase observed polarization.

Additionally, although all available dimensions are politicized in equilibrium, some are more politicized than others. Equation 6 shows that the sample covariance matrix $W_+'W_+$ determines the magnitude of the parties' divergence on each dimension. The larger is the correlation between policy and covariate in the history, the larger will be the parties' polarization on that dimension. This statement is Corollary 7.

Corollary 7. *The stronger the (within-sample) correlation between x^j and z, the larger is the magnitude of the difference in equilibrium coefficients on the j dimension, $|\beta_R^{j^*} - \beta_L^{j^*}|$.*

Endogenous Salience

In the previous section, the set of dimensions available for parties to include in their ideologies is exogenous. Furthermore, it is costless for political parties to propose more complex models. As a consequence, even an arbitrarily small correlation with the policy dimension z will allow parties to reap the benefits of including an additional variable in their models of the world, and ideologies are thus always total. All available dimensions become relevant in the ideological competition (though some are more relevant than others).

Here, we will instead endogenize the set of politically salient dimensions. For simplicity of exposition we will consider the choice between the baseline setting where only the policy variable is available in the history, and one with a single additional scalar covariate, x.[13]

We model the endogenous choice of salience by adding an additional stage to the game, occurring prior to the ideological contest. In this stage, either party may pay a cost C to make the dimension x available in the history. If at least one party pays the cost, then the parties play the multidimensional version of the game with $k = 1$; otherwise they play the baseline game. We interpret C as the cost of informing the voter of the past history of x, through a campaign of public education (or propaganda, depending on one's point of view).[14] Absent such costly efforts, the voter is unaware of the history of x, and will not be able to use the x dimension when computing the predicted outcomes under the parties' models of the world. As such, parties cannot gain from including this additional dimension in their ideologies unless and until the voter is made aware of its past history.

[13] This is equivalent to the case with $k = 1$ in Proposition 3.

[14] Qualitatively equivalent results would obtain if each party must decide whether to invest a cost in order to increase the dimensionality of its model, as long as this investment has to be completed before the ideological contest.

Analysis

We note that once the set of salient dimensions is determined, the game proceeds as in the previous section. Thus, the results from Propositions 2 or 3 will apply, depending on which path is chosen in the first stage.

Moving one step backward, we therefore analyze when political parties choose to pay the cost of making x politically salient. Denote β^z_{OLS} the (expected) OLS estimate for β^z if x is included in the model, and $\beta^z_{Omitted}$ the estimate from the single-dimensional model. The difference $\beta^z_{Omitted} - \beta^z_{OLS}$ is the omitted variable bias, which we denote Δ.

Proposition 4. *There exists a unique threshold $|\widehat{\Delta}|$ such that the non-policy dimension x becomes endogenously salient if and only if $|\Delta| \geq |\widehat{\Delta}|$. Furthermore, only one party at a time is willing to pay the cost of making x salient:*

- *R chooses to pay the cost if and only if the omitted variable bias is negative, and sufficiently large;*
- *L chooses to pay the cost if and only if the omitted variable bias is positive, and sufficiently large.*

The threshold $|\widehat{\Delta}|$ is a function of the cost C, the parties' bliss points z^R and z^L and the amount of noise in the electoral process $\frac{1}{\psi}$. The intuition is analogous to that described in 6. Recall that parties always win the ideological contest with equal probability in equilibrium, and that their ideologies always center on the (expected) OLS estimate for β^z on the z dimension. Further, each party's utility is linear in the winning ideology β^z_w (increasing in β^z_w for the right-wing party, and decreasing for the left-wing one). Thus, when choosing whether to pay the cost C, parties only consider how the number of salient dimensions influences the location of this estimate, i.e., the omitted variable bias. Straightforwardly, the right-wing party will be wiling to pay the cost C only if $\beta^z_{Omitted} - \beta^z_{OLS} < 0$. Otherwise, R would want to exploit the omitted variable bias to move both equilibrium ideologies to the right. Analogously, L is willing to pay C only when the bias is positive.

The next Corollary indicates that the extremism in the parties' policy preferences plays a crucial role:

Corollary 8. *The larger the distance in the parties' preferred policy $z^R - z^L$, the higher the likelihood that x will become politically salient (in the sense of set inclusion).*

As the distance in the parties' bliss points increases, the electoral stakes are higher for the parties: losing the election becomes more costly. Thus, parties are willing to pay a higher cost C in order to move the voter's ideology in their preferred direction, thereby improving their electoral chances.

Analogously, as the electoral process becomes more noisy, parties have lower incentives to invest the cost C in order to improve their prospects:

Corollary 9. *The lower the noise in the electoral process (i.e., the lower the variance of the valence shock $\frac{1}{\psi}$), the higher the likelihood that x will become politically salient (in the sense of set inclusion).*

The results of this section highlight that random events may have a crucial importance in the genesis of political ideologies. Recall that parties have no control, or preferences, over the non-policy dimension. Therefore, if a strong enough historical correlation with the policy variable (and thus a large enough omitted variable bias) emerges, it is purely by chance. Political parties, acting as ideological entrepreneurs, will exploit such correlations between variables to their favor, constructing complex models of the world that support their favored policy position. Furthermore, Proposition 4 indicates that the sign of the historical correlation with the policy dimension will determine how the parties position themselves on other issues. Thus, random historical events will determine how social issues, religion, race, and other issues become integrated into political ideologies.

Conclusion

We develop a model of hierarchical policy preference formation, where parties provide causal explanations (ideologies) for patterns observed in voters' past experience. Voters evaluate these ideologies according to their fit to experience, and then use the best-fitting one when deciding which policies to support. Winning the ideological battle thus confers an advantage in subsequent electoral competition. The advantage of political parties over voters in our model is a technological rather than an informational one: their exclusive control over the creation and dissemination

of ideologies grants parties some ability to persuade even in the absence of any private information.

This persuasive ability is constrained by the features of the common-knowledge history of outcomes. A lower degree of extremism in past policy means the data is less informative about the effects of policy, which in turn allows parties to offer more divergent ideologies. The persuasive ability of elites is also determined by the history that is considered politically relevant. We showed how ideology can be shaped by variables that are beyond the control of politicians, and even issues that are not inherently political, as long as they are correlated and, to an uninformed voter's mind, plausibly related.

The model thus helps to understand why the construction of consensus history and its dissemination in schools and universities is so often hotly politicized. Even in a world of perfect information about parties' policy agendas and no commitment to platforms — where what parties will do if given power is perfectly predictable ex ante — what citizens know about the past can change how they respond to party rhetoric, and ultimately which party they are willing to support. Strategic politicians and political movements understand this dependence well, and intervene when they can to shape the "dataset" that future voters draw upon.

An illustrative example comes from the Reconstruction era in the United States. In the years following the US civil war, technological advances and the distribution of vast tracts of land and mineral rights in the nation's interior allowed for the creation of enormous fortunes which quickly overpowered the federal and state governments' regulatory capacity (Skowronek, 1982). An age of spectacular public corruption followed.

At the same time, formerly enslaved Black freedmen gained political rights and wielded substantial political power in the legislatures of Southern states for the first time. White Southern politicians took advantage of the coincidence to argue that Black empowerment had been the cause of the corruption and insider dealing that characterized legislative behavior across the country in this period, and therefore that Black political rights should be curtailed (DuBois, 2007). "The Negro vote and graft were indissolubly linked in the public mind by incessant propaganda" launched by Democratic Redeemers (pp. 511). Despite emerging out of historical coincidence, this association between Black empowerment and political corruption in the Reconstruction pe-

riod, literally written into the history books by Southern historians, provided the intellectual foundation for Jim Crow and one-party rule in the South for nearly a century thereafter.

The Reconstruction era left a profound mark on American, and particularly Southern, politics. Much research has established how it affects politics today (Acharya et al., 2018) and it is natural to think it still runs through political ideologies. Other episodes may be more ephemeral, with parties opportunistically exploiting a variable while it is fresh and dropping it once it fades from memory. Generational turnover too implies a greater importance of the recent past on ideology.

Intriguingly, this suggests that the relationship between voter knowledge and political outcomes is conditional on the type of information held. On one hand, a longer memory of policy history reduces polarization as voters draw upon a larger time series of data to tighten their beliefs about the underlying world. On the other hand, a broader knowledge of other variables, whether policy related or not, may increase polarization as parties are better able to fit the policy variable to their preferred ideology. Proposals to improve policy making by expanding voter knowledge depend critically, therefore, on the direction of knowledge expansion.

To focus on the ideological competition between the parties, we limit attention to a single representative voter. In practice, voters differ in their preferences, their policy knowledge, as well as in their lived experiences. Our model is easily extended to accommodate this heterogeneity. One striking finding of voter behavior is that even those who refuse party labels are rarely indifferent over policy choices (Klar and Krupnikov, 2016). Our model, in distinguishing ideology from policy choice, sheds light on why the middle is empty despite the otherwise smooth heterogeneity across the population. Through this lens it becomes clear why voters with similar objective conditions nevertheless tilt systematically one way over the other. If voters first choose an ideology, then that choice, even when itself a coin flip, guides the later vote choice. It is the underlying, oftentimes obscured ideology, that drives vote choice to systematically favor either Republicans or Democrats.

Exploring this logic more thoroughly seems to us an interesting application of our framework and a fruitful direction for future work. Beyond how the choice of ideology shapes vote choice, we can use our model to understand how different experiences and different information, from education levels to media consumption (Martin and Yurukoglu, 2017), in turn, shape the choice of ideology. The formal separation of ideology from policy preferences that we propose connects

many strands of empirical work in American politics and opens up the tools of formal theory to help organize and deepen our understanding of this fundamental political phenomenon.

References

Acharya, Avi, Matthew Blackwell, and Maya Sen, *Deep Roots: How Slavery Still Shapes Southern Politics*, Princeton University Press, 2018.

Aragones, Enriqueta, Itzhak Gilboa, Andrew Postlewaite, and David Schmeidler, "Fact-Free Learning," *American Economic Review*, December 2005, *95* (5), 1355–1368.

Brady, H.E. and P.M. Sniderman, "Attitude attribution: A group basis for political reasoning," *The American Political Science Review*, 1985, pp. 1061–1078.

Brown, Wendy, *In the ruins of neoliberalism: the rise of antidemocratic politics in the West*, Columbia University Press, 2019.

Bullock, John G., "Politics, ideology, and belief systems," *The American political science review*, 2011, *105* (3), 496–515.

Bénabou, Roland and Jean Tirole, "Belief in a Just World and Redistributive Politics," *The Quarterly Journal of Economics*, 2006, *121* (2), 699–746.

Callander, Steven, "Searching for good policies," *American Political Science Review*, 2011, *105* (04), 643–662.

Calvert, Randall L., "Robustness of the Multidimensional Voting Model: Candidate Motivations, Uncertainty, and Convergence," *American Journal of Political Science*, 1985, *29* (1), 69–95.

Chong, Dennis and James N. Druckman, "Framing Theory," *The Annual Review of Political science*, 2007, *10*, 103–126.

Converse, Philip E., "The Nature of Belief Systems in Mass Publics," in D.E. Apter, ed., *Ideology and Discontent*, Free Press, 1964.

Downs, Anthony, "An economic theory of democracy," 1957.

DuBois, W. E. B., *Black Reconstruction in America.* The Oxford W.E.B. DuBois, Oxford: Oxford University Press, 2007.

Eliaz, Kfir and Ran Spiegler, "A Model of Competing Narratives," *The American Economic Review*, 2020, *110* (12), 3786–3816.

Festinger, Leon, *A theory of cognitive dissonance*, Vol. 2, Stanford university press, 1957.

Goffman, Erving, *Frame Analysis*, Harvard University Press, 1974.

Gramsci, Antonio, *I Quaderni dal Carcere*, Einaudi, 1929.

Hafer, Catherine and Dimitri Landa, "Deliberation as self-discovery and institutions for political speech," *Journal of Theoretical Politics*, 2007, *19* (03), 329–360.

Kinder, Donald R. and Nathan P. Kalmoe, *Neither liberal nor conservative: Ideological innocence in the American public*, University of Chicago Press, 2017.

Klar, Samara and Yanna Krupnikov, *Independent politics*, Cambridge University Press, 2016.

Lemke, Thomas, "'The birth of bio-politics': Michel Foucault's lecture at the Collège de France on neo-liberal governmentality," *Economy and society*, 2001, *30* (2), 190–207.

Lenz, Gabriel S., *Follow the leader?: how voters respond to politicians' policies and performance*, University of Chicago Press, 2013.

Levy, Gilat, Ronny Razin, and Alwyn Young, "Misspecified Politics and the Recurrence of Populism," March 2020. Working paper.

Martin, Gregory J. and Ali Yurukoglu, "Bias in cable news: Persuasion and polarization," *The American Economic Review*, 2017, *107* (9), 2565–2599.

Mason, Lilliana, *Uncivil agreement: How politics became our identity*, University of Chicago Press, 2018.

Mullainathan, Sendhil, Joshua Schwartzstein, and Andrei Shleifer, "Coarse Thinking and Persuasion," *The Quarterly Journal of Economics*, 2008, *123* (2), 577–619.

Ortoleva, Pietro, "Modeling the Change of Paradigm: Non-Bayesian Reactions to Unexpected News," *The American Economic Review*, 2012, *102* (6), 2410–2436.

Petty, Richard E. and John T. Cacioppo, "The Elaboration Likelihood Model of Persuasion," *Advances in Experimental Social Psychology*, 1986, *19*, 123–205.

Sanders, Bernie, *Our revolution: A future to believe in*, Macmillan, 2016.

Sartori, Giovanni, "Politics, ideology, and belief systems," *The American political science review*, 1969, *63* (2), 398–411.

Simon, Herbert A., "A Behavioral Model of Rational Choice," *The Quarterly Journal of Economics*, 1955, *69* (1), 99–118.

Skowronek, Stephen, *Building a new American state: The expansion of national administrative capacities, 1877-1920*, Cambridge University Press, 1982.

Thatcher, Margaret, *The Downing Street Years*, 1st ed. ed., New York, N.Y.: HarperCollins, 1993.

Wittman, Donald, "Candidate motivation: A synthesis of alternative theories," *The American political science review*, 1983, pp. 142–157.

Zaller, John R., *The nature and origins of mass opinion*, Cambridge university press, 1992.

Ideological Competition: Appendix

Contents

A Proof of Proposition 1 **1**

B Proof of Propositions 2 and 3 **2**

B.1 Preliminaries 2

B.2 Existence 4

B.3 Characterization 6

C Proofs of Corollaries 1-7 **8**

D Proofs of Proposition 4 and Corollaries 8-9 **12**

A Proof of Proposition 1

Fixing the winning ideology β_w^z, the voter's expected utility from electing the right-wing party is $\beta_w^z z^R$. Similarly, her expected utility from electing the left-wing party is $\beta_w^z z^L$. Further, recall that her evaluation of the right-wing party is subject to the shock ξ. Thus, the voter will elect R if and only if $\beta_w^z z^R + \xi \geq \beta_w^z z^L$ (the indifference breaking assumption is without loss of generality since ξ is drawn from a continuous distribution). Thus, the probability that R wins the election is:

$$\mathbb{P}(R\ elected) = \mathbb{P}(\xi \geq \beta_w^z(z^L - z^R)).$$

Exploiting the fact that $\xi \sim U[-\frac{1}{2\psi}, \frac{1}{2\psi}]$, the above reduces to:

$$\frac{1}{2} + \psi\beta_w^z(z^R - z^L).$$

In what follows, we will assume that the bounds of the ideological space are such that this probability is always interior.

B Proof of Propositions 2 and 3

We start from the general (multidimensional) version of the model. The result in the one-dimensional case presented in the main text (Proposition 2) follows as an immediate corollary of the general proof (Proposition 3). The existence statement is Property B1, part 1 is Property B2, part 2 is Property B4, and part 3 is Property B3.

B.1 Preliminaries

We begin with the common-knowledge history h consisting of triples $(y_t, \mathbf{x}_t, z_t)$ and the final period covariates and policy $(\mathbf{x}_0, z_0)$. The extended history h^1 observed only by the citizen contains an additional observation of the outcome y_0 corresponding to the final period policy. Each $\mathbf{x_i} \in \mathcal{R}^k$.

The parties' maximization problems are:

$$\max_{\beta_R} \ (\beta_R^z - \beta_L^z)\ \mathbb{P}\left(\log \mathcal{L}(y|h^1, \beta_R) - \log \mathcal{L}(y|h^1, \beta_L) \geq 0\right) + \beta_L^z$$

$$\max_{\beta_L} \ -(\beta_R^z - \beta_L^z)\ \mathbb{P}\left(\log \mathcal{L}(y|h^1, \beta_R) - \log \mathcal{L}(y|h^1, \beta_L) \geq 0\right) - \beta_L^z$$

The probability appearing in the maximand is the probability that the citizen adopts R's model, $\mathbb{P}(R\ win)$. This can be written as:

$$\mathbb{P}(R\ win) \equiv \mathbb{P}(\log \mathcal{L}(y|h^1, \beta_R) - \log \mathcal{L}(y|h^1, \beta_L) \geq 0) =$$
$$\mathbb{P}\left(\sum_{t=-T}^{-1} \log\left(\frac{f(y_t|x_t, z_t, \beta_R)}{f(y_t|x_t, z_t, \beta_L)}\right) + \log\left(\frac{f(y|x_0, z_0, \beta_R)}{f(y|x_0, z_0, \beta_L)}\right) \geq 0\right)$$

Plugging in the distributional assumptions, and letting $w_t \equiv \begin{bmatrix} z_t \\ x_t \end{bmatrix}$, this reduces to:

$$\mathbb{P}\left(w_0'(\beta_R - \beta_L)\left((w_0'\beta_0 + \epsilon) - \frac{w_0'(\beta_R + \beta_L)}{2}\right) \geq -\sum_{i=1}^{n} \sigma^2 \log r_t\right),$$
$$\sigma^2 \log r_t = w_t'(\beta_R - \beta_L)\left(y_t - \frac{w_t'(\beta_R + \beta_L)}{2}\right)$$

The right hand side of the inequality above can be more compactly written in matrix notation as

$$-\frac{1}{\sigma}(\beta_R - \beta_L)'W'(y - W\frac{\beta_R + \beta_L}{2})$$

Where y is the vector of T outcome observations, and W is the $(T \times k)$ matrix of observed data.

Rearranging, we get

$$\mathbb{P}\left(\epsilon/\sigma \geq \frac{1}{\sigma w_0'(\beta_R - \beta_L)}\left(-(\beta_R - \beta_L)'W'(y - W\frac{\beta_R + \beta_L}{2})\right) - w_0'\frac{(\beta_0 - (\beta_R + \beta_L)/2)}{\sigma}\right)$$

for $w_0'(\beta_R - \beta_L)$ positive or the reverse condition if $w_0'(\beta_R - \beta_L)$ is negative. So:

$$\mathbb{P}(R\ wins) = \begin{cases} (1 - \Phi(\xi(\beta_R, \beta_L))), & w_0'(\beta_R - \beta_L) > 0 \\ \Phi(\xi(\beta_R, \beta_L)), & w_0'(\beta_R - \beta_L) < 0 \end{cases} \tag{1}$$

where

$$\xi(\beta_R, \beta_L) = \frac{1}{\sigma w_0'(\beta_R - \beta_L)} \left(-(\beta_R - \beta_L)' W' (y - W \frac{\beta_R + \beta_L}{2}) \right) - w_0' \frac{(\beta_0 - (\beta_R + \beta_L)/2)}{\sigma} \tag{2}$$

Remark 1. *The derivation above extends directly to the case with many citizens with heterogeneous outcome histories, where parties maximize the population-average adopted ideology.*

Suppose that instead of a single citizen, we have a population of citizens of size N. From the above, $\xi(\beta_R, \beta_L)$ is a sufficient statistic for the probability of adoption of party R's ideology. Conditionally on ξ, realizations of the random variable $R\ wins$ are iid. Hence in a large population of citizens with constant ξ (i.e. all observed data is identical and only the last-period outcome realization may differ), the share of the population adopting R's ideology converges to the probability in (1) almost surely. The solution to the single-citizen problem defined above is thus identical to the large-electorate problem with constant ξ where parties solve:

$$\max_{\beta_R} \frac{\beta_R^z - \beta_L^z}{2k} \frac{1}{N} \sum_{i=1}^{N} \mathbf{1} \left(\log \mathcal{L}(y_i | h_i^1, \beta) - \log \mathcal{L}(y_i | h_i^1, \beta_L) \geq 0 \right) + \frac{\beta_L^z}{2k}$$

$$\max_{\beta_L} -\frac{\beta_R^z - \beta_L^z}{2k} \frac{1}{N} \sum_{i=1}^{N} \mathbf{1} \left(\log \mathcal{L}(y_i | h_i^1, \beta) - \log \mathcal{L}(y_i | h_i^1, \beta_L) \geq 0 \right) - \frac{\beta_L^z}{2k}$$

B.2 Existence

We now show that a Nash equilibrium exists in the first (ideological) stage of the contest.

Lemma 1. *The payoff functions are continuous in β_R and β_L.*

Proof. Both parties' utilities conditional on the model adopted by the citizen are linear in the β's (Proposition 1), so the only question concerns the probability $\mathbb{P}(R\ wins)$. $\xi(\beta_R, \beta_L)$ is continuous everywhere except the hyperplane with $w_0'(\beta_R - \beta_L) = 0$. However, the switch from $1 - \Phi(\cdot)$ to $\Phi(\cdot)$ that occurs here counteracts this and preserves the continuity of the payoff function. Taking limits from below and above respectively:

$$\lim_{w_0'(\beta_R-\beta_L)\uparrow 0}\mathbb{P}(win)=\lim_{w_0'(\beta_R-\beta_L)\uparrow 0}\Phi(\xi(\beta_R,\beta_L))=\begin{cases}1, & (\beta_R-\beta_L)'W'(y-W\frac{\beta_R+\beta_L}{2})>0\\ 0, & (\beta_R-\beta_L)'W'(y-W\frac{\beta_R+\beta_L}{2})<0\end{cases}$$

$$\lim_{w_0'(\beta_R-\beta_L)\downarrow 0}\mathbb{P}(win)=\lim_{w_0'(\beta_R-\beta_L)\downarrow 0}1-\Phi(\xi(\beta_R,\beta_L))=\begin{cases}1, & (\beta_R-\beta_L)'W'(y-W\frac{\beta_R+\beta_L}{2})>0\\ 0, & (\beta_R-\beta_L)'W'(y-W\frac{\beta_R+\beta_L}{2})<0\end{cases}$$

□

Property B1. *If W has full rank and $w_0 \neq 0$, then a Nash equilibrium (possibly in mixed strategies) exists.*

Proof. We apply Theorem 1.3 of Fudenberg and Tirole (1991). Lemma 1 established continuity of the payoff functions, the first condition of the theorem. The second condition, compactness of the choice space, reduces to closedness and boundedness in our setting. We show now that the equilibria where strategies are restricted to a closed and bounded subset of $\mathbb{R}^k$ always coincide with the unbounded equilibria for some arbitrary large but finite bound.

We first show that β_R^z times the probability of winning goes to zero as $\beta_R^z \to \pm\infty$ for fixed β_L. An identical argument applies in the opposite case. Hence the best response to any finite strategy is also finite.

Given the structure of payoffs, we need to ensure that $\lim_{\beta_R^z\to\infty}\xi(\beta_R,\cdot)$ is $+\infty$ when $w_0'(\beta_R-\beta_L)>0$ and $-\infty$ when $w_0'(\beta_R-\beta_L)<0$. This is guaranteed if $W'W$ is positive definite, e.g. if W has full rank, as in this case $-(\beta_R-\beta_L)'W'(y-W\frac{\beta_R+\beta_L}{2})$ is strictly convex. $\xi(\beta_R,\cdot)$ is $O(\beta_R^z)$ and hence $\beta_R^z\phi(\xi(\beta_R,\cdot))$ converges to 0 in the limit as required.

To show that there can be no equilibrium where both strategies are infinite, note that any such equilibrium requires that $\xi(\beta_R,\beta_L)$ remains finite in the limit. Otherwise, one player earns expected payoffs of $-\infty$, which cannot be an equilibrium since she can guarantee finite payoffs by proposing any finite β.

If W is full rank, then ξ remains finite in the limit only if $(\beta_R+\beta_L)/2=(W'W)^{-1}W'y$. In this case $\xi=w_0'(W'W)^{-1}W'y$. R's payoff is $\beta_R^z\Phi(\xi)+\beta_L^z(1-\Phi(\xi))$, which can be finite only if $\xi=0$. When W is full rank, $\xi=0$ only if $w_0=0$.

□

B.3 Characterization

The first-order conditions generate a system of equations:

$$\nabla_R \mathbb{P}(R\ wins) = \begin{bmatrix} -\frac{\mathbb{P}(R\ wins)}{\beta_R^z - \beta_L^z} \\ 0 \end{bmatrix} \tag{3}$$

$$\nabla_L \mathbb{P}(R\ wins) = \begin{bmatrix} -\frac{1-\mathbb{P}(R\ wins)}{\beta_R^z - \beta_L^z} \\ 0 \end{bmatrix} \tag{4}$$

$$\tag{5}$$

where ∇_i is the gradient with respect to β_i. By inspection, we have immediately the non-convergence result:

Property B2. *For $T < \infty$, $\beta_R^{z^*} > \beta_L^{z^*}$.*

Proof. From the first row of (3) and (4), $\beta_R^z = \beta_L^z$ implies that the derivative of $\mathbb{P}(R\ wins)$ with respect to β_R^z and β_L^z is infinite. This is possible only if the likelihood surface degenerates to a step function (e.g. in the limit $T \to \infty$). For any finite T the slope is finite and hence $\beta_R^z \neq \beta_L^z$. Given the utility functions it must be that $\beta_R^{z^*} > \beta_L^{z^*}$ in equilibrium. □

We next show that equilibrium proposals are symmetric and center around the expected OLS estimate on the z dimension. It will be useful first to establish the property that divergence appears on all dimensions:

Property B3. *Ideological polarization on all dimensions between the parties is given (implicitly) by the solution to the vector equation:*

$$\beta_R - \beta_L = \frac{\sigma}{\phi(0)} \frac{|w_0'(\beta_R - \beta_L)|}{\beta_R^z - \beta_L^z} \left(W_+' W_+\right)^{-1} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{6}$$

Equilibrium platforms diverge on all dimensions with probability 1.

Proof. Add the two FOC's (3) and (4), plug in the definition in (1) and rearrange to get (6). The ith entry of the vector on the right hand side of (6) can be zero only if $(W_+'W_+^{-1})_{(z,i)}$ is zero. This is a zero probability event if the w's are iid from a distribution with full support on $\mathcal{R}^k$. □

Property B4. *The parties always win the ideological contest with equal probability, and propose models whose z-component centers around the expected OLS estimate for $\hat{\beta}^z$:*

$$\frac{\beta_R^z + \beta_L^z}{2} = \left((W_+'W_+)^{-1}W_+'y_+\right)^z,$$
$$y_+ \equiv \begin{bmatrix} y \\ w_0'\beta_0 \end{bmatrix}$$

Proof. From the first row of the FOCs, it is evident that any symmetric equilibrium (with $\mathbb{P}(R\ wins) = \frac{1}{2}$) requires $\frac{\partial}{\partial \beta_R^z}\mathbb{P}(R\ wins) = \frac{\partial}{\partial \beta_L^z}\mathbb{P}(R\ wins)$. We now show that this is guaranteed to hold in equilibrium.

From the FOC's we have that $\mathbb{P}(A\ wins) = \frac{1}{2}$ iff $\nabla_A\mathbb{P}(R\ wins) = \nabla_B\mathbb{P}(R\ wins)$. From the expression in (1) we see that $\nabla_i\mathbb{P}(R\ wins) = \alpha\nabla_i\xi(\beta_R, \beta_L)$ for some strictly nonzero scalar α, hence checking equality reduces to checking $\nabla_A\xi(\beta_R, \beta_L) = \nabla_B\xi(\beta_R, \beta_L)$. Plugging in the definition from (2) and taking the differences between the gradients with respect to the choice of each party, we get that the condition $\nabla_A\xi(\beta_R, \beta_L) - \nabla_B\xi(\beta_R, \beta_L) = 0$ reduces to:

$$w_0'(\beta_R - \beta_L)'W'(y - W\frac{\beta_R + \beta_L}{2}) - w_0'(\beta_R - \beta_L)W'(y - W\frac{\beta_R + \beta_L}{2}) = 0 \tag{7}$$

Which is true for any β_R, β_L, hence any equilibrium must be symmetric. This result implies that in equilibrium, $\xi(\beta_R, \beta_L) = 0$. Plugging in the expression in (2), and multiplying through by $\sigma w_0'(\beta_R - \beta_L)$ we get:

$$
\begin{aligned}
0 =&((\beta_R - \beta_L)'w_0)'w_0'(\beta_0 - \frac{\beta_R + \beta_L}{2}) + (\beta_R - \beta_L)'W'(y - W\frac{\beta_R + \beta_L}{2}) \\
=&(\beta_R - \beta_L)'w_0w_0'(\beta_0 - \frac{\beta_R + \beta_L}{2}) + (\beta_R - \beta_L)'W'(y - W\frac{\beta_R + \beta_L}{2}) \\
=&(\beta_R - \beta_L)'\left(w_0w_0'(\beta_0 - \frac{\beta_R + \beta_L}{2}) + W'y - W'W(\frac{\beta_R + \beta_L}{2})\right) \\
=&(\beta_R - \beta_L)'\left(W_+'y_+ - W_+'W_+\frac{\beta_R + \beta_L}{2}\right)
\end{aligned}
$$

where the first step uses the fact that $w_0'(\beta_R - \beta_L)$ is a scalar.

We know from Equation (6) that $\beta_R - \beta_L$ is generically nonzero in all components. Plugging in the expression from (6) and canceling the scalar terms gives:

$$
\begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix} (W_+'W_+)^{-1}W_+'y_+ = \frac{\beta_A^z + \beta_B^z}{2} \tag{8}
$$

In other words, the z components of the vectors β_R and β_L must average to the z component of the OLS estimate on the observed data $\hat{\beta}_{OLS} = (W_+'W_+)^{-1}W_+'y_+$. The average on the other dimensions is unconstrained. Hence while (6) pins down the difference between A and B's positions on all dimensions, only the z component of the sum is restricted.

Payoffs are thus unique, but equilibrium strategies are unique only on the z dimension. We note that a natural focal point is the choice of the OLS estimate on all dimensions:

$$
\frac{\beta_R + \beta_L}{2} = (W_+'W_+)^{-1}W_+'y_+ \tag{9}
$$

This choice solves for $\left(W_+'y_+ - W_+'W_+\frac{\beta_R+\beta_L}{2}\right) = 0$ and thus satisfies (7) for any choice of $\beta_R - \beta_L$.

□

C Proofs of Corollaries 1-7

Corollary 1. *For any $T < \infty$, $\beta_R^{z^*}, \beta_L^{z^*} \neq \beta_0^z$ with probability 1.*

Proof. The inequality is immediate from Proposition 3. The proposals center around the expected OLS estimate, which can be written as:

$$\begin{aligned}\hat{\beta}_{OLS} &= (W_+'W_+)^{-1}W_+'y_+ \\ &= (W_+'W_+)^{-1}W_+'(W_+\beta_0 + \varepsilon) \\ &= \beta_0 + (W_+'W_+)^{-1}W_+'\varepsilon\end{aligned}$$

The last term converges in probability to zero but is nonzero in all components with probability 1 for any finite T and full-rank W_+.

□

Corollary 2. *Equilibrium ideological polarization is not a function of the history of outcomes.*

Proof. By inspection of (6), in which y does not appear. □

To get comparative statics on W_+, σ we apply the implicit function theorem around the point where $w_0 = [z_0, 0, 0, \ldots, 0]$, i.e. close to the mean of $\mathbf{x}$. We prove some comparative static statements only in a neighborhood of this line. In one dimension the condition is satisfied trivially and hence the comparative statics hold without qualification.

We define $\Delta \equiv \beta_R - \beta_L$ and using D_Δ to denote the Jacobian matrix of (6) with respect to Δ, we note that

$$D_\Delta = I - \frac{\sigma}{\phi(0)}(W_+'W_+)^{-1}\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix}\left(\frac{w_0'}{\Delta^z} - \frac{|w_0'\Delta|}{(\Delta^z)^2}\right),$$

which when evaluated at $w_0 = [z_0, 0, 0, \ldots, 0]$ is just I. Hence by the implicit function theorem the Jacobian matrix of Δ with respect to the parameters q is just $-D_q$, and it suffices to differentiate (6) with respect to the parameter of interest to get local comparative statics.

Corollary 3. *in an open neighborhood around the line $w_0 = [z_0, 0, 0, \ldots, 0]$, equilibrium ideological polarization:*

- *Decreases as past policies become more extreme (i.e., move away from 0);*
- *Increases as today's policy (z_0) becomes more extreme.*

Proof. At $w_0 = [z_0, 0, 0, \ldots, 0]$, $|w_0'(\beta_R - \beta_L)| = |z_0|(\beta_A^z - \beta_B^z)$.

Hence, (6) reduces at this point to:

$$\beta_R^z - \beta_L^z = \frac{\sigma}{(T+1)\phi(0)}(W_+'W_+)_{zz}^{-1}|z_0| \tag{10}$$

$$\beta_R^j - \beta_L^j = \frac{\sigma}{(T+1)\phi(0)}(W_+'W_+)_{zj}^{-1}|z_0|,\ \forall j \tag{11}$$

The statement follows by inspection of (10).

□

Corollary 4. *In an open neighborhood around the line $w_0 = [z_0, 0, 0, \ldots, 0]$, equilibrium ideological polarization increases in the variance of the distribution of the error (ε), σ^2.*

Proof. By inspection of (10), as $(W_+'W_+)_{zz}^{-1}$, T, $|z_0|$, and the normal density $\phi(0)$ are all positive.

□

Corollary 5. *Suppose the element of the within-sample inverse covariance matrix $(W_+'W_+)^{-1}$ corresponding to the pair (z, x^j) is positive. Then, in an open neighborhood around the line $w_0 = [z_0, 0, 0, \ldots, 0]$, $\beta_R^{j^*} > \beta_L^{j^*}$. The converse also holds.*

Proof. By inspection of (6), as σ, T, $|w_0'(\beta_R - \beta_L)|$, $\beta_R^z - \beta_L^z$, and the normal density $\phi(0)$ are all positive. □

Corollary 6. *Let $\tilde{W}_+ \equiv \begin{bmatrix} W_+ & x^{k+1} \end{bmatrix}$, and let $w_0 = [z_0, 0, 0, \ldots, x_0^{k+1}]$. The ratio of equilibrium ideological polarization on the z dimension under $\tilde{W}_+$ to that under W_+ is:*

$$\frac{\tilde{\beta}_R^z - \tilde{\beta}_L^z}{\beta_R^z - \beta_L^z} = \begin{cases} 1, & x_0^{k+1} = 0 \\ r > 1, & \textit{sign}(z_0) = \textit{sign}(x_0^{k+1}(W_+'W_+)_{z,k+1}^{-1}) \\ r > 0, & \textit{sign}(z_0) \neq \textit{sign}(x_0^{k+1}(W_+'W_+)_{z,k+1}^{-1}) \end{cases}$$

Proof. (10) shows that at $w_0 = [z_0, 0, 0, \ldots, 0]$, $\beta_R^z - \beta_L^z$ is a function only of the variance on the z dimension. We consider adding an additional element with $x_0^{k+1} \neq 0$. The first row of (6) now becomes:

$$\tilde{\beta}_R^z - \tilde{\beta}_L^z = \frac{\sigma}{\phi(0)} \frac{\left| z_0(\tilde{\beta}_R^z - \tilde{\beta}_L^z) + x_0^{k+1}(\tilde{\beta}_R^{k+1} - \tilde{\beta}_L^{k+1}) \right|}{\tilde{\beta}_R^z - \tilde{\beta}_L^z} \left(W_+' W_+\right)_{zz}^{-1}$$

While under the lower dimension the corresponding expression is:

$$\beta_R^z - \beta_L^z = \frac{\sigma}{\phi(0)} |z_0| \left(W_+' W_+\right)_{z,z}^{-1}$$

(Note that while the full covariance matrices are different under the different dimensions, the top left (z, z) element is the same in both). Dividing the two, we get that the ratio $r = \frac{\tilde{\beta}_R^z - \tilde{\beta}_L^z}{\beta_R^z - \beta_L^z}$ satisfies:

$$r = \frac{\left| z_0(\tilde{\beta}_R^z - \tilde{\beta}_L^z) + x_0^{k+1}(\tilde{\beta}_R^{k+1} - \tilde{\beta}_L^{k+1}) \right|}{|z_0| \left(\tilde{\beta}_R^z - \tilde{\beta}_L^z\right)}$$

Obviously, r must be positive, and if $x_0^{k+1} = 0$ then $r = 1$. By Corollary 5, the sign of $\beta_R^{k+1} - \tilde{\beta}_L^{k+1}$ is the same as the sign of $\left(W_+' W_+\right)_{z,j}^{-1}$. If $x_0^j \left(W_+' W_+\right)_{z,j}^{-1}$ has the same sign as z_0 then the absolute value in the numerator can be separated and the ratio becomes $r = 1 + R$ where R is strictly positive. Otherwise, it cannot, and we can only show $r \geq 0$. □

Corollary 7. *The stronger the (within-sample) correlation between x^j and z, the larger is the magnitude of the difference in equilibrium coefficients on the j dimension, $|\beta_R^{j^*} - \beta_L^{j^*}|$.*

Proof. By inspection of (11), as σ, T, $|z_0|$, and the normal density $\phi(0)$ are all positive. □

D Proofs of Proposition 4 and Corollaries 8-9

Proceeding by backwards induction, we can compute the parties' expected utility as a function of the set of salient dimensions. Given property B4, we know that parties always win the ideological contest with equal probability. Given properties B3 and B4, the z-component of the equilibrium ideologies is a function of the set of covariates. Furthermore, we know from Proposition 1 that each party's probability of winning the election is a function of β_w^z. Thus, the set of salient dimension influences the parties' expected utility only via the z-component of their equilibrium models.

In particular, let $\mathbb{I}_x$ a binary indicator taking value 1 if x is salient, and 0 otherwise. Then, we can denote $\beta_R^{z^*}(\mathbb{I}_x)$ and $\beta_L^{z^*}(\mathbb{I}_x)$ the parties' equilibrium proposed coefficients on the z dimension as a function of the set of salient dimensions.

With this, we can write the parties' expected utility as:

$$E[U_R] = -\frac{1}{2}(z^R - z^L)^2(\frac{1}{2} - \beta_R^{z^*}(\mathbb{I}_x)\psi(z^R - z^L)) - \frac{1}{2}(z^R - z^L)^2(\frac{1}{2} - \beta_L^{z^*}(\mathbb{I}_x)\psi(z^R - z^L)) \tag{12}$$

and

$$E[U_L] = -\frac{1}{2}(z^R - z^L)^2(\frac{1}{2} + \beta_R^{z^*}(\mathbb{I}_x)\psi(z^R - z^L)) - \frac{1}{2}(z^R - z^L)^2(\frac{1}{2} + \beta_L^{z^*}(\mathbb{I}_x)\psi(z^R - z^L)) \tag{13}$$

Thus, the right-wing party's expected utility is higher when x is salient if and only if

$$\begin{aligned} -\frac{1}{2}(z^R - z^L)^2(\frac{1}{2} - \beta_R^{z^*}(1)\psi(z^R - z^L)) - \frac{1}{2}(z^R - z^L)^2(\frac{1}{2} - \beta_L^{z^*}(1)\psi(z^R - z^L)) > \\ -\frac{1}{2}(z^R - z^L)^2(\frac{1}{2} - \beta_R^{z^*}(0)\psi(z^R - z^L)) - \frac{1}{2}(z^R - z^L)^2(\frac{1}{2} - \beta_L^{z^*}(0)\psi(z^R - z^L)) \end{aligned} \tag{14}$$

Which reduces to

$$\beta_R^{z^*}(1) + \beta_L^{z^*}(1) > \beta_R^{z^*}(0) + \beta_L^{z^*}(0) \tag{15}$$

From Property B4, we know that in equilibrium the parties always propose models whose z-component is centered around the expected OLS estimate for β^z. Denote β^Z_{OLS} the expected OLS estimate for β^z if x is salient, and $\beta^Z_{Omitted}$ the estimate if x is excluded by both models.

The above can be rewritten as

$$\beta^z_{OLS} > \beta^z_{Omitted} \tag{16}$$

Thus, the right-wing party's expected equilibrium utility increases when x is salient if and only if the omitted variable bias is negative. Similarly, we can verify that the left-wing party's expected equilibrium utility increases when x is salient if and only if the omitted variable bias is positive. Therefore, only one party at a time may be willing to pay the cost of making x salient.

Let us focus on the right-wing party, with symmetric results holding for the left-wing one. In equilibrium, R chooses to pay the cost C if and only if

$$\begin{aligned} -\frac{1}{2}(z^R - z^L)^2(\frac{1}{2} - \beta^{z^*}_R(1)\psi(z^R - z^L)) - \frac{1}{2}(z^R - z^L)^2(\frac{1}{2} - \beta^{z^*}_L(1)\psi(z^R - z^L)) - C > \\ -\frac{1}{2}(z^R - z^L)^2(\frac{1}{2} - \beta^{z^*}_R(0)\psi(z^R - z^L)) - \frac{1}{2}(z^R - z^L)^2(\frac{1}{2} - \beta^{z^*}_L(0)\psi(z^R - z^L)) \end{aligned} \tag{17}$$

Which reduces to

$$C < (z^R - z^L)^3\psi\frac{\beta^{z^*}_R(1) + \beta^{z^*}_L(1) - \beta^{z^*}_R(0) - \beta^{z^*}_L(0)}{2} \tag{18}$$

Which, noticing that the z-components of the equilibrium models are always centered around the expected OLS estimate for β^z can be rewritten as

$$C \quad < -\psi(z^R - z^L)^3\Delta, \tag{19}$$

$$\Delta \quad \equiv \beta^z_{Omitted} - \beta^z_{OLS} \tag{20}$$

The proof of Corollaries 8 and 9 follows straightforwardly from inspection of (19).