Motivated Reasoning and Democratic Accountability

ANDREW T. LITTLE   *University of California, Berkeley, United States*
KEITH E. SCHNAKENBERG   *Washington University in St. Louis, United States*
IAN R. TURNER   *Yale University, United States*

*D*oes motivated reasoning harm democratic accountability? Substantial evidence from political behavior research indicates that voters have "directional motives" beyond accuracy, which is often taken as evidence that they are ill equipped to hold politicians accountable. We develop a model of electoral accountability with voters as motivated reasoners. Directional motives have two effects: (1) divergence—voters with different preferences hold different beliefs, and (2) desensitization—the relationship between incumbent performance and voter beliefs is weakened. While motivated reasoning does harm accountability, this is generally driven by desensitized voters rather than polarized partisans with politically motivated divergent beliefs. We also analyze the relationship between government performance and vote shares, showing that while motivated reasoning always weakens this relationship, we cannot infer that accountability is also harmed. Finally, we show that our model can be mapped to standard models in which voters are fully Bayesian but have different preferences or information.

A rich and accurate model of voter behavior needs to account for sensible voter responsiveness but also these [psychological] biases, identifying the conditions under which retrospective voting achieves effective democratic accountability and when it fails to do so.

Healy and Malhotra (2013)

Voters are often misinformed about political facts. Furthermore, biases in voter beliefs are systematic in a way that suggests they may interpret information using directionally motivated reasoning that reinforces preferences for candidates or policies they already favored (Redlawsk 2002; Taber and Lodge 2006). These findings lead some to worry that standard models of electoral accountability do not describe real electorates and that voters are not competent enough to induce desirable behavior by politicians (e.g., Achen and Bartels 2017). However, the implications of voters' motivated reasoning for democratic performance are not well understood because improving voters' information about politicians' performance need not always improve democratic performance (Ashworth and Bueno de Mesquita 2014). The voter-centric orientation of the political psychology literature and the focus of political economy models on rational voters leaves us without answers to some basic questions. Is directionally motivated reasoning a concern for democracy? Which motives and patterns of beliefs are of greatest concern? What is the connection between the influence of motivated reasoning on electoral accountability and electoral outcomes?

To answer these questions, we analyze a variant of a standard political economy model of elections. An incumbent politician decides how hard to work on behalf of voters. The incumbent's performance in office is increasing in this effort as well as her competence and a random shock. Voters observe the performance, form a conclusion about the incumbent's competence, and prefer to reelect more competent representatives.[1] As long as these conclusions are increasing in incumbent performance, she will have an incentive to exert effort to increase her chance of reelection. The accountability mechanism works better when these incentives for effort are strong.

In the standard model, voters form their beliefs about the incumbent's performance according to Bayes' rule. Or, in the terminology from the behavioral literature we draw on, voters only have *accuracy motives*. Our main innovation is to assume that in addition to accuracy motives, voters also have *directional motives*: conclusions about the incumbent politician that they like more than others, independent of incumbent performance. Our model highlights two potential behavioral effects of this kind of motivated reasoning and spells out the implications for both politicians' performance in office and aggregate electoral outcomes.[2]

Andrew T. Little [iD], Associate Professor, Department of Political Science, University of California, Berkeley, United States, andrew.little@berkeley.edu.
Keith E. Schnakenberg [iD], Associate Professor, Department of Political Science, Washington University in St. Louis, United States, keschnak@wustl.edu.
Ian R. Turner [iD], Assistant Professor, Department of Political Science, Yale University, United States, ian.turner@yale.edu.

---

[1] We use the term *conclusion* to refer to an estimate of incumbent competence. We say conclusion rather than belief to distinguish this estimate from a full probability distribution over types.

[2] We also analyze how motivated reasoning affects whether less competent politicians are chosen when voters have stronger directional motives in the "Electoral Selection and Voter Welfare" section. The analysis of selection is more complicated but broadly similar.

DIVERGENCE AND DESENSITIZATION

The first effect of motivated reasoning is *divergence* of voters' conclusions about the incumbent. Voters with different preexisting affinity toward or against the incumbent form different conclusions about her competence, and increasing divergence refers to the phenomenon where these differences increase in voters' directional motives. Divergence is a theoretical analogue to many empirical results that have long been seen as a serious risk to democratic accountability. For instance, members of different parties have different perceptions of economic indicators (Bartels 2002), of whether weapons of mass destruction were found in Iraq (Jacobson 2010), or of what proportion of the federal budget is allocated to welfare programs (Kuklinski et al. 2000).

Our model also formalizes an important caveat to these findings raised by some scholars of partisanship: divergence does not imply that voters are unresponsive to information, and in fact they may all respond to changes in incumbent performance "in parallel" (Gerber and Green 1999; Green, Palmquist, and Schickler 2004). In our conception of motivated reasoning, no matter how much voter conclusions diverge for a fixed performance level, all individual conclusions respond to performance in the "correct" direction—better performance improves (or at least does not harm) assessments of incumbent competence. Divergence is not driven by a "backlash" effect where, say, those with an affinity toward the incumbent respond to negative information about her performance by doubling down on their support.

However, this does not imply that motivated reasoning has no influence on how voters respond to information. The second effect of motivated reasoning we identify, *desensitization*, is a weakening of this response. Independent of divergence effects, stronger directional motives can weaken the connection between actual performance and voters' conclusions about incumbent quality. Desensitization captures, for example, what Hill (2017) calls "cautious Bayesians"—voters whose beliefs move toward the correct beliefs but not far enough (see Benjamin 2019 for a more general discussion of the evidence that experimental subjects underreact to new information in many settings). This effect is consistent with empirical patterns that suggest information about the economy influencing vote choice has weakened in recent years (e.g., Donovan et al. 2020; Freeder 2019). Our first set of results shows how several natural formulations of voters' accuracy and directional motives lead to different combinations of belief divergence and belief desensitization.

MOTIVATED REASONING AND ACCOUNTABILITY

Our core theoretical contention is that desensitization has more uniform consequences for democratic performance than divergence. The reasoning for this is twofold. First, the mechanism through which voters' beliefs affect politicians' incentives for effort is by changing the relationship between performance in office and the likelihood of reelection. Desensitization leads voters to have a weaker response to changes in performance, which causes lower effort by politicians. Divergence alone need not have this effect: if some voters evaluate the incumbent more favorably than others but each respond to changes in performance in the same way, politicians' behavior may be approximately the same as with fully Bayesian voters. Second, electoral incentives depend mostly on the behavior of voters in the middle. If centrist voters are not inclined toward one side or the other, divergence will not have a substantial effect on their conclusions and, therefore, will not affect politicians' behavior. In contrast, desensitization affects all voters, including centrists.

This observation leads to the following contention: although motivated reasoning may undermine politicians' incentives to work on behalf of voters, the reason that this is true is not what one might expect. It is common to lament the consequences of deeply committed partisans holding systematically divergent beliefs for democratic accountability. However, our results suggest that the more pernicious threat to democratic accountability from motivated reasoning is that even moderate voters become less responsive to new information.

MOTIVATED REASONING AND ELECTORAL OUTCOMES

We also tie our results to the larger empirical literature that studies the relationship between performance indicators (economic growth, crime, educational outcomes) and incumbent vote shares. This literature often interprets the strength of this relationship as a proxy for the strength of politician incentives to provide good outcomes. Further, some recent research argues that a decline in this relationship in the United States is driven by increased partisan-motivated reasoning (e.g., Freeder 2019). Our model highlights that the degree to which this kind of inference is warranted depends on the kind of motivated reasoning in question. We show that *both* divergence and desensitization will weaken the correlation between performance and vote shares, as this relationship depends on more than just the median voter. However, only desensitization unambiguously leads to worse incumbent performance. Thus, if the primary effect of motivated reasoning is that it drives belief divergence, then it can weaken the relationship between performance and vote shares with minimal effect on electoral accountability.

CONTRIBUTIONS

Our paper makes three main contributions. First, we make a methodological contribution by integrating a game-theoretic model with a model of motivated reasoning. Hill (2017) argues that "formal models of

accountability, which usually assume that citizens update as perfect Bayesians, may benefit by considering the implications of … biased Bayesian processing of signals about incumbent performance" (1404). We provide exactly such a model. The model of belief formation is based on Little (2019), which conceptualizes motivated reasoning as an optimization problem, balancing accuracy and directional motivations (Kunda 1990).[3] Ours is the first paper to incorporate this model of motivated beliefs into a game-theoretic or decision-theoretic model, which we see as a promising avenue for synthesizing work in political behavior and political economy.

A concrete aspect of this contribution is to decompose the effects of motivated reasoning into belief divergence and belief desensitization, which could apply well beyond the particular accountability model we study. This decomposition proves useful theoretically, and also provides clarity to the empirical literature on motivated reasoning that largely focuses on the experience of individual voters. In experimental terms, one can conceptualize belief divergence as an across-subject phenomenon and belief desensitization as a within-subject phenomenon. A treatment that increases the strength of directional motives among subjects may increase divergence across subjects and desensitization within subjects. As we show below, some commonly discussed behavioral biases exhibit only one of divergence or desensitization and others may exhibit both, which may provide practical leverage to isolate what kinds of directional motives are common.

Second, within the context of democratic accountability, our model helps to clarify normative implications of contemporary work in political psychology and political behavior. Our model not only explains when voter biases might matter for electoral accountability but also helps explain *which* biases affect accountability and suggests a way to categorize different findings in the literature to help think through their potential electoral consequences.

Moreover, this decomposition of effects also illustrates how motivated reasoning can affect aggregate electoral outcomes without negatively affecting the incentives provided by elections. While studies that focus on identifying and exploring the flaws of individual voters are crucial for understanding how voters experience the political world, on their own they cannot provide normative conclusions about democratic accountability. In order to understand how and when directionally motivated reasoning threatens the quality of democracy, the interdependence between voter and politician behavior needs to be explicitly considered (Ashworth and Bueno de Mesquita 2014).

Third, we show that the effects of motivated reasoning on voters and politicians parallel effects that are studied in "standard" models. While motivated reasoning matters, these effects are not qualitatively different than the effects of more standard parameters like heterogeneity of voter preferences, the quality of the information environment, and how much voters weigh performance versus other non-performance-related factors. The risks to effective democratic governance driven by motivated reasoning are important to consider, but they are not uniquely troubling in terms of voters' ability to hold their representatives accountable.

RELATED LITERATURE

To develop a model of democratic accountability with motivated reasoning, we build on a voluminous theoretical literature in political economy studying electoral accountability (e.g., Fearon 1999).[4] This literature has provided insight into how and when elections can produce high-quality governance. The ability of elections to produce good government follows from citizens' ability to select quality politicians and hold incumbents accountable through the threat of removing them from office through voting. Typically these two effects depend on voters updating their beliefs about incumbent quality based on performance information. This improves selection by increasing the likelihood that good politicians are reelected and improves performance incentives because incumbents want to produce signals that they are high quality.

Empirical research on electoral accountability is often more skeptical about the ability of citizens to effectively hold their representatives to account (e.g., Lupia and McCubbins 1998). Much of this work has focused on whether voters are sufficiently informed (Delli Carpini and Keeter 1996) or sufficiently rational (Achen and Bartels 2017; Healy, Malhotra, and Mo 2010; Woon 2012) to hold politicians accountable in ways predicted by the models noted above.[5] Additionally, evidence on whether providing voters with more information about political behavior affects behavior is mixed (e.g., Ashworth 2012; Dunning et al. 2019).

Most closely related to our purpose are studies arguing that directionally motivated reasoning leads voters to form biased beliefs about politics (e.g., Lodge and Taber 2013),[6] often caused by partisan attachments coloring voter perceptions (Bolsen, Druckman, and Cook 2014).[7] These deviations may be particularly

---

[3] This is also conceptually similar to the dual motivations argument in Groenendyk (2013). Mayraz (2019) presents a related model of "wishful thinking" where agents update beliefs using Bayes' rule but treat their directional motives as an extra signal, which Thaler (2020) adapts for an experimental test of motivated reasoning.

[4] See Ashworth (2012) and Duggan and Martinelli (2017) for reviews of the literature.

[5] See Healy and Malhotra (2013) for a review.

[6] Among many others, see also Bisgaard (2019), Druckman and McGrath (2019), Groenendyk (2013), Groenendyk and Krupnikov (2020), Lau and Redlawsk (2006), Nyhan and Reifler (2010), and Prior, Sood, and Khanna (2015).

[7] Though see Bullock and Lenz (2019), which suggests survey answers may sometimes reflect "cheerleading" rather than sincere and strongly held beliefs. See also Bullock et al. (2015) and Roush and Sood (2020) for work on partisan knowledge gaps. De Vries,

harmful to accountability if voters with different partisanship or ethnicity only react to information they are already predisposed to believe (Adida et al. 2017). Redlawsk (2002) characterizes motivated reasoning specifically as "a direct challenge to the notion of candidate evaluation as a Bayesian updating process in which voters readily modify their prior expectations based on the value of new information" (1041).

Yet, an established result in political science is that incumbent politicians tend to fare better electorally when government or the economy is generally performing well (e.g., de Benedictis-Kessner and Warshaw 2020; Erikson 1989; Ferraz and Finan 2008; Payson 2017). This seems to be consistent with voters' ability to hold their representatives accountable, largely in line with the "standard" political economy model predictions.

A broad reading of these literatures suggests that voters are individually flawed but aggregate outcomes reward good performance in predictable ways. Ashworth and Fowler (2019) provide one way to reconcile these findings, showing that the presence of flawed individual voters does not preclude electorates from enforcing democratic accountability. Even when the typical voter is biased, the electorate is nonetheless able to provide positive effort incentives for incumbents. Our results about the aggregate consequences for accountability and vote shares given less than fully rational individual voters are complementary. However, by explicitly incorporating motivated-reasoning processes into the model we also provide more detailed results on the consequences of different voter biases.

Methodologically, we contribute to a burgeoning literature in political economy that incorporates psychological concepts into formal models. For example, Ashworth and Bueno de Mesquita (2014) show how voters who fail to form correct conjectures about politician effort affect the equilibrium behavior and welfare properties of a similar model. Diermeier and Li (2017) study an accountability model where voters reward politicians for good performance beyond what they would in a standard prospective voting model. Previous work has also considered voters who use adaptive rules (Bendor et al. 2011), experience cognitive dissonance (Acharya, Blackwell, and Sen 2018), focus more or less on costs/benefits (Nunnari and Zápal 2017), are affected by partisan affect (Diermeier and Li 2019), selectively perceive new information (Gerber and Green 1999), or make random mistakes in their belief formation (Ogden 2016).[8]

A MODEL OF ELECTIONS WITH MOTIVATED REASONERS

We study a model of elections with voters who are directionally motivated reasoners. The players are an incumbent politician (I, she/her), a nonstrategic electoral challenger (C), and a finite set of citizens $N=\{1,\ldots,n\}$ with n odd. We refer to a generic voter with j and male pronouns.[9]

The incumbent politician has competence θ_I. The value of θ_I is unknown to all players, including the incumbent. The common prior belief is that θ_I is drawn from a normal distribution, with mean normalized to zero and variance σ_θ^2. The challenger has analogous competence θ_C, drawn from a common prior distribution that is normal with mean μ_C and variance σ_θ^2.[10] Because the mean of the incumbent competence is normalized to 0, $\mu_C < 0$ will capture cases when the incumbent is ex ante viewed as more competent than the challenger and $\mu_C > 0$ cases where the challenger is ex ante more competent.

The incumbent exerts effort $e \geq 0$, which is valued by the public but is not observed directly. This effort term captures a variety of things that an incumbent politician can choose to do to improve government performance indicators. Accountability is "working well" when the incumbent exerts high levels of effort. The cost of effort is $c(e)$, which we assume is increasing and convex.

Though voters do not observe the incumbent's competence or effort directly, they observe a public signal correlated with both,

$$s = \theta_I + e + \varepsilon,$$

where ε is normally distributed with mean 0 and variance σ_ε^2. A natural way to interpret s is as a performance indicator like GDP growth. The ε term could correspond to factors outside of the incumbent's control that nonetheless also affect the outcome. Alternatively, ε could represent noise in citizens' perception of this signal, in which case $\theta_I + e$ is the "real" outcome.[11]

After the public signal is realized, there are also common shocks to voters' utilities for reelecting the incumbent (η_I) or instead electing the challenger (η_C). We assume $\eta_C - \eta_I$ follows a normal distribution with mean μ_η and variance σ_η^2. These shocks capture swings in candidate- or party-specific preferences that are unrelated to performance. They can be thought of as partisan "national tides" that change aggregate

---

Hobolt, and Tilley (2018) also show that presentation of unambiguous real-world conditions increases accurate answers across partisans.

[8] See also Levy and Razin (2015) for a model of information aggregation where voters don't account for the correlation between their sources of information, Glaeser and Ponzetto (2017) for models in which voters suffer from "fundamental attribution errors," and, more generally, Minozzi (2013) for a model in which beliefs are chosen at the beginning of the game to maximize the total payoff from the game.

[9] Our use of gendered pronouns is merely to ease exposition, and is not intended to imply different beliefs or behavior based on gender identity.

[10] The variance of the belief about the challenger does not play a role in our analysis, so setting it equal to the variance of the incumbent belief is innocuous.

[11] The informational environment in our model, i.e., normally distributed types with additive and normally distributed random noise, resembles the global games literature (Morris and Shin 2003), with the exception that signal realizations are common to all players rather than private. We thank an anonymous reviewer for pointing out this connection.

attitudes toward the parties. Alternatively, as the difference in shocks is what matters for behavior, these shocks could capture the idea that incumbents do not yet know who the challenger will be and may face a more- or less-popular electoral opponent. They can also represent any unpredictable shocks to the incumbent's popularity that may arise through media coverage, spillover effects from copartisan behavior, or down-ballot effects of other elections.

Each voter j also has an *affinity* for the incumbent, denoted by $a_j \in \mathbb{R}$, which directly affects his utility. Voter affinities capture voters' general taste $(a_j > 0)$ or distaste $(a_j < 0)$ for the incumbent that is independent of actual performance. Prominent examples include partisanship and party polarization (Druckman, Peterson, and Slothuus 2013) and party identification as social identity (Huddy, Mason, and Aarøe 2015).[12] The main innovation of our model is to tie this preference parameter to belief formation, which we formalize below.

Finally, after observing s, η_I, and η_C, each citizen $j \in N$ decides whether to vote for the incumbent $(r_j = 1)$ or the challenger $(r_j = 0)$. The candidate receiving the majority of votes wins. We denote whether the incumbent is reelected by $R \in \{0, 1\}$, where $R = 1$ means the incumbent is reelected and $R = 0$ means the challenger is elected. Each voter's second-term payoff sums the winning politician's competence, the voter's affinity toward the winner, and a random shock.[13]

Payoffs. The incumbent seeks reelection but also pays for any effort. Accordingly, her utility is given by[14]

$$u_I(e; R) = R - c(e).$$

In exerting higher effort, the incumbent trades off improving the performance signal, and thus the voter's assessments of her quality, with the cost of that effort.

The utility of each voter $j \in N$ is given by

$$u_j(R) = \underbrace{s + a_j}_{\text{Period 1}} + \underbrace{R(\theta_I + a_j + \eta_I)}_{\text{Period 2 with } I} + \underbrace{(1-R)(\theta_C + \eta_C)}_{\text{Period 2 with } C}.$$

The $s + a_j$ terms capture citizen payoffs from the first period, which are equal to the outcome s plus his affinity for the incumbent. The remaining terms capture the second-period payoff based on incumbent and challenger behavior described above.

Incorporating Motivated Reasoning

To capture citizens as motivated reasoners, we assume that rather than forming beliefs about incumbent competence using only Bayes' rule, as in standard accountability models, the voter forms a *conclusion* about incumbent competence. We refer to a generic conclusion as $\tilde{\theta}_I$ and assume voters reach an *optimal conclusion* $\tilde{\theta}_I^*$:

$$\tilde{\theta}_I^*(s; a_j) \in \arg\max_{\tilde{\theta}_I} \quad \log f_{\theta_I|s}(\tilde{\theta}_I|s) + \delta v(a_j, \tilde{\theta}_I), \quad (1)$$

where the first term (the log-likelihood of the Bayesian posterior belief) captures the accuracy motive and the second term captures the directional motive.

The accuracy motive captures the fact that the voters prefer conclusions that are objectively more likely.[15] By standard calculations, when the voter expects effort level $\hat{e}$, the density in the first term is normally distributed with a mean that is a weighted average of the signal minus the expected effort and the prior mean (which is normalized to 0),

$$\bar{\mu}(s) = \frac{\sigma_\varepsilon^2(s - \hat{e})}{\sigma_\varepsilon^2 + \sigma_\theta^2},$$

and variance,

$$\bar{\sigma}_\theta^2 = \frac{\sigma_\varepsilon^2 \sigma_\theta^2}{\sigma_\varepsilon^2 + \sigma_\theta^2}.$$

That is, $\bar{\mu}(s)$ and $\bar{\sigma}_\theta^2$ represent the mean and variance of voters' Bayesian posteriors given performance signal s.

The second term, $\delta v(a_j, \tilde{\theta}_I)$, measures each voter's directional motive: $\delta \geq 0$ is a scalar that measures the general strength of directionally motivated reasoning, and the $v(\cdot)$ function dictates the relationship between voter affinity and preferred conclusions. We assume that this function has the following properties.

Assumption 1. The directional motive function $v(a_j, \tilde{\theta}_I)$ is (i) continuous and (weakly) concave in $\tilde{\theta}_I$ and (ii) $\frac{\partial^2 v(a_j, \tilde{\theta}_I)}{\partial a_j \partial \tilde{\theta}_I} \geq 0$.

The continuity assumption in part (i) ensures that small changes in voter affinity and conclusion have small effects on the desirability of the conclusion. The concavity assumption loosely means that there are "diminishing returns" to forming a conclusion closer to one's ideal. Assumption 1(i) only requires weak concavity, which is satisfied for instance by a linear directional motive, which is one of the examples we use below. Part (ii) states that as a voter's affinity for the incumbent increases, he has an intrinsic reason, aside

---

[12] Affinities can also capture emotional states that may affect decision making (Albertson and Gadarian 2015).

[13] These payoffs capture what the average performance in period 2 would be with no incumbent effort, which is what would happen in equilibrium since the game ends after second period effort choice.

[14] For simplicity, we abstract away from other motivations that might influence effort. For example, some politicians may care strongly about their own legacy, which may lead to positive incentives for effort. The qualitative nature of all of our results would continue to hold because for a fixed level of "legacy motivations," reducing effort incentives based on reelection concerns will still weaken effort incentives overall.

[15] See Little (2019) for an explanation of why the log likelihood is a natural way to capture this and for a more general discussion of this formulation.

from objective performance, to conclude that the incumbent is more competent. This ensures that directional motivations are increasing in voter affinity for the incumbent.

In effect, we make two departures from a standard model of belief formation, though only one of them is consequential in our setting. First, rather than computing an expected payoff given a belief about θ_I, the voter picks a single conclusion about the incumbent's competence. That is, even if $\delta = 0$, the voters here form a conclusion at the mode of their belief rather than considering the full distribution. However, because θ_I enters voter utility in a linear fashion and the mean and mode of a normal distribution are equal, a voter with $\delta = 0$ behaves in an identical manner to a voter using standard Bayesian beliefs and expected utility maximization. The second, consequential difference is the addition of the directional motive, which we illustrate through a series of substantive examples in the next section.

Solution Concept

Other than the model of belief formation, our solution concept is standard, requiring that voters behave optimally given incumbent effort choice and their beliefs and the incumbent makes an effort choice that maximizes her utility given voter behavior. Formally, a *sincere motivated reasoning equilibrium* is a pure strategy profile consisting of an effort choice for the incumbent and a voting strategy for each voter $(e^*, \{r_j^*(s, \eta_I, \eta_C)\}_{j \in N})$, combined with a profile of optimal conclusions for each voter, $\{\{\tilde{\theta}_I^*(s, a_j)\}_{j \in N}\}$ such that each conclusion $\tilde{\theta}_I^*(s, a_j)$ is a solution to Equation 1, each r_j assigns j's vote choice to j's most-preferred candidate under the conclusion $\tilde{\theta}_I^*(s, a_j)$,[16] and e^* maximizes u_I given the citizens' strategies.

This definition implies that although voters do not have standard beliefs about the incumbent competence, they do form correct conjectures about the incumbent equilibrium effort choice (for a model that loosens this assumption, see Ashworth and Bueno de Mesquita 2014). It also means that the incumbent understands the directional motives of voters, or at least plays a best response to their voting rule.

MOTIVATED REASONING AND VOTER CONCLUSIONS

In this section we analyze the novel mechanism of our model: how motivated reasoning affects the distribution of voter conclusions about incumbent performance.

Examples of Directional Motives

To better understand how our model of motivated reasoning works and generate intuitions for how it will affect voter behavior, we first provide three examples of the v function, which capture different notions of directional motives.

Example 1 (*Polarized partisanship*). It is natural to conceptualize citizens with positive affinity for the incumbent, $a_j > 0$, as the incumbent's copartisans and those with negative incumbent affinity, $a_j < 0$, as citizens from the opposite party. Under this conceptualization directional motives represent the idea that copartisans are motivated to believe that the incumbent is competent and those in the other party are motivated to believe that the incumbent is incompetent. Furthermore, those who more strongly identify with their partisan label (as represented by the magnitude of $|a_j|$) are more strongly motivated. This can be represented by the linear directional motive $v(a_j, \tilde{\theta}_I) = a_j \tilde{\theta}_I$. If voter j has affinity $a_j > 0$ $(a_j < 0)$, then he is motivated to conclude that the incumbent is highly competent (incompetent). With this assumption, the optimal conclusion maximizes

$$\log f_{\theta_I | s}(\tilde{\theta}_I | s) + \delta v(a_j, \tilde{\theta}_I) = \kappa - \frac{(\tilde{\theta}_I - \overline{\mu}(s))^2}{2\overline{\sigma}_\theta^2} + \delta a_j \tilde{\theta}_I,$$

where $\kappa = -\log(\overline{\sigma}_\theta^2) - \frac{1}{2}\log(2\pi)$ is a constant that does not depend on $\tilde{\theta}_I$. This is maximized at

$$\tilde{\theta}_I^{part}(s, a_j) = \overline{\mu}(s) + \delta \overline{\sigma}_\theta^2 a_j.$$

The first term is what a Bayesian would conclude $(\overline{\mu}(s))$, and so the second term captures the bias away from this ideal. Voter conclusions approach the mean of the Bayesian posterior $(\tilde{\theta}_I^{part}(s, a_j) \to \overline{\mu}(s))$ under three conditions. First, when the directional motive is very small, or $\delta \to 0$. Second is if voter affinity toward the incumbent is neutral, or $a_j \to 0$. Both of these follow from the fact that the v term approaches zero and the maximum of a normal density is at the mean. The third condition is if the Bayesian belief is very precise, or $\overline{\sigma}_\theta \to 0$, which leads to steep losses in the accuracy motive for even small deviations from the Bayesian mean. Conversely, a motivated reasoner of this form will tend to form a belief far from the Bayesian mean when δ, $|a_j|$ and $\overline{\sigma}_\theta$ are large.

Example 2 (*Confirmation bias*). Another type of motivation unrelated to affinity for the incumbent is reluctance to process information in order to confirm preexisting beliefs.[17] Because voters share a common prior with mean zero, this can be represented by

---

[16] By a standard argument, voters must behave sincerely if we rule out weakly dominated strategies.

[17] See Rabin and Schrag (1999) for a microfoundation for this kind of belief formation and Lockwood (2017) for a related study of confirmation bias in a political agency model with a focus on pandering dynamics.

$v(a_j, \tilde{\theta}_I) = -\tilde{\theta}_I^2$. In this case, voter j is motivated to conclude that his initial assessment of incumbent competence was accurate. He still forms a new conjecture about competence but is biased toward confirming his prior. Under this directional motive, the first-order condition for voter j's optimal conclusion becomes

$$-\frac{\tilde{\theta}_I - \overline{\mu}(s)}{\overline{\sigma}_\theta^2} - 2\delta\tilde{\theta}_I = 0,$$

which is solved by

$$\tilde{\theta}_I^{\text{conf}}(s, a_j) = \frac{1}{1 + 2\delta\overline{\sigma}_\theta^2}\overline{\mu}(s).$$

Confirmation bias creates a common bias for all citizens toward the prior that does not depend on a_j. Instead it takes the form of a weighted average of a fully Bayesian conclusion, $\overline{\mu}(s)$, and the (common) prior expectation of incumbent competence, which here is normalized to zero. Similar to the polarized partisanship motive, the optimal conclusion approaches that of a Bayesian as directional motives approach zero, $\delta \to 0$, and the precision of the Bayesian belief increases, $\overline{\sigma}_\theta \to 0$. However, in this case voter j's directional motives push him to confirm his preexisting assessment of the incumbent rather than forming highly positive or negative assessments. Thus, his affinity for the incumbent has no effect on distortions away from the Bayesian posterior mean. The manner in which voters form optimal conclusions in this example is analogous to the description of "cautious Bayesians" in Hill (2017). Conclusions are formed in the correct direction of the information received, due to weight on accuracy motives, but except under particular circumstances the conclusions will not reach those of a Bayesian given positive directional motives.

Example 3 (*Spatial motivations*). Another interpretation of directional motives is that there is a one-to-one correspondence between a citizen's affinity for the incumbent and his most preferred conclusion about incumbent competence. This can be represented by the quadratic directional motive $v(a_j, \tilde{\theta}_I) = -(a_j - \tilde{\theta}_I)^2$. Voter j is motivated to form a conclusion about incumbent quality that justifies his preexisting affinity for the incumbent. His affinity is like an "ideal conclusion target," much like an ideal point in standard spatial models. The directional motive in this case leads to the following optimal conclusion:

$$\tilde{\theta}_I^{\text{spat}}(s, a_j) = \frac{1}{1 + 2\delta\overline{\sigma}_\theta^2}\overline{\mu}(s) + \frac{2\delta\overline{\sigma}_\theta^2}{1 + 2\delta\overline{\sigma}_\theta^2}a_j.$$

As with the confirmation bias example, this can be thought of as a weighted average of what a Bayesian would conclude ($\overline{\mu}(s)$) and what a "pure motivated reasoner" ($\delta \to \infty$) would conclude (a_j). Also similar to the previous examples, as the directional motive

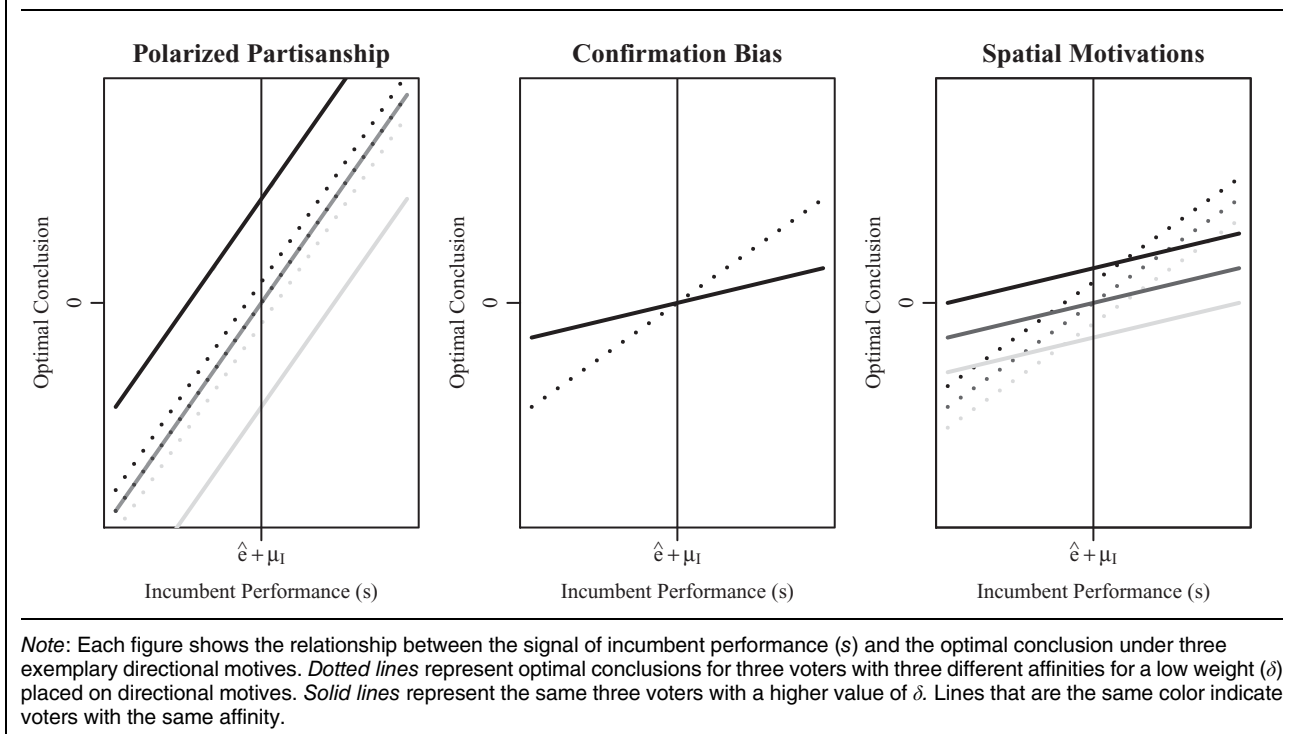**TABLE 1. Properties of Different Directional Motives**

Directional motive example	Belief divergence	Belief desensitization
Polarized partisanship	✓	X
Confirmation bias	X	✓
Spatial motivations	✓	✓

weakens, $\delta \to 0$, or precision of the Bayesian belief increases, $\overline{\sigma}_\theta \to 0$, the optimal conclusion approaches the Bayesian mean. An important difference with the spatial directional motive is that a voter with a neutral affinity (a_j close to zero) does *not* approximate the Bayesian mean (as long as $\overline{\mu}(s) \neq a_j$). This is because a centrist who does not strongly support or oppose the incumbent *does not lack a directional motive* but has a directional motive to form a conclusion close to zero. In other words, they are *motivated moderates* who intrinsically like holding a neutral view of the incumbent. As we will see when analyzing the voting model, this has important consequences for the politician's effort incentives.

As these examples illustrate, our model of motivated reasoning is very flexible and can incorporate a variety of voter biases. Citizens' directional motives have two principal effects on the relationship between a politician's performance and citizens' assessment of the politician. First, motivated reasoning may lead to *divergence*, meaning that voters' conclusions about politician quality differ based on their existing affinities for the incumbent. The second potential effect from directionally motivated reasoning is *desensitization*, which is a weakening of the relationship between incumbent performance and citizens' beliefs.

The two effects of directionally motivated reasoning are evident in the main examples from the fact that each optimal conclusion is decomposed into two components. One component is a "slope" term that multiplies the correct posterior mean/mode $\overline{\mu}(s)$: this represents the potential desensitization effects. In the partisan polarization example, the posterior mode is multiplied by one, so there is no desensitization effect. In the confirmation bias and spatial motivations examples, the posterior mode is multiplied by $\frac{1}{1 + 2\delta\overline{\sigma}_\theta^2} \in (0, 1)$, which indicates some level of desensitization in the sense that conclusions will depend less on the signal than under full Bayesian updating. Another component of each solution is an "intercept" term, which does not change the dependence of conclusions on the signal but does shift the baseline conclusion up or down. In examples 1 and 3, this intercept term is strictly increasing in a_j, which indicates a divergence effect. Table 1 displays the properties of the directional motives in Examples 1–3, and Figure 1 illustrates these three directional motives.

FIGURE 1. Illustration of the Three Main Examples of Directional Motives

Note: Each figure shows the relationship between the signal of incumbent performance (*s*) and the optimal conclusion under three exemplary directional motives. *Dotted lines* represent optimal conclusions for three voters with three different affinities for a low weight (δ) placed on directional motives. *Solid lines* represent the same three voters with a higher value of δ. Lines that are the same color indicate voters with the same affinity.

Each panel of Figure 1 plots the optimal conclusion for three voters as a function of the signal of incumbent performance. One voter has a positive affinity toward the incumbent $(a_j = .3,$ black lines), one a negative affinity toward the incumbent $(a_j = -.3$, light gray lines), and one is neutral $(a_j = 0,$ dark gray lines). The dotted lines correspond to a case with a relatively weak directional motive $(\delta = 1)$, and the solid lines a relatively strong directional motive $(\delta = 5)$. The vertical line corresponds to the average (expected) signal. Note this is a function of the voters' conjecture about the incumbent effort, $\widehat{e}$, which we formally derive in the following section.

The left panel illustrates the polarized partisan motive. All of the lines (between voters and across levels of motivated reasoning) are parallel, indicating no desensitization. However, comparing the dotted with the solid lines, we see there is divergence, as the solid lines are further apart from each other than the dotted lines. Another important aspect of this graph is that the centrist voter reaches the same conclusions (which correspond to the Bayesian mean) regardless of the level of motivated reasoning.

The middle panel illustrates the confirmation bias motive. Here, there is just one line because all of the voters reach the same conclusion independent of a_j. However, increasing the strength of the directional motive affects all voters, as this decreases the slope of the relationship between performance and conclusion. Thus, this motive illustrates a case of pure desensitization without divergence.

Finally, the right panel illustrates the spatial motivation. As strength of directional motive increases

(dotted to solid lines), the slopes of the conclusions decrease, again indicating desensitization. However, in contrast to confirmation bias, there is also divergence in this case, which is easiest to see by looking at the vertical line (average signal), where the pro- and anti-incumbent voters move further apart as the directional motives become stronger.

For all three of these examples, motivated reasoning is consistent with different voters responding to information in parallel (Guess and Coppock 2020), though it can change the slope of their updates (desensitization) and the difference between their intercept terms (divergence). As we will argue below, motivated reasoning that leads to desensitization presents more uniformly serious problems for politicians' responsiveness than does divergence.

Beyond the Examples

We now provide a more formal analysis of voter conclusions. Proposition 1 establishes the existence of a unique optimal conclusion for each voter and provides general results about how increasing the directional motive affects these conclusions.

Proposition 1. *Under Assumption 1,*
(i) there exists a unique optimal conclusion $\tilde{\theta}_I^*(s, a_j, \delta; \widehat{e})$ *for each voter* $j \in N$,

(ii) if $\frac{\partial^2 v(a_j, \bar{\theta}_I)}{\partial a_j \partial \bar{\theta}_I} > 0$, *then the optimal conclusion is strictly increasing in voter affinity* $\left(\frac{\partial \tilde{\theta}_I^*}{\partial a_j} > 0\right)$, *and the*

strength of this relationship is increasing in the directional motive $\left(\frac{\partial^2 \tilde{\theta}_I^*}{\partial \delta \partial a_j} > 0\right)$, and

(iii) the optimal conclusion is strictly increasing in the signal of performance $\left(\frac{\partial \tilde{\theta}_I^*}{\partial s} > 0\right)$, and if v is strictly concave in θ, then the strength of this relationship is decreasing in the directional motive $\left(\frac{\partial^2 \tilde{\theta}_I^*}{\partial \delta \partial s} < 0\right)$.

The first part proves analytically convenient: just as there is one Bayesian belief, there is a unique optimal conclusion for each voter (though this will depend on their affinity). The second part formalizes the notion of belief divergence: those who want to like the incumbent more will reach a more favorable conclusion, and this dependence grows stronger as the δ parameter increases. The third part formalizes the notion of belief desensitization: although voters with motivated beliefs still form more favorable conclusions about incumbents who perform better, this relationship becomes weaker when δ increases (provided there are diminishing returns to reaching more favorable conclusions).

We now have a general statement of how directional motives affect voters conclusions, both individually and in aggregate. There are two distinct effects—divergence and desensitization—which could in turn affect politician behavior. Further, Proposition 1 shows how these two components also correspond to the initial assumptions placed on the v function. As we move toward studying incumbent behavior, it will prove useful to be able to isolate these two effects so that we can see what aspects of motivated reasoning affect democratic accountability. To do so, we will place a stronger assumption on the directional motive, which allows for a clean decomposition.

Assumption 2. $\frac{\partial v}{\partial \theta_I}$ is linear in both $\tilde{\theta}_I$ and a_j.

Assumption 2 is satisfied for all of our main examples, and it is flexible enough to capture a wide range of directional motives.

Lemma 1 shows that these assumptions allow for a linear form of the optimal conclusion and provides a convenient way to characterize the two ways in which directional motives affect optimal conclusions, which correspond exactly to the discussion of Figure 1.

Lemma 1. *Under Assumptions 1 and 2,*

(i) the optimal conclusion is linear in a_j and s. In particular, it can be written as

$$\tilde{\theta}_I^*(s, a_j, \delta; \hat{e}) = \alpha_0 + \alpha_1 a_j + \beta \cdot (s - \hat{e}), \qquad (2)$$

where $\alpha_1 \geq 0$ and $\beta \geq 0$.

(ii) α_1 is strictly increasing in δ if and only if $\frac{\partial^2 v(a_j, \tilde{\theta}_I)}{\partial \tilde{\theta}_I \partial a_j} > 0$.

(iii) β is strictly decreasing in δ if and only if v is strictly concave in $\tilde{\theta}_I$.

The first term of the conclusion is an "intercept" term independent of the signal, $\alpha_0 + \alpha_1 a_j$, which depends more on the voter affinities (i.e., divergence effects are

stronger) as α_1 increases. As shown in Proposition 1, belief divergence is stronger as the directional motive increases (higher δ) if and only if those with strictly higher affinity for the incumbent want to form strictly higher conclusions. In our leading examples, this is true in the partisan polarization and spatial bias motives, but not confirmation bias.

The second term of the conclusion is a "slope" term $\beta \cdot (s - \hat{e})$ that captures how sensitive a voter is to changes in observed incumbent performance. As shown in Proposition 1, this term is strictly decreasing in δ (meaning more desensitization) if and only if the v function is strictly concave in $\tilde{\theta}_I$. In our leading examples this is true with the spatial bias and confirmation bias examples but not polarized partisanship.

MOTIVATED REASONING AND INCUMBENT PERFORMANCE

Now that we understand how motivated reasoning affects voters' conclusions, an analysis of incumbent behavior parallels standard models. As $\tilde{\theta}^*$ is increasing in a_j, voters' preferences satisfy single crossing: if voter j weakly prefers to retain the incumbent and $a_j < a_k$ for some other voter k, then voter k strictly prefers to retain the incumbent. Assuming indifference is broken in favor of the incumbent, voter j votes for the incumbent if and only if

$$\tilde{\theta}_I^*(s, a_j, \delta; \hat{e}) + a_j + \eta_I \geq \mu_C + \eta_C, \qquad (3)$$

where the left-hand side is strictly increasing in a_j (as above, we let $\hat{e} > 0$ denote the effort level that voters expect from the incumbent). As a result, the voter with the median affinity is decisive. Let a_m represent the median voter's affinity for the incumbent and r_m^* his equilibrium vote choice.[18]

Corollary 1. *In any equilibrium, the median voter is decisive: $R = 1$ if and only if $\tilde{\theta}_I^*(s, a_m, \delta; \hat{e}) + a_m + \eta_I \geq \mu_C + \eta_C$.*

Incumbent effort

To analyze which aspects of directionally motivated reasoning matter for incumbent effort, we rely on the characterization of the voter conclusion from Lemma 1. This allows us to study the effect of divergence by asking how changes in α_1 affect incumbent behavior

---

[18] A primary assumption driving median responsiveness is that the shocks η_I and η_C affect all voters the same rather than being voter specific as in other models (see Coughlin 1992). The exact nature of the shocks, however, is not crucial to the general results. A model with aggregate and individual uncertainty (introduced through individual preference shocks) would resemble our model with a large number of voters, where uncertainty for the politician diminishes with the number of voters. A model with no shocks leads to qualitatively similar results except in knife-edge cases where motivated reasoning never influences politician behavior.

and to study the effect of desensitization by varying the β parameter. Further, because the median voter is decisive we can analyze how these terms influence the optimal conclusion of the median voter.

What is the effect of divergence on effort? Formally, does increasing α_1 harm incumbent effort incentives? The answer to this question is ambiguous and characterized in Proposition 2. For this result, we say that the incumbent is "ahead" ("behind") if her ex ante probability of reelection given the expected level of effort before any random variables are realized is greater (less) than one half. If this probability is exactly one half, we say the election is a "dead heat." (A formal statement of this condition is in the proof of the following proposition.)

Proposition 2. *Under Assumption 2,*
(i) If $a_m = 0$ or the election is a dead heat, then divergence has no influence on incumbent effort.
(ii) If $a_m \neq 0$, then increasing divergence (i.e., increasing α_1) increases effort when either the incumbent is behind and $a_m > 0$ or the incumbent is ahead and $a_m < 0$ and decreases effort otherwise.

Because the median is decisive, divergence effects only matter for politician behavior if the median is drawn to one side or the other. We may expect that the voter with median affinity, by virtue of his position in the middle, is not strongly biased toward one candidate or another. If the median voter is exactly neutral ($a_m = 0$), then divergence has no effect on incumbent effort.[19] However, if the median voter is biased for or against the incumbent, then divergence mainly affects the marginal return on effort by making the election either closer or less close. For this reason, the effect of α_1 is ambiguous: if the incumbent is behind in the election ex ante, then increasing α_1 will reduce effort if the median voter dislikes the incumbent ($a_m < 0$) and increase effort if he likes the incumbent ($a_m > 0$). Conversely, if the incumbent is ahead ex ante, then increasing α_1 will tend to decrease effort when the median voter likes the incumbent ($a_m > 0$) and increase effort when he dislikes her ($a_m < 0$) because electoral competitiveness increases. Overall, the effect of divergence on incumbent effort depends on whether ex ante outcomes are pushed closer or further away from an even vote share.

The effect of desensitization on incumbent effort is much more straightforward. Whenever belief desensitization increases (i.e., β decreases) incumbent effort decreases.

Proposition 3. *Under Assumption 2, incumbent effort is reduced by desensitization effects of motivated reasoning (e^* is increasing in β).*

Given Lemma 1, this implies that whenever the directional motive is strictly concave in $\tilde{\theta}_I$, stronger directional motives will lead to less effort through this channel. In the extreme, sufficiently strong directional motives completely diminish the return to effort for the incumbent.

Our final result in this section combines this observation with the fact that, absent desensitization, making the directional motive arbitrarily strong will make all voters with $a_m \neq 0$ insensitive to the election result (and thus incumbent effort) because they will always vote for the candidate for which they have stronger affinity. As a result, except for the special case where there is no desensitization and the median voter is unaffected by divergence, when directional motives become arbitrarily strong, the overall effect of motivated reasoning is to eliminate all incentives for effort.

Corollary 2. *Under Assumptions 1 and 2, if there is desensitization $\left(\frac{\partial^2 v(a_j, \tilde{\theta}_I)}{\partial \tilde{\theta}_I^2} < 0 \right)$ or divergence influences the median voter $\left(a_m \neq 0 \text{ and } \frac{\partial^2 v(a_j, \tilde{\theta}_I)}{\partial \tilde{\theta}_I \partial a_j} > 0 \right)$, then as $\delta \to \infty$, $e^* \to 0$.*

To illustrate, Propositions 2 and 3 help us make predictions regarding Examples 1–3. What effect would each type of directionally motivated reasoning by voters have on politician behavior? Let us first consider the case in which $a_m = 0$ (i.e., the median voter has no clear preference for one candidate over the other). In such a case, the fact that voter conclusions diverge due to directional motives makes no difference for incumbent effort. In fact, from the incumbent's perspective, the polarized partisanship in Example 1 is no different at all from when voters are fully Bayesian. However, Examples 2 and 3 also exhibit belief desensitization. In those cases, the presence of directionally motivated reasoning decreases incumbent effort because all voters, including the median, respond less to government performance information, which weakens incumbent effort incentives.

When $a_m \neq 0$ our predictions are more subtle. Figure 2 illustrates an example of how the partisan directional motive (Example 1) affects equilibrium effort and the incumbent's reelection probability. In each panel, the x-axis is the strength of the directional motive, so these plots illustrate how increasing the strength of voters' partisan motivations affects equilibrium outcomes. There are three cases: when the prior (Bayesian) belief is that the challenger is weaker than the incumbent ($\mu_C < 0$; light gray), of equal competence on average ($\mu_C = 0$; dark gray), or more competent than the incumbent ($\mu_C > 0$; black).

The left panel shows that if the challenger is believed to be weak or equally competent, stronger motivated reasoning always decreases incumbent effort. This is because, as we can see in the right panel, these are cases

---

[19] The result that divergence does not affect incentives when $a_m = 0$ is partially driven by the common-shocks assumption in that some alternative shock structures might lead to a postshock median with some other affinity. However, in that case as the voting population grows large and the incumbent becomes increasingly confident that the median will be the voter with $a_m = 0$, the effect of divergence on effort incentives will shrink toward zero, in line with the qualitative thrust of this result.

FIGURE 2.   Equilibrium Effort and Reelection Probability with Partisan Directional Motives and a Median Voter with a Small Positive Affinity for the Incumbent
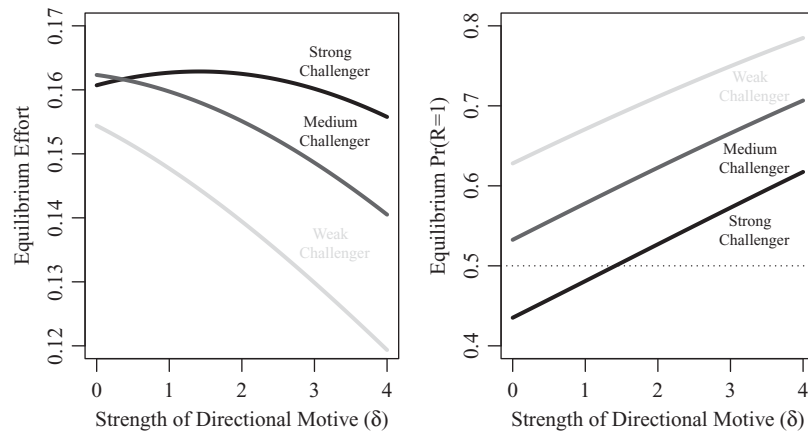
FIGURE 3.   Equilibrium Effort and Reelection Probability with Spatial Directional Motives and a Median Voter with a Small Affinity for the Incumbent
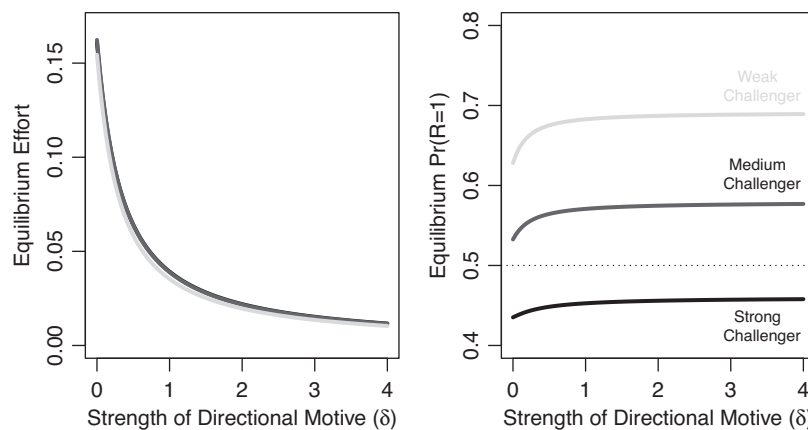
where the incumbent is likely to win. Because the median voter has a slight positive affinity for the incumbent, stronger directional motives make this advantage stronger, decreasing returns to effort.

When the challenger is expected to be more competent than the incumbent, as in the right panel, we can see that with no directionally motivated reasoning (low δ) the incumbent is more likely to lose the election. For this part of the parameter space, increasing directionally motivated reasoning makes the election "closer," as the median voter affinity for the incumbent counteracts the fact that she is objectively less competent than the challenger. Thus, at low levels of δ, increasing directionally motivated reasoning leads to *more* incumbent effort. However, as Corollary 2 shows, when the directional motives become strong enough, eventually the incumbent is advantaged despite his relatively low competence and further increases in δ strengthen this advantage, which leads to decreased equilibrium effort.

Figure 3 provides the same illustrations but using the spatial directional motive (Example 3), where there is desensitization in addition to divergence. Here the effect of increasing the directional motive on equilibrium effort is so strong (and, as expected, negative) that it swamps any differences based on how strong the challenger is. Though, as we can see in the right panel, there are still large differences across reelection probabilities.

Comparing across the figures, increased directionally motivated reasoning that manifests primarily as desensitization leads to strong and unambiguous decreases in incumbent effort. On the other hand, increases in motivated reasoning that only affect belief divergence have generally weaker effects that can run in either direction.

ELECTORAL SELECTION AND VOTER WELFARE

So far we have discussed the effects of motivated reasoning on politicians' incentives to invest in good

governance, but we have not discussed the overall effect of motivated reasoning on voter welfare. The total effect of motivated reasoning on voter welfare includes its effect on the incumbent's effort level as well as its effect on the ability of voters to select good politicians. Consider the ex ante welfare of the median voter:[20]

$$
W_m\big(e^*, r_m^*(s, \eta_I, \eta_C)\big) =
$$
$$
\mathbb{E}_{(\theta_I, \varepsilon)}\big[\theta + e^* + \varepsilon + \mathbb{E}_{(\eta_I, \eta_C)}\big[r_m^*(\theta_I + e^* + \varepsilon, \eta_I, \eta_C)
$$
$$
(\theta_I + a_j + \eta_I) + \big(1 - r_m^*(\theta_I + e^* + \varepsilon, \eta_I, \eta_C)\big)(\mu_C + \eta_C)\big]\big].
$$
$$(4)$$

Motivated reasoning has two effects on the median voter's welfare. First, as we have already discussed, motivated reasoning may affect equilibrium effort, e^*. Second, motivated reasoning may affect second-period expected utility by changing the retention rule r_m^*, which depends on the optimal conclusion given δ. By Lemma 1, the optimal conclusion $\tilde{\theta}_I^*$ is linear in $s - \hat{e}$. Because $e^* = \hat{e}$ in equilibrium, this implies that the selection term in voter welfare is independent of effort. Thus, the effect of motivated reasoning decomposes the total effect of δ into two separable effects on accountability and selection.

Examination of Equation 4 reveals that motivated reasoning must always have a negative effect on the selection component of voter welfare. When $\delta = 0$ the voters' decisions are identical to those of Bayesian decision makers, so voter welfare becomes

$$
\mathbb{E}_{(\theta_I, \varepsilon)}\big[\theta + e^* + \varepsilon + \mathbb{E}_{(\eta_I, \eta_C)}[max\{\mathbb{E}[\theta_i|\theta_I + \varepsilon]
$$
$$
+ a_j + \eta_I, \mu_C + \eta_C\}]\big].
$$

In other words, when $\delta = 0$ the median voter's retention strategy is the one that maximizes second-period expected utility given the available information because $\tilde{\theta}_I^*(s, a_j, \delta; \hat{e}) = \mathbb{E}[\theta_i|\theta_I + \varepsilon]$ (i.e., the optimal conclusion is the same as the correct conditional expectation). Adding motivated reasoning ($\delta > 0$) can only decrease second-period expected utility either by retaining the incumbent when $\mathbb{E}[\theta_i|\theta_I + \varepsilon] + a_j + \eta_I < \mu_C + \eta_C$ or removing him when $\mathbb{E}[\theta_i|\theta_I + \varepsilon] + a_j + \eta_I > \mu_C + \eta_C$.

Several conclusions follow about how motivated reasoning affects voter welfare. First, motivated reasoning has two effects on voter welfare: a strategic effect on effort by politicians and a statistical effect on selection of good politicians. Second, when motivated reasoning decreases effort by politicians it must always decrease voter welfare overall. Finally, when motivated reasoning increases effort by politicians it may still decrease voter welfare through its effect on selection

(though this could be counterbalanced by positive strategic effects).[21]

Our analysis of voter welfare takes an objective perspective on the competence of the incumbent. Alternatively we could consider welfare from the perspective of voters with biased beliefs. The main implication of doing so would be to erase the negative selection effects: because voters are selecting optimally in their own opinions, they see selection as improved relative to selection under Bayesian beliefs. This would lead us to more often conclude that motivated reasoning is good for voter welfare.

THE RELATIONSHIP BETWEEN PERFORMANCE AND VOTE SHARES

So far we have studied how different kinds of directional motives affect accountability. In our model this largely depends on the behavior of the median voter. However, a large body of empirical work on electoral accountability focuses on the relationship between aggregate vote shares—which depend on the behavior of all voters—and measures of incumbent success, most frequently economic performance (e.g., Erikson 1989; Fair 1996; Healy and Lenz 2017; Hopkins and Pettingill 2018; Markus 1992; Tufte 1976).

In this section we adapt our model to show that motivated reasoning can affect incumbent vote share without affecting accountability. To do so, we explore how different directional motives influence the relationship between expected incumbent vote share given government performance. For a fixed signal of government performance s, the probability that voter j votes to retain the incumbent is

$$
Pr\big(\tilde{\theta}_I^*(s, a_j, \delta; \hat{e}) + a_j + \eta_I \geq \mu_C + \eta_C\big).
$$

To convert this observation into expected vote shares, we maintain Assumption 2 and further assume that voter affinities are normally distributed with mean μ_a and variance σ_a^2. Under these assumptions, we can write the condition for j to vote for the incumbent as

$$
\alpha_0 + \alpha_1 a_j + \beta_j \cdot (s - e^*) + a_j + (\eta_I - \eta_C) \geq \mu_C. \quad (5)
$$

The left-hand side of Equation 5 (for a fixed realization of s) is normally distributed with mean $\alpha_0 + \beta_j(s - e^*) + (1 + \alpha_1)\mu_a + \mu_\eta$ and variance $(1 + \alpha_1)^2 \sigma_a^2 + \sigma_\eta^2$, so the incumbent's average vote share is given by

[21] The possibility of differing effects of motivated reasoning on selection versus accountability also raises interesting possibilities about what may happen if politicians hold incorrect beliefs about δ. For instance, if politicians believe that voters are motivated reasoners when they are not, we might see the accountability effects of motivated reasoning with none of the negative selection effects. This raises the possibility that such incorrect beliefs may sometimes be beneficial to voters. Similarly, if politicians believe that voters are rational when they are not, we will see the selection effects of motivated reasoning with none of the strategic accountability effects.

$$\mathbb{E}[VS|s] = \Phi\left(\frac{\alpha_0 + \beta_j(s-e^*) + (1+\alpha_1)\mu_a + \mu_\eta - \mu_C}{\sqrt{(1+\alpha_1)^2\sigma_a^2 + \sigma_\eta^2}}\right).$$

As noted above, many empirical papers studying electoral accountability examine the relationship between incumbent vote shares and some performance indicator (in our model, s) such as change in GDP, change in real income, unemployment, crime, etc. In our formulation the theoretical prediction for this relationship is given by differentiating $\mathbb{E}[VS|s]$ with respect to s, which we denote with $\Delta^{VS}(s)$:

$$\Delta^{VS}(s) = \frac{\partial\mathbb{E}[VS|s]}{\partial s} = \frac{\beta_j}{\sqrt{(1+\alpha_1)^2\sigma_a^2 + \sigma_\eta^2}}\phi$$
$$\left(\frac{\alpha_0 + \beta_j(s-e^*) + (1+\alpha_1)\mu_a + \mu_\eta - \mu_C}{\sqrt{(1+\alpha_1)^2\sigma_a^2 + \sigma_\eta^2}}\right).$$

When $\Delta^{VS}(s) = 0$, there is no relationship between observed government performance and incumbent vote share. When $\Delta^{VS}(s) \neq 0$, performance information does influence the incumbent's average vote share. The next result establishes that even when the effects of motivated reasoning do not influence equilibrium, effort there can nonetheless be an effect on vote shares.

Remark 1. *Suppose Assumption 2 holds and voter affinities are normally distributed with mean μ_a and variance σ_a^2. Further, let $a_m = 0$, $\mu_a = 0$, and $\frac{\partial\beta}{\partial\delta} = 0$ so that there is divergence but no desensitization. Then motivated reasoning can affect incumbent vote share even when it does not affect equilibrium effort.*

The preceding analysis showed that when $a_m = 0$ and motivated reasoning only causes divergence in voter conclusions, there is no change in incumbent effort. However, this does not mean divergence does not affect the relationship between performance and vote share.

For example, consider the polarized partisan directional motive in Example 1: as the strength of the directional motive grows arbitrarily large ($\delta \to \infty$), $\alpha_1 \to \infty$, which leads to $\Delta^{VS}(s) \to 0$. This implies that the relationship between performance s and incumbent vote share diminishes as directional motives dominate voter conclusions. Moreover, given part (i) of Proposition 2, this implies that even when there is no effect of directional motives on incumbent effort there can still be an effect on the relationship between performance and vote shares. Intuitively, this is because when $a_m = 0$ and the belief divergence effect is very strong, nearly all voters with positive affinity for the incumbent $a_j > 0$ vote to retain him and nearly all voters with negative affinity for the incumbent $a_j < 0$ vote to remove him, yet the median voter remains decisive for both the electoral outcome and for providing incumbent effort incentives. More generally, motivated reasoning
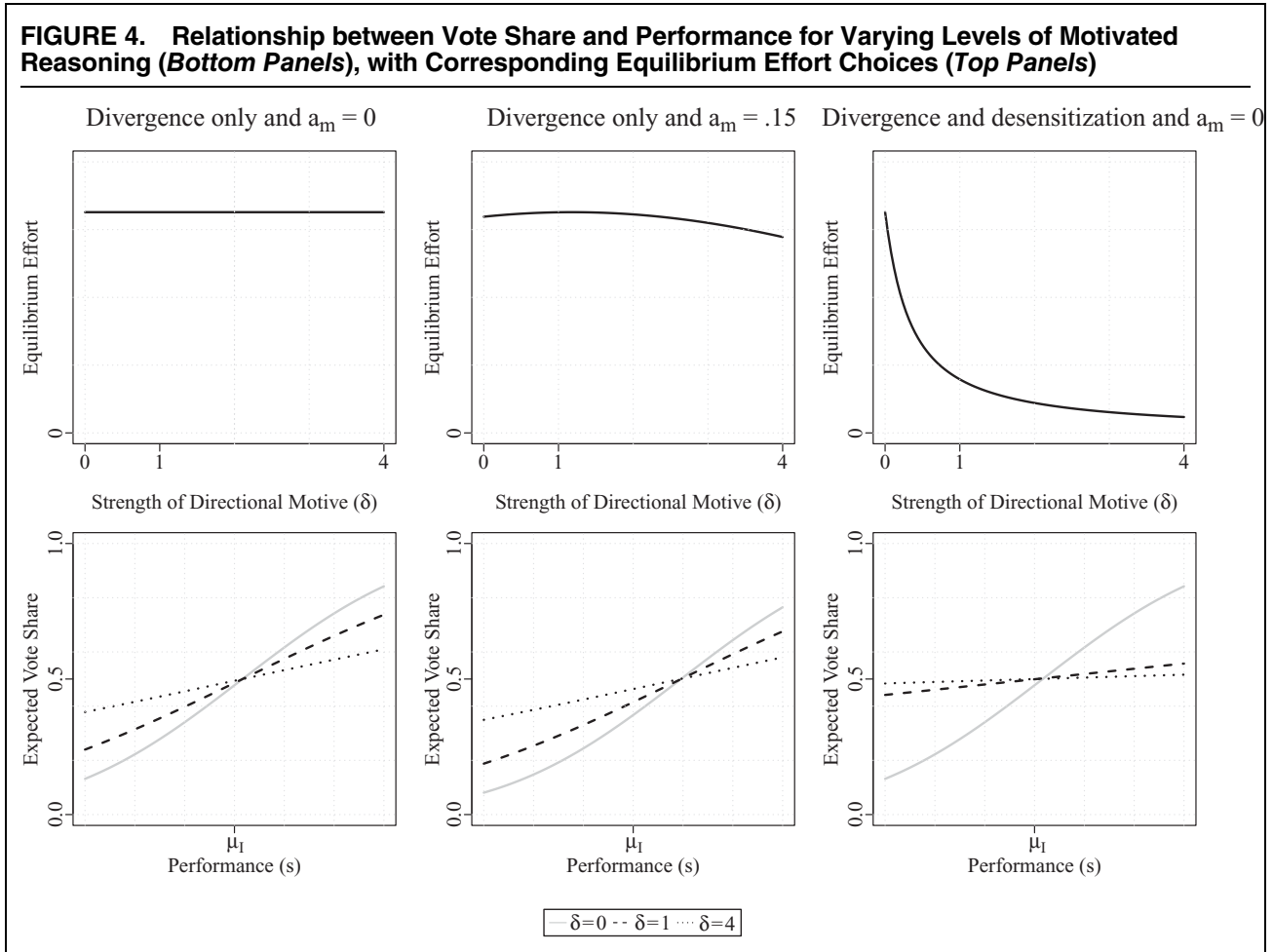
producing belief divergence may change the relationship between incumbent performance and *vote shares* even if it does not affect accountability.

Figure 4 provides illustrative examples. The top panels show the relationship between the strength of the directional motive (δ) and effort choice of the incumbent for different model specifications. Each panel illustrates equilibrium effort as a function of the directional motive δ. On the x-axis we highlight three values of δ, each corresponding to the bottom panel, which plots the relationship between incumbent performance s and average vote share at those three levels of directional motive. The solid curve corresponds to the case with no directional motives ($\delta = 0$), the dashed curve to "low" directional motives ($\delta = 1$), and the dotted curve to "high" directional motives ($\delta = 4$).

The left-most panels illustrate the environment described by Remark 1. There is a partisan directional motive as in Example 1, with a median voter who is an exact centrist, $a_m = 0$. In this case, variation in directional motive δ has no effect on effort incentives. However, the bottom-left panel shows that increasing the directional motive (solid to dashed to dotted curve) weakens the relationship between performance and vote share. Thus, the relationship between performance and vote share can become weaker *while having no effect on accountability.*

The middle panel also models the partisan motive, but now the median voter has a slight positive affinity for the incumbent ($a_m = .15$). The challenger is also ex ante more likely to be competent ($\mu_C = .4$) and therefore enjoys a competence-based electoral advantage when there are weak directional motives. As the analysis above shows, increasing δ, which intensifies the median's positive affinity for the incumbent, "tightens" the election by reducing the challenger's advantage. This translates into an initial increase in effort. However, looking now at the corresponding bottom panel, the relationship between vote shares and performance is weaker moving from $\delta = 0$ (solid curve) to $\delta = 1$ (dashed curve). This illustrates that the connection between incumbent success and vote shares can be loosened even when there is a positive effect on effort incentives. Once δ increases sufficiently, the median's positive affinity pushes the incumbent "ahead," thereby reducing effort incentives. Intuitively, in this case a weaker relationship between performance and vote shares corresponds to weakened effort incentives. This can be seen by comparing either $\delta = 0$ with $\delta = 4$ (dotted curve) or $\delta = 1$ with $\delta = 4$. It is worth noting, however, that although the slope of the dotted line is relatively flat compared with either the dashed or solid curves, it corresponds to about 90% of the effort incentives. In this case at least, very different relationships between performance and vote share are indicative of relatively modest changes in accountability.

Finally, the right panels represent a case with the spatial directional motive, with an exact centrist median voter ($a_m = 0$). This is a case then where motivated

FIGURE 4. Relationship between Vote Share and Performance for Varying Levels of Motivated Reasoning (*Bottom Panels*), with Corresponding Equilibrium Effort Choices (*Top Panels*)

reasoning leads to both belief divergence and desensitization. Accordingly, we see that effort decreases rapidly in directional motive (δ). In this case the relationship between performance and incumbent vote share is also strongly affected, moving from $\delta = 0$ to either $\delta = 1$ or $\delta = 4$, with only modest difference between the latter two levels. As directional motives intensify, the dual effects of divergence and desensitization lead to weak accountability and a weak relationship between incumbent performance and vote shares.

A weakening relationship between incumbent performance and vote shares might be a symptom of weakening democratic accountability, but the degree to which this is true depends on what exactly drives this attenuation. If driven by desensitization, then a weakening relationship between incumbent performance and vote shares is cause for alarm. However, if it is mostly driven by belief divergence then this is not necessarily the case. Further, because effort can sometimes increase in belief divergence, it is possible to see a weaker performance–vote share relationship even while the incumbent faces stronger accountability incentives.

DISCUSSION: BEHAVIORAL EQUIVALENCIES

We have shown different ways in which motivated reasoning might harm democratic accountability. But an important question is whether these are unique to our model in which voters trade off accuracy and directional motivations. A benefit of reducing the total effect of motivated reasoning to two channels—belief divergence and belief desensitization—is that we can map any realization of our model with voters who are motivated reasoners to a "standard" model where voters process information only using Bayes' rule.

In particular, take any set of parameters to the model meeting Assumption 2, where $\delta > 0$ and voters have affinities $\{a_j\}$. In this starting model, the voters choose to reelect if and only if

$$\left(\alpha_0 + \alpha_1 a_j + \beta(s - \widehat{e})\right) + a_j \geq \mu_C.$$

Now consider a modified version of the model where voters form their beliefs about θ_I using Bayes' rule. All other parameters are kept fixed except we add a parameter w that scales how much the voter cares about

incumbent performance (relative to affinity). Further, let the voters have a different set of affinities $\{a'_j\}$. By standard analysis, voters in such a model will vote to reelect if and only if

$$w(\alpha_0 + \beta'(s - \widehat{e})) + a'_j \geq \mu_C,$$

where α_0 is the same as in our main model and $\beta' = \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + \sigma_\theta^2}$. If we set

$$w = \beta/\beta'$$
$$a'_j = a_j(1 + \alpha_1),$$

then the behavior of all voters is identical in this modified game (for any fixed conjecture $\widehat{e}$), which implies the incumbent has equivalent incentives and will choose the same effort level. That is, for any version of our model with motivated reasoning, there is a modified version of the model with fully Bayesian voters where (1) voters put less weight on incumbent performance and (2) have more "spread-out" affinities and equilibrium behavior is identical.[22] It is possible to obtain similar results by changing the variances of the prior belief (how informed voters are about politics at the outset) and noise term ε (how much attention voters pay to the signal of performance or how much control politicians have over performance).[23] For some results, it would also be possible to do away with the affinities and give citizens private signals about the incumbent or opposition competence, where those with more favorable signals will behave like those with higher affinity.

In other words, while motivated reasoning certainly matters for democratic accountability, the effects are not qualitatively distinct from reformulating voter preferences or information in more standard models. As directional motives strengthen in our motivated-reasoning model, voters behave as though they are more polarized in terms of preferences and/or place less emphasis on, or pay less attention to, objective performance indicators when reaching conclusions about incumbent desirability. These general dynamics can also manifest in a standard accountability model through changes to the preference environment. Thus, many of the comparative statics from canonical models of electoral accountability mimic the effects of non-Bayesian belief formation present in our model.

CONCLUSION

Ultimately, we offer qualified agreement with the argument that voter biases that arise through directionally motivated reasoning may be bad news for democracy. Our results suggest that we should expect some aspects of directionally motivated reasoning to have negative

consequences for democratic accountability. However, the findings most often seen as the primary subject of concern are that partisans of different stripes hold different beliefs on statements of fact. This type of belief divergence does not seem to have a uniform, consistent effect on politicians' performance incentives. The larger concern, we argue, is the extent to which motivated reasoning may weaken the relationship between a politician's performance and conclusions about that politician. If motivated reasoning is a problem for democracy, the corrective may be to help voters become more sensitive to new information rather than focusing on reducing partisanship.

How might voters' sensitivity to information be increased? A standard answer to this question is to improve the informativeness of the signals voters receive: here, decreasing σ_ε^2. The addition of motivated reasoning does not overturn this result, and, in fact, it provides another channel through which it is true. With both the spatial motivations and confirmation bias examples, the weight that is placed on the posterior belief is decreasing in the variance of the posterior belief about θ_I. More precise beliefs lead to larger penalties in the accuracy motive in these examples when moving away from the fully Bayesian belief. Even those with strong directional motives will tend toward the fully Bayesian conclusion when anything far from it is highly implausible, which is the scenario when σ_ε^2 is very small. Substantively, we may improve voter informational sensitivity through improved factual misperception correction (or fact-checking), improving the quality of political news and also the access to high-quality political news. Interventions of this sort are likely to reduce uncertainty, which will in turn increase the value for voters to reach accurate conclusions even if they also have directional motivations.

SUPPLEMENTARY MATERIALS

To view supplementary material for this article, please visit http://doi.org/10.1017/S0003055421001209.

CONFLICT OF INTEREST

The authors declare no ethical issues or conflicts of interest in this research.

---

[22] Formally, as we have shown that in the main model β is decreasing in δ and α_1 is increasing in δ, we know that $w < 1$ in the modified game. So, for any voters k and l such that $a_k \neq a_l$, $|a'_k - a'_l| > |a_k - a_l|$.

[23] This equivalence is weaker in the sense that different priors entail changes to the distribution of the incumbent performance signals and thus the distribution of voter behavior across realizations of the game (if not for a particular realization of performance).

ETHICAL STANDARDS

The authors affirm this research did not involve human subjects.

REFERENCES

Acharya, Avidit, Matthew Blackwell, and Maya Sen. 2018. "Explaining Preferences from Behavior: A Cognitive Dissonance Approach." *The Journal of Politics* 80 (2): 400–11.

Achen, Christopher H., and Larry M. Bartels. 2017. *Democracy for Realists: Why Elections Do Not Produce Responsive Government*. Princeton, NJ: Princeton University Press.

Adida, Claire, Jessica Gottlieb, Eric Kramon, and Gwyneth McClendon. 2017. "Reducing or Reinforcing In-Group Preferences? An Experiment on Information and Ethnic Voting." *Quarterly Journal of Political Science* 12 (4): 437–77.

Albertson, Bethany, and Shana Kushner Gadarian. 2015. *Anxious Politics: Democratic Citizenship in a Threatening World*. New York: Cambridge University Press.

Ashworth, Scott. 2012. "Electoral Accountability: Recent Theoretical and Empirical Work." *Annual Review of Political Science* 15: 183–201.

Ashworth, Scott, and Anthony Fowler. 2019. "Electorates vs. Voters." Unpublished Manuscript. University of Chicago.

Ashworth, Scott, and Ethan Bueno de Mesquita. 2014. "Is Voter Competence Good for Voters? Information, Rationality, and Democratic Performance." *American Political Science Review* 108 (3): 565–87.

Bartels, Larry M. 2002. "Beyond the Running Tally: Partisan Bias in Political Perceptions." *Political Behavior* 24 (2): 117–50.

Bendor, Jonathan, Daniel Diermeier, David A. Siegel, and Michael M. Ting. 2011. *A Behavioral Theory of Elections*. Princeton, NJ: Princeton University Press.

Benjamin, Daniel J. 2019. "Errors in Probabilistic Reasoning and Judgment Biases." Chap. 2 in *Handbook of Behavioral Economics: Applications and Foundations 1*, eds. B. Douglas Bernheim, Stefano DellaVigna, and David Laibson. Princeton, NJ: North–Holland.

Bisgaard, Martin. 2019. "How Getting The Facts Right Can Fuel Partisan-Motivated Reasoning." *American Journal of Political Science* 63 (4): 824–39.

Bolsen, Toby, James N. Druckman, and Fay Lomax Cook. 2014. "The Influence of Partisan Motivated Reasoning on Public Opinion." *Political Behavior* 36 (2): 235–62.

Bullock, John G., Alan S. Gerber, Seth J. Hill, and Gregory A. Huber. 2015. "Partisan Bias in Factual Beliefs about Politics." *Quarterly Journal of Political Science* 10 (4): 519–78.

Bullock, John G., and Gabriel Lenz. 2019. "Partisan Bias in Surveys." *Annual Review of Political Science* 22: 325–42.

Coughlin, Peter J. 1992. *Probabilistic Voting Theory*. Cambridge: Cambridge University Press.

de Benedictis-Kessner, Justin, and Christopher Warshaw. 2020. "Accountability for the Local Economy at All Levels of Government in United States Elections." *American Political Science Review* 114 (3): 660–76.

de Vries, Catherine E., Sara B. Hobolt, and James Tilley. 2018. "Facing up to the Facts: What Causes Economic Perceptions?" *Electoral Studies* 51: 115–22.

Delli Carpini, Michael X., and Scott Keeter. 1996. *What Americans Know about Politics and Why It Matters*. New Haven, CT: Yale University Press.

Diermeier, Daniel, and Christopher Li. 2017. "Electoral Control with Behavioral Voters." *The Journal of Politics* 79 (3): 890–902.

Diermeier, Daniel, and Christopher Li. 2019. "Partisan Affect and Elite Polarization." *American Political Science Review* 113 (1): 277–81.

Donovan, Kathleen, Paul M Kellstedt, Ellen M. Key, and Matthew J. Lebo. 2020. "Motivated Reasoning, Public Opinion, and Presidential Approval." *Political Behavior* 42: 1201–21.

Druckman, James N., Erik Peterson, and Rune Slothuus. 2013. "How Elite Partisan Polarization Affects Public Opinion Formation." *American Political Science Review* 107 (1): 57–79.

Druckman, James N., and Mary C. McGrath. 2019. "The Evidence for Motivated Reasoning in Climate Change Preference Formation." *Nature Climate Change* 9 (2): 111–19.

Duggan, John, and César Martinelli. 2017. "The Political Economy of Dynamic Elections: Accountability, Commitment, and Responsiveness." *Journal of Economic Literature* 55 (3): 916–84.

Dunning, Thad, Guy Grossman, Macartan Humphreys, Susan D. Hyde, Craig McIntosh, and Gareth Nellis. 2019. *Information, Accountability, and Cumulative Learning: Lessons from Metaketa I*. New York: Cambridge University Press.

Erikson, Robert S. 1989. "Economic Conditions and the Presidential Vote." *American Political Science Review* 83 (2): 567–73.

Fair, Ray C. 1996. "The Effect of Economic Events on Votes for President: 1992 Update." *Political Behavior* 18 (2): 119–39.

Fearon, James D. 1999. "Electoral Accountability and the Control of Politicians: Selecting Good Types versus Sanctioning Poor Performance." In *Democracy, Accountability, and Representation*, eds. Adam Przeworski, Susan C. Stokes, and Bernard Manin, 55–97. Cambridge: Cambridge University Press.

Ferraz, Claudio, and Frederico Finan. 2008. "Exposing Corrupt Politicians: The Effects of Brazil's Publicly Released Audits on Electoral Outcomes." *The Quarterly Journal of Economics* 123 (2): 703–45.

Freeder, Sean. 2019. "It's No Longer the Economy, Stupid: Selective Perception and Attribution of Economic Outcomes." Unpublished Manuscript. University of California, Berkeley.

Gerber, Alan, and Donald Green. 1999. "Misperceptions about Perceptual Bias." *Annual Review of Political Science* 2: 189–210.

Glaeser, Edward L., and Giacomo A. M. Ponzetto. 2017. "Fundamental Errors in the Voting Booth." *NBER* Working Paper 23683.

Green, Donald P., Bradley Palmquist, and Eric Schickler. 2004. *Partisan Hearts and Minds: Political Parties and the Social Identities of Voters*. New Haven, CT: Yale University Press.

Groenendyk, Eric W. 2013. *Competing Motives in the Partisan Mind: How Loyalty and Responsiveness Shape Party Identification and Democracy*. New York: Oxford University Press.

Groenendyk, Eric W., and Yanna Krupnikov. 2020. "What Motivates Reasoning? A Theory of Goal-Dependent Political Evaluation." *American Journal of Political Science* 65 (1): 180–96.

Guess, Andrew, and Alexander Coppock. 2020. "Does Counter-Attitudinal Information Cause Backlash? Results from Three Large Survey Experiments." *British Journal of Political Science* 50(4): 1497–515.

Healy, Andrew, and Gabriel S. Lenz. 2017. "Presidential Voting and the Local Economy: Evidence from Two Population-Based Data Sets." *The Journal of Politics* 79 (4): 1419–32.

Healy, Andrew, and Neil Malhotra. 2013. "Retrospective Voting Reconsidered." *Annual Review of Political Science* 16: 285–306.

Healy, Andrew J., Neil Malhotra, and Cecilia Hyunjung Mo. 2010. "Irrelevant Events Affect Voters' Evaluations of Government Performance." *Proceedings of the National Academy of Sciences* 107 (29): 12804–809.

Hill, Seth J. 2017. "Learning Together Slowly: Bayesian Learning about Political Facts." *The Journal of Politics* 79 (4): 1403–18.

Hopkins, Daniel J., and Lindsay M. Pettingill. 2018. "Retrospective Voting in Big-City US Mayoral Elections." *Political Science Research and Methods* 6 (4): 697–714.

Huddy, Leonie, Lilliana Mason, and Lene Aarøe. 2015. "Expressive Partisanship: Campaign Involvement, Political Emotion, and Partisan Identity." *American Political Science Review* 109 (1): 1–17.

Jacobson, Gary C. 2010. "Perception, Memory, and Partisan Polarization on the Iraq War." *Political Science Quarterly* 125 (1): 31–56.

Kuklinski, James H., Paul J. Quirk, Jennifer Jerit, David Schwieder, and Robert F. Rich. 2000. "Misinformation and the Currency of Democratic Citizenship." *The Journal of Politics* 62 (3): 790–816.

Kunda, Ziva. 1990. "The Case for Motivated Reasoning." *Psychological Bulletin* 108 (3): 480–98.

Lau, Richard R., and David P. Redlawsk. 2006. *How Voters Decide: Information Processing in Election Campaigns*. New York: Cambridge University Press.

Levy, Gilat, and Ronny Razin. 2015. "Correlation Neglect, Voting Behavior, and Information Aggregation." *American Economic Review* 105 (4): 1634–45.

Little, Andrew T. 2019. "The Distortion of Related Beliefs." *American Journal of Political Science* 63 (3): 675–89.

Lockwood, Ben. 2017. "Confirmation Bias and Electoral Accountability." *Quarterly Journal of Political Science* 11 (4): 471–501.

Lodge, Milton, and Charles S. Taber. 2013. *The Rationalizing Voter*. New York: Cambridge University Press.

Lupia, Arthur, and Mathew D. McCubbins. 1998. *The Democratic Dilemma: Can Citizens Learn What They Need to Know?* New York: Cambridge University Press.

Markus, Gregory B. 1992. "The Impact of Personal and National Economic Conditions on Presidential Voting, 1956-1988." *American Journal of Political Science* 36 (3): 829–34.

Mayraz, Guy. 2019. "Priors and Desires: A Bayesian Model of Wishful Thinking and Cognitive Dissonance." Unpublished Manuscript. University of Sydney. http://mayraz.com/papers/PriorsAndDesires.pdf.

Minozzi, William. 2013. "Endogenous Beliefs in Models of Politics." *American Journal of Political Science* 57 (3): 566–81.

Morris, Stephen, and Hyun Song Shin. 2003. Global Games: Theory and Applications. In *Advances in Economics and Econometrics: Theory and Applications, 8th World Congress of the Econometric Society*, eds. Mathias Dewatripont, Lars Peter Hansen, and Stephen J. Turnovsky, 56–114. Cambridge: Cambridge University Press.

Nunnari, Salvatore, and Jan Zápal. 2017. "A Model of Focusing in Political Choice." *CEPR* Discussion Paper DP12407.

Nyhan, Brendan, and Jason Reifler. 2010. "When Corrections Fail: The Persistence of Political Misperceptions." *Political Behavior* 32 (2): 303–30.

Ogden, Benjamin. 2016. "An Imperfect Beliefs Voting Model." Unpublished Manuscript. Texas A&M University. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2431447.

Payson, Julia A. 2017. "When Are Local Incumbents Held Accountable for Government Performance? Evidence from US School Districts." *Legislative Studies Quarterly* 42 (3): 421–48.

Prior, Markus, Gaurav Sood, and Kabir Khanna. 2015. "You Cannot Be Serious: The Impact of Accuracy Incentives on Partisan Bias in Reports of Economic Perceptions." *Quarterly Journal of Political Science* 10 (4): 489–518.

Rabin, Matthew, and Joel L. Schrag. 1999. "First Impressions Matter: A Model of Confirmatory Bias." *The Quarterly Journal of Economics* 114 (1): 37–82.

Redlawsk, David P. 2002. "Hot Cognition or Cool Consideration? Testing the Effects of Motivated Reasoning on Political Decision Making." *The Journal of Politics* 64 (4): 1021–44.

Roush, Carolyn E., and Gaurav Sood. 2020. "A Gap in Our Understanding? Reconsidering the Evidence for Partisan Knowledge Gaps." Unpublished Manuscript. Florida State University. https://www.gsood.com/research/papers/partisan_gap.pdf.

Taber, Charles S., and Milton Lodge. 2006. "Motivated Skepticism in the Evaluation of Political Beliefs." *American Journal of Political Science* 50 (3): 755–69.

Thaler, Michael. 2020. "The Fake News Effect: An Experiment on Motivated Reasoning and Trust in News." Unpublished Manuscript. Harvard University.

Tufte, Edward R. 1976. *Political Control of the Economy*. Princeton, NJ: Princeton University Press.

Woon, Jonathan. 2012. "Democratic Accountability and Retrospective Voting: A Laboratory Experiment." *American Journal of Political Science* 56 (4): 913–30.