Economics Faculty Publications   Faculty Scholarship and Creative Work

4-1-2019

"Unmotivated bias" and partisan hostility: Empirical evidence

Daniel F. Stone
Bowdoin College

Follow this and additional works at: https://digitalcommons.bowdoin.edu/economics-faculty-publications

"Unmotivated bias" and partisan hostility: Empirical evidence☆

Daniel F. Stone

Department of Economics, Bowdoin College, Brunswick, ME, USA

ARTICLE INFO

ABSTRACT

Extreme partisan animosity has been on the rise in the US and is prevalent around the world. This hostility is typically attributed to social group identity, motivated reasoning, or a combination thereof. In this paper, I empirically examine a novel contributing factor: the "unmotivated" cognitive bias of overprecision (over-confidence in precision of beliefs). Overprecision could cause partisan hostility indirectly via inflated confidence in one's own ideology, partisan identity, or perceptions of social distance between the parties. Overprecision could also cause this hostility directly by causing excessively strong inferences from observed information that is either skewed against the out-party or simply misunderstood. Using a nationally representative sample, I find consistent support for direct effects of overprecision and mixed support for indirect effects. The point estimates imply a one standard deviation increase in a respondent's overprecision predicts as much as a 0.71 standard deviation decline in relative out-party favorability.

"There's no sense in being precise when you don't even know what you're talking about."
-John von Neumann

1. Introduction

Recent political polarization in the US has attracted much attention from scholars across the social sciences, with good reason. It is clear now that the actions and rhetoric of politicians of the two major parties have diverged sharply over the last half-century (Barber and McCarty, 2015; Andris et al., 2015; Gentzkow et al., 2015). It is more ambiguous to what extent, if any, voter ideologies have also become polarized (Hill and Tausanovitch, 2015; Lelkes, 2016). In just the last several years, however, more clear evidence has emerged regarding another type of voter polarization—*affective* polarization—that voters' *feelings* about the parties have polarized over time (Iyengar et al., 2012; Lelkes, 2016). In particular, hostility toward the "out-party," sometimes called partyism for short (Sunstein, 2015), has grown sharply, even controlling for partisan sorting, as shown in Fig. 1. Partyism contributes to partisan conflict and gridlock, and has other significant political and social implications.[1]

What causes partyism, and changes in partyism? Much of the literature in this area focuses on growth in "social distance" between the parties (see, e.g., Iyengar and Westwood, 2015) and strengthened partisan identity and "tribalistic" thinking (Greene, 2014; Mason, 2018). In many papers it is taken as given that people automatically dislike those in different social groups, more so as we identify more strongly with the in-group and/or perceive larger differences from the out-group. Papers that discuss or analyze the cognitive basis for such

[1] Hetherington and Rudolph (2015) discuss adverse welfare effects of affective polarization via reduced cooperation and compromise among policymakers, leading to lower productivity. Azzimonti (2018) provides evidence of adverse macroeconomic effects of partisan conflict. Partisan hostility has led to numerous violent incidents (e.g. the "MAGA bomber" and Congressional softball game shooting) and see Huber and Malhotra (2017) for evidence of other adverse social effects. See Potrafke (2018) for discussion of benefits of polarization vis-a-vis offering voters clearer policy choices, and Iyengar et al. (2019) for a broader review of affective polarization. I note also that while recent partyism in the US is the focus of this paper, partyism is certainly not unique to this environment. Westwood et al. (2018) show that partyism is prevalent around the world and see, for example, Selinger (2016) for discussion of historical partisan hostility.
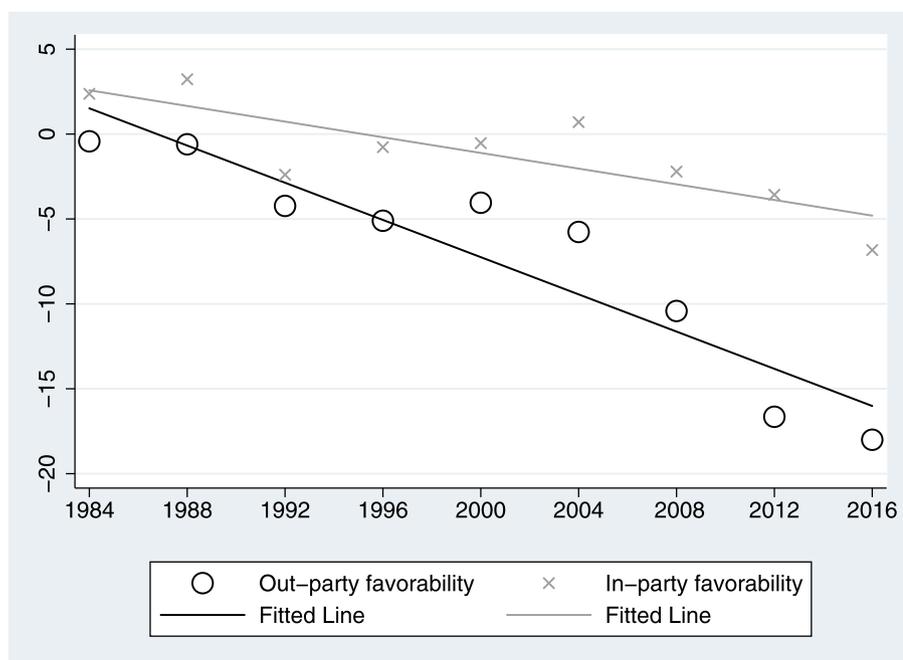
Fig. 1. Ideology, demographic, and partisanship-adjusted changes in out-party and in-party favorability ratings versus 1980, with linear trends. Each data point is an estimated year fixed effect (with omitted year 1980) from regression of dependent variable of in- or out-party thermometer ratings (0–100; 0 = "coldest" and 100 = "warmest"; the standard measure of feelings about the parties used in this literature) with fixed effect controls for 7-point party identity and 7-point ideology (to control for party sorting), age, education, income, race, gender, and region, for American National Election Studies cumulative file data from presidential election years from 1980 to 2016.

dislike focus on social group-based motivated beliefs—that partisans prefer (consciously or otherwise) to believe in their own group's strengths, and the out-group's weaknesses in various dimensions.[2]

In this paper, I empirically study another possible cause of partisan hostility: "unmotivated" cognitive bias. The distinction from motivated bias is significant, as misperceptions held due to unmotivated bias may be more likely to be mitigated by exposure to information contradicting one's beliefs. That is: I study whether partisan conflict is due partly to relatively unintentional, and thus perhaps relatively correctible, misunderstanding.

It might seem intuitive that partyism is due to cognitive biases beyond social identity-driven motivated reasoning. We all know from experience that when people disagree, this tends to make them unfairly judge and dislike one another, whether they come from different social groups or not. The 2016 US presidential campaign was characterized by much conflict and hostility within both parties (Sanders versus Clinton for the Democrats, and Trump versus the field for Republicans). Thus, group identity theory does not seem to offer the complete explanation for partisan animosity.

However, to the best of my knowledge, there is no existing work (in economics or beyond) that directly empirically studies the relationship between partisan hostility and unmotivated cognitive bias of any type. The broader literatures on hostility between social groups, and on interpersonal animosity more generally, also largely neglect such cognitive bias.[3] The lack of research on the relationship between bias and affective outcomes may be due to the historical distinction drawn between cognitive (thinking) and affective (feeling) psychological processes, which in recent years has been called into question.[4]

The specific bias that I study is overconfidence in precision of beliefs, a.k.a., overprecision. As I argue in more detail in Section 2, overprecision is an important bias that may affect partisan feelings for several reasons, both indirect and direct. Overprecision could lead to greater hostility indirectly by enhancing one's own political extremism and/or perceived social distance between the parties, in turn directly causing greater partyism for other reasons, such as motivated reasoning. Overprecision could also lead to greater out-party hostility directly by increasing confidence in information that reflects negatively on the out-party. Available information about the out-party could be skewed toward the negative for several reasons, which I discuss. Some of these reasons (extremism of political actions and slanted media) have become more prevalent over time, so they are also potentially relevant to the observed growth over time in partyism. (The empirical analysis is cross-sectional due to the nature of the data, as I discuss just below; I discuss the relationship between this analysis and the growth in partyism further at the end of Section 2.)

In Section 3, I describe the data, a 2011 module of the Cooperative Congressional Election Study (CCES), a representative sample with 1,000 respondents. This data set was developed and first used by Ortoleva and Snowberg (2015), henceforth referred to as OS.[5] In Section 3, I also discuss the method developed to calculate individual-level overprecision, and validation of this method. I describe the methods used for the empirical analysis in Section 4. Since motivated reasoning is the key alternative explanation for the outcome of partyism, I take extra steps to account for this confounding factor, including use of a control variable constructed to capture individual-level strength of motivated reasoning, and analysis of comparative statics that differ for motivated reasoning and overprecision. I also use pure confidence (in the accuracy of one's beliefs) as both a control variable and a placebo independent variable.

---

[2] See Haidt (2012) and Greene (2014) for books summarizing related psychology and neuroscience research. See also, e.g., Lelkes et al. (2017) and Rogowski and Sutherland (2016) for examples of political science papers that discuss the relevance of motivated reasoning to partisan hostility. Motivated beliefs and reasoning have recently begun to attract more attention in economics and were the subject of a symposium in a recent issue of the *Journal of Economic Perspectives* (see, e.g., Bénabou and Tirole, 2016).

[3] See, e.g., Tajfel (1978) and Rempel and Sutherland (2016) for work on social groups and on animosity more generally.

[4] For an example of the historical divide in psychology between cognition and affect, see Neisser (1971), and for an example of recent work critical of this divide, Pessoa (2008). Bénabou and Tirole (2016) refer to work "on the

---

(*footnote continued*)

interplay of emotions and information-processing … [as] sometimes referred to as the 'affective revolution' or 'second cognitive revolution.'" Another reason for the lack of literature on bias and out-group dislike may be lack of clarity regarding the distinction between motivated and unmotivated bias. I discuss this issue further in the next section, and attempt to provide clarification.

[5] The data set is now publicly available on the American Economic Review website.

I present results in Section 5. I find mixed evidence regarding the indirect effects of bias on partyism, and consistent support for direct effects. Overall, the results suggest that overprecision plays a substantial role in driving partyism: the point estimates imply a one standard increase deviation in a respondent's overprecision predicts as much as a 0.71 standard deviation decline in relative out-party favorability. The direct effect results are robust to inclusion of detailed controls for ideology, partisanship, perceived social distance between the parties, motivated reasoning, and pure confidence. I also show that overprecision effects do not increase with education (which they would likely do if they were merely proxying motivated reasoning effects; Kahan, 2015), and present evidence of a specific cognitive channel, belief in the Obama "birther" theory (the theory that he was not born in the US), by which overprecision increases dislike (indicating that overprecision is not just a proxy for "mindless" emotional effects). Both results, and an additional placebo analysis, support the interpretation of the main direct effect results representing causal effects. However, I acknowledge that the establishment of causality is not completely clear, and should be studied further in future work. In the final section of the paper, I provide brief concluding remarks, including discussion of potential related future work.

2. Related literature and hypotheses

2.1. Related political economics literature

Before proceeding to discussion of related work on overprecision, I briefly discuss two related literatures from the broader fields of public choice and political economy. One is the literature on biases in voter beliefs. See Caplan (2002) for documentation of such biases on policy issues, and Caplan (2001) for a defense of the individual rationality of biased beliefs due to low private costs and potentially substantial private benefits, and Levy (2014) for more recent related work. See, e.g. Congleton (2001), for related work on rational ignorance. The present paper is similar in that I also study biases in political beliefs—about the Birther issue in particular, and implicitly about character traits and actions more broadly that might drive favorability ratings (Stone, 2016). Just as voters may face low private costs of holding biased beliefs on issues such as the benefits of free trade, the private costs of holding biased beliefs underlying favorability ratings of the parties are likely also low. My paper differs from this literature in that I do not focus on private benefits as the source of the biased beliefs, often (implicitly) assumed to be driven by motivated reasoning, but instead, study more incidental, unmotivated cognitive bias as the cause.[6]

The economics literature on political polarization is also related. There are numerous strands of this literature. For example, there is a body of theory and empirical work on rational divergence of opinions (e.g., Andreoni and Mylovanov, 2012; Cheng and Hsiaw, 2017), and there are a number of empirical papers on polarization of politicians and/or voters on other cognitive mechanisms (e.g., de Leon and Rizzi, 2016; DeVault, 2013; Jha et al., 2018). Some papers study the relationship between non-standard information processing and polarization in various ways; for example, Esponda and Pouzo (2016) show that politicians may rationally polarize in response to behavioral voters, and Fryer et al. (2013) show that polarization can occur when individuals are Bayesian but process ambiguous signals in a simplified way. However, besides Ortoleva and Snowberg's work, there is perhaps surprisingly little empirical work linking cognitive bias to political polarization, and again no work, to my knowledge, on cognitive bias and affective polarization except (Stone, 2016).

2.2. Background on overprecision

Overconfidence is an especially ubiquitous and well-studied cognitive bias. It is now widely recognized in the behavioral economics literature that there are two major types of overconfidence: overprecision (over-estimation of the probability of holding correct beliefs or the precision of one's distribution of beliefs) and overoptimism (over-estimation of the probability of something one wishes to be true) (see, e.g., Grubb, 2015). Overprecision shrinks the variance of beliefs, while overoptimism shifts the mean.[7]

Overoptimism is more directly related to motivated reasoning, while overprecision is the most robust and widely observed type of overconfidence (Moore et al., 2015). Overprecision is perhaps even an over-arching bias, i.e., a root cause of other more narrow biases.[8] Empirically, overprecision indeed predicts other unmotivated biases (Hoppe and Kusterer, 2011; Noori, 2016).

2.3. Unmotivated versus motivated bias

The difference between overprecision and overoptimism reflects the general difference between unmotivated and motivated bias: the latter is due to believing what one wishes to be true, the former is due to cognitive heuristics and other "quirks." While overprecision can appear to reflect motivational forces, it does not have to, as we may hold overly-precise beliefs about a negative outcome. For example, overprecision can cause a student who lacks confidence in his ability to be excessively certain in his pessimism and unwilling to exert effort to learn his true potential. In general, while unmotivated and motivated bias may be empirically correlated, they are certainly not collinear, and they are conceptually quite distinct. See Moore et al. (2015) for further discussion and evidence of overprecision not being driven by motivational effects. Moore et al. (2016) do note a "conspicuous" lack of literature directly addressing the relationship between motivation and overprecision, so this issue is still unresolved, and I acknowledge that overprecision in this context could also be influenced by motivated forces.

It should also be noted that the term unmotivated bias is not original to this paper, but is not used often in prior literature.[9] But the term conveys a key distinction. An important comparative static that varies for the two types of bias is that overall cognitive ability and education should enhance the effects of motivated bias, as this ability makes us more likely to obtain cognitive motivational goals (see, e.g.,

---

[6] See Schnellenbach and Schubert (2015) for a useful survey of the broader behavioral political economy literature with an emphasis on the Virginia School of Public Choice in particular.

[7] Moore and Healy (2008) discuss three types of overconfidence: overprecision, overplacement, and overestimation. Numerous later papers follow this taxonomy. In Grubb's taxonomy, which I follow because it is simpler, the latter two types are subsumed by over-optimism.

[8] Informally, the general importance of overprecision is reflected in several famous, perhaps even clichéd quotes: "The more you know, the more you don't know," (Aristotle); "Not to be absolutely certain is, I think, one of the essential things in rationality," (Bertrand Russell); "It ain't what you don't know that gets you into trouble. It's what you know for sure that just ain't so," (Mark Twain). Tetlock and Gardner (2015) write that WYSIATI ("what you see is all there is"), which is conceptually similar to overprecision (both imply overconfidence in one's knowledge), is "the mother of all cognitive illusions." The intuition for this idea, and related ideas reflected in the other quotes, is that in order to hold a biased belief, one must be unaware of the full extent of the bias (otherwise one would account for the bias). And this unawareness of bias is essentially equivalent to over-estimation of the accuracy of one's beliefs.

[9] Laibson and Zeckhauser (1998) may have been the first paper in economics to refer to this term. Since then, there have been very few others in economics to use it; it is used more often, though still sporadically, in other fields. The terms hot and cold cognition are used somewhat more often, but are not equivalent to motivated and unmotivated; motivated reasoning does not require a "hot" or aroused mental state to occur. Bénabou and Tirole (2016) refer to the class of "unmotivated" biases as: "'mechanical failures' of inference due to bounded rationality or limited attention."

Kahan, 2015), while we would not expect overprecision effects to be enhanced this way.

2.4. Partisism hypotheses: indirect effects

I first discuss hypotheses of indirect effects of overprecision on partyism, then discuss direct effects. While I argue that unmotivated bias as a cause of partisan hostility has been relatively under-studied, it has not been completely neglected. There is a strand of the recent political science literature on affective polarization that studies this topic implicitly, focusing on biased perceptions of the out-party. In particular, several papers have shown that perceptions of differences between the parties are biased upward (Ahler, 2014; Rogowski and Sutherland, 2016; Levendusky and Malhotra, 2016), leading to greater social distance and thus, plausibly, greater out-party dislike (Enders and Armaly, 2018). Similar literature from psychology includes (Chambers and Melnyk, 2006) and Westfall et al. (2015).

Though not discussed by these papers, there are at least two specific mechanisms that could cause overprecision to increase perceived social distance from the out-party. The more straightforward mechanism is that overprecision could cause overconfidence in beliefs after observing signals of party differences that tend to be biased upwards. This bias could occur for various reasons, such as a focus by the media on issues that parties disagree on rather than what they have in common. A more subtle mechanism could occur independent of signals, via overprecision having asymmetric effects on the distribution of beliefs about the out-party's type if the tails of this distribution are thin or the support is bounded. Overprecision could then have a relatively strong effect on "ruling out" the out-party being a moderate type, leading to an exaggerated expectation of the differences between the parties.[10]

This literature and its relation to overprecision thus motivates the first hypothesis that I examine for an indirect effect of overprecision on out-party dislike:

Hypothesis 1.A. Partisans who are more overconfident in the precision of their beliefs perceive greater social distance from the out-party versus that of the in-party, which in turn increases relative dislike of the out-party, ceteris paribus.

H1.A conditions on holding an individual's own politics fixed. But it is also possible that bias drives dislike indirectly by playing a role in determining one's politics. This issue is the focus of OS's work: they study a formal model implying that overconfidence increases ideological extremism and strength of partisan identity, and present empirical results consistent with this prediction.[11] From other disciplines, Fernbach et al. (2013) provide similar experimental evidence, showing that individuals who hold extreme issue positions are subject to the illusion of understanding, and Rollwage et al. (2018) show that extremists demonstrate less meta-cognitive sensitivity. Enhanced extremism and strength of partisan identity increase social distance and out-party dislike (Mason, 2015; Abramowitz and Webster, 2015). Thus the second and third indirect effects that I hypothesize are:

Hypothesis 1.B. Partisans who are more overconfident in the precision of their beliefs are more politically extreme, which in turn increases relative out-party dislike, ceteris paribus.

Hypothesis 1.C. Partisans who are more overconfident in the precision of their beliefs have stronger partisan identities, which in turn increases relative out-party dislike, ceteris paribus.

2.5. Partisism hypotheses: direct effects

The primary focus of this paper is direct effects of overprecision on partyism (effects holding fixed social distance and group identity, i.e., holding fixed one's own ideology, partisanship, and perception of the out-party's ideology). While overprecision may not lead to hostility on its own, it plausibly exacerbates several other biases that likely enhance hostility, and the totality of such exacerbation effects could be substantial.[12] One such factor that overprecision could interact with to enhance dislike is negativity bias—the well-known tendency to focus on negative information (Rozin and Royzman, 2001). If partisans in general focus on negative aspects of the out-party, then those who hold more overconfident beliefs based on these negative aspects would likely be more hostile toward the out-party. See Iyengar and Krupenkin (2018) for discussion of negativity bias in the context of partisan affective polarization.[13]

The hypothesis of direct effects is also motivated partly by a companion paper (Stone, 2016). In that paper, a model is presented showing how two (unmotivated) cognitive biases can cause misinterpretation of political actions leading to increased dislike of the out-party, more so as the actions grow more extreme. I present a simplified version of this model in Appendix A.1 and show how these effects are exacerbated by overprecision. The basic intuition is that political actions reflect one's willingness to trade off private for social gains. More extreme actions reflect a stronger preference for private gains, i.e., being more self-serving, and less socially-minded. But the reasons for these extreme actions are more complex than observers realize.[14] Consequently, observers who tend to, in general, over-estimate their knowledge and understanding, tend to draw excessively negative character inferences from observed political actions.

In the companion paper, I discuss how the model applies to other bilateral relationships outside of politics, such as spouses, in which motivational and group identity factors should be much less relevant. The fact that unjustified dislike—dislike based on misguided beliefs—does seem to occur regularly in such contexts (this is the source of plot tension in many fictional works, and why there is a market for marriage counselors) implies that motivated reasoning and tribalistic thinking are not the only elements underlying this phenomenon. It is very plausible that unmotivated biases contribute to discord in relationships more generally, and that overprecision would exacerbate these effects.

Another way that overprecision could cause out-party dislike directly is via judgments based on a skewed distribution of available information. It is well known that people tend to prefer and be more likely to be exposed to politically congenial information via news media, and to avoid uncongenial information; see, e.g., Martin and Yurukoglu (2017) for recent evidence. Information from friends, family, and acquaintances could also be skewed in a similar way. While motivational forces could cause one to be insufficiently skeptical of biased information from such sources, overprecision would very plausibly have the same effect. Indeed, the paper that coined the term "affective polarization" (Iyengar et al., 2012) speculated that the increase in

---

[10] To see this idea more clearly, it helps to use a numerical example: suppose the interval of types is [0,1], and the true leftist type is 0.25 and rightist is 0.75. Suppose all partisans believe the mode for the out-party's type is the extreme value (1 for the rightist's type, and 0 for the leftist's). But the precision of beliefs varies; some partisans are certain that out-party members are extremists, while others put some (say, 1/3) weight on the correct value. Then partisans with certain (i.e., more overprecise) beliefs will expect the difference in types to be 0.75, while those partisans with more uncertain beliefs will have a lower expectation of this difference ((1/3)*0.5 + (2/3)*0.75 = 2/3).

[11] I refer the reader to OS's paper for the details of the mechanisms driving these results.

[12] To be clear, "direct" refers to the distinction from the indirect effects discussed in Section 2.4, and even the direct effects discussed in this subsection involve interactions between overprecision and other plausible factors or biases.

[13] Negativity bias may be less likely to affect beliefs about the in-party due to ego and/or identity concerns.

[14] Haidt's (2012) work on unintentional under-estimation of differences in moral "tastes" across parties reflects related ideas.

negative campaign advertising was a key cause of the phenomenon. That paper did not discuss the cognitive mechanisms underlying this effect but implied that media consumers would likely be naïve in interpreting this information.

It is also worth noting the substantial literature in psychology (Robinson et al., 1995; Graham et al., 2009; Chambers et al., 2013) showing that members of different parties think about moral values differently and misunderstand the moral thinking of out-party members. This type of misunderstanding could lead to an unduly pessimistic view of the out-party's morality and hence greater hostility between the parties, if failure to understand the out-party's moral values leads to inference of an absence of moral values. While I am not aware of any papers that directly discuss how unmotivated bias could play a role in forming these judgments, it seems plausible, and again overprecision would exacerbate negative effects of such inferences. Thus, there are several ways in which overprecision could directly increase partyism holding fixed other relevant factors such as ideological extremism. This hypothesis is stated formally as follows.

Hypothesis 2.A (*Direct effects of overprecision on relative out-party dislike*). Partisans who are more overconfident in the precision of their beliefs feel more relative out-party dislike, holding fixed ideology, partisan identity, and perceived social distance (and other relevant factors).Several important and testable auxiliary hypotheses follow from the above discussion. First, the discussion in Section 2.2 implied that motivated reasoning effects would increase with cognitive ability, while this would not be true for overprecision (direct) effects. Thus, testing the next hypothesis would shed light on whether observed overprecision effects might be driven by motivated reasoning forces.[15]

Hypothesis 2.B. The direct effects of overprecision on relative out-party dislike do not increase as cognitive ability increases.

Next, the discussion earlier in this subsection suggests that if overprecision direct effects are due to media exposure, then the overprecision effect should increase in the level of media exposure.

Hypothesis 2.C. The direct effects of overprecision on relative out-party dislike increase as media exposure increases.

Finally, one may be concerned that observed overprecision effects could reflect other unobserved psychological factors. For example, partisans who have higher levels of overprecision may be more likely to be emotional and simply have stronger feelings about out-groups, even though those feelings are not based on false beliefs influenced by overprecision per sé. On the other hand, if overprecision does affect partyism for the reasons discussed above, then overprecision should increase confidence in false beliefs that reflect negatively on the out-party. I state this final hypothesis as follows.

Hypothesis 2.D. Partisans who are more prone to overprecision are more likely to hold false beliefs that lead to greater relative out-party dislike.

Two additional remarks are in order before proceeding. First, as noted in Section 1, the hypotheses all refer to the cross-sectional relationship between overprecision and dislike, as the data set is cross-sectional, and an important question at hand is why is partyism so severe today. However, these hypotheses also relate to growth in dislike (affective polarization), as the mechanisms underlying the hypotheses (extremism in politician behavior, availability of anti-out-party information, and party sorting) have been exacerbated over time.[16]

Second, to be clear, I am not assuming that (relative) out-party dislike is due to bias; this dislike may be justified, at least for one party, and the empirical analysis allows for this (since the analysis focuses on relative dislike—disliking the other party more than one's own—it cannot be objectively justified for both parties).

3. Data and measurement

The data set is the Caltech module of the 2011 CCES (1,000 respondents), again, first used by OS. This survey module includes questions on four general knowledge topics and on four economic topics, and corresponding questions on confidence in knowledge. OS use modules from the 2010 CCES as well; I use only the 2011 data for three reasons: (1) the 2011 data set includes general knowledge questions, while the 2010 modules include just economic questions (general knowledge questions are more standard in studies of overprecision); (2) the 2011 data set has questions on twice as many knowledge topics as the 2010 data set (eight versus four); (3) only the 2011 CCES includes party thermometer questions (questions on the favorability of each party, rated on a 0–100 scale; this is the standard variable used to study affective polarization). Table 1 presents the eight knowledge questions from the 2011 CCES, corresponding correct answers, and summary statistics for respondent answers.[17]

Other papers on overconfidence typically measure this for individuals by comparing their actual accuracy of answers or coverage of confidence intervals to expected accuracy or coverage. OS discuss how overconfidence cannot be measured this way with the CCES data because it does not include a sufficient number of questions, as this would be infeasible for the survey, and because the CCES economic topic confidence questions are qualitative, which has the advantage of being more clearly understood by survey respondents than quantitative measures.

OS describe their method for measuring overprecision (which they refer to as overconfidence, but do note that the type of overconfidence that they study is overprecision) as follows (p. 511): "Confidence reflects both knowledge and overconfidence, so subtracting knowledge from confidence leaves overconfidence. To subtract knowledge, we deduct points from a respondent's reported confidence based on his or her accuracy, and thus knowledge, on the corresponding factual question. This is implemented conservatively: we regress confidence on an arbitrary, fourth-order polynomial of accuracy, and use the residual as a measure of overconfidence." Overprecision is measured this way separately for each knowledge topic, for each respondent, and then an aggregate measure is constructed for each respondent as the first principal component of the topic-level measures.[18]

(*footnote continued*)

forms of political behavior likely perceived as extreme such as government shut-downs, refusals to raise the federal debt ceiling, and blocked court appointments, including an extended delay for a Supreme Court Justice.

[17] All general knowledge questions were followed by "Even if you are not sure, please give us your best guess." All economic questions were preceded by statements of the mean, high and low values of the variable since World War II. Both types of questions include qualitative measures of confidence, coded as categorical variables ranging from 1 to 6, and the general knowledge questions also include quantitative measures of confidence ("what do you think is the percent chance that your best guess, entered above, is within _ of the actual answer?", with the range referred to varying across the questions). For the general knowledge questions, to use all available confidence information, I use the average of the qualitative and percentage questions (each normalized to a scale of 0–1); results are similar when just the qualitative questions are used.

[18] Moore and Dev (2017) note that while it is intuitively and theoretically plausible that overconfidence would be a stable individual-level characteristic (as assumed by OS), prior literature has failed to support this idea. However, for overprecision in particular there is at least one study Moore and Swift (2010) that "does seem to demonstrate more intra-individual consistency."

---

[15] This issue is much less relevant to indirect overprecision effects since we would not expect motivated reasoning to lead to evidence consistent with Hypothesis 1.B or 1.C. The threat is more plausible for Hypothesis 1.A, but it turns out that this hypothesis is not supported so we do not need to be concerned about alternative explanations for results supporting the hypothesis.

[16] Regarding political extremism, in addition to the well-documented polarization of politician actions and language, there has been an uptick in other

Table 1
Knowledge questions and summary statistics for responses.

Question	Correct answer	Mean answer	Mean confidence	5th percentile	95th percentile
General knowledge questions					
In what year was the telephone invented?	1896	1876.6	0.49	1825	1930
What is the population of Spain, in millions?	46	60.0	0.31	2	200
In what year was the playwright William Shakespeare born?	1564	1642.7	0.34	1466	1897
What percent of the US population lives in California?	12.1	17.9	0.33	4	45
Economic questions					
What is your best guess about the no. of jobs gained (in millions) in the last year?	0.4	−5.5	0.22	−34	1.8
" best guess about the inflation rate in the United States today?	2.9	9.9	0.30	−1	45
" best guess about the no. of jobs that will be gained or lost over the next year?	3.1	−0.9	0.26	−5.2	2.8
What do you expect the inflation rate to be a year from now?	2.3	6.8	0.29	−2	25

Confidence normalized to 0–1 scale as described in text. Mean answers are calculated from raw data (without top-coding). 998–1000 observations for all values.

While this is a very reasonable procedure, it yields estimates of overprecision that are very highly correlated with pure confidence, so analysis using this measure could largely pick up pure confidence effects and not those of overconfidence. Their measure of overprecision has a 0.96 correlation with an analogous measure of confidence. OS note in their paper that all of their theoretical and empirical results hold when overconfidence is replaced with confidence, implying that separating confidence and overconfidence is not as crucial for their purposes as it is for mine in this paper.

I propose an alternative measure of overprecision that more directly accounts for the difference between confidence and knowledge (i.e., the difference between perceived and actual precision of beliefs). To do this, I first normalize confidence for respondent i on topic t to range from 0 to 1, denoted by C_{it}. This is equal to the fraction of possible confidence the respondent could have for the topic. For each topic, knowledge is bounded above (the correct answer) and bounded below (following OS, I top-code errors; more on this below). A natural definition of knowledge (for respondent i on topic t) then with the same support and units as C_{it} is $K_{it} = \frac{e_t^* - e_{it}}{e_t^*}$, in which e_{it} denotes the absolute value of i's error for topic t, and e_t^* denotes the maximum error for t.

Given that K_{it} and C_{it} have the same scale and (lack of) units, I can now apply OS's conceptual definition of overprecision directly. Letting OP_{it} denote overprecision for i on topic t:

$$OP_{it} = C_{it} - K_{it}. \tag{1}$$

This definition, while straightforward, makes a linearity assumption on the proper calibration of confidence: the marginal effects of (fractional) changes in both confidence (perceived "precision") and knowledge (actual "precision") are constant.[19] Still, in addition to mapping directly to OS's conceptual definition of overconfidence, this definition of OP has several attractive properties beyond OS's definition:

1. OP is at its upper bound (for a given respondent-topic) when confidence is at its upper bound and knowledge at its lower bound (for that respondent-topic) ($OP_{it} = 1$ if $K_{it} = 0$ and $C_{it} = 1$).
2. OP is at its lower bound when confidence is at its lower bound and knowledge at its upper bound ($OP_{it} = -1$ if $K_{it} = 1$ and $C_{it} = 0$).
3. $OP_{it} = 0$ when knowledge and confidence are "equal" ($K_{it} = C_{it}$).

For this measure of OP, I follow OS in using the first principal component of the respondent-topic-level estimates of overconfidence to estimate respondent-level overconfidence (with all eight topics for most of the analysis). As mentioned above, OS top-code the absolute value of error for each question at the level approximately equal to the errors at the 5th and 95th percentiles for each question. The measure of OP that I

use in the paper uses the same top-coding as OS.[20] OP has a correlation of 0.62 with pure confidence, substantially lower than OS's correlation. In the appendix I discuss and present evidence of further validation of OP.

Fig. 2 presents kernel densities of favorability ratings for each of the parties for respondents who are members of each party. The in-party and out-party ratings are similar across the two parties, but Democrats have a higher modal rating of the in-party. Fig. 3 presents a smoothed plot of the relationship between OP and the difference in party favorability ratings, F_D (favorability difference), equal to $F_O - F_I$ (F_O, F_I = out (in)-party favorability), providing a preview of the more detailed econometric results. Note that, to the extent that OP is measured with error, this would attenuate results regarding its effects on the various outcome variables, strengthening the validity of such non-null results (and reducing the validity of null results).

4. Empirical methods

The methods that I use to estimate OP effects are mostly straightforward: various dependent variables are regressed on OP, the set of economic and demographic control variables used by OS, and myriad alternative controls and interactions. The standard controls used in all of the analyses are fixed effects for income groups, education groups, union member (current or former), union household, home ownership, stock ownership, state, fourth order-polynomials in both age and a political knowledge/interest index.[21] I also generally include as many ideology and partisanship control variables as possible to attempt to conservatively account for omitted factors that could bias OP coefficients.

---

[19] As OS discuss, this proper, or correct, calibration is unobservable with this data, and thus estimating overconfidence requires us to make some type of assumption.

[20] However, as Table 1 shows, the absolute values of errors for responses below and above the correct answer can vary substantially. The table also indicates that for some questions, the sign of the error is relevant to the interpretation. For example, when estimating the population of Spain, an answer of less than 5 million is closer to correct (44 million), in levels, than an answer of 100 million, but the latter could be seen as more reasonable. To account for these issues, I construct an additional measure of OP in which I use the actual (rounded) 5th and 95th percentiles for top-coding of errors, and use one value of e^* for estimates greater than the true value (e at the 95th percentile answer) and one for estimates less than the true value (e at the 5th percentile) for each topic, and find that using this measure instead of the original one has very minor effects on all of the results.

[21] My sample size is smaller than OS's, so to preserve degrees of freedom I replace age fixed effects (which OS use in some of their analyses) with a fourth order polynomial. The political knowledge/interest index (henceforth just political knowledge) is constructed as follows. The CCES includes eight questions on knowledge of current political representatives: the party of ["that has majority of seats of" in 2010 CCES]: (1) House; (2) Senate; (3) state Senate; (4) state House; (5) governor; (6) US Senator 1; (7) US Senator 2; (8) US House Representative. I create dummy variables for the answer "not sure" for each of these, and then use the first principal component of these eight dummies as the index.
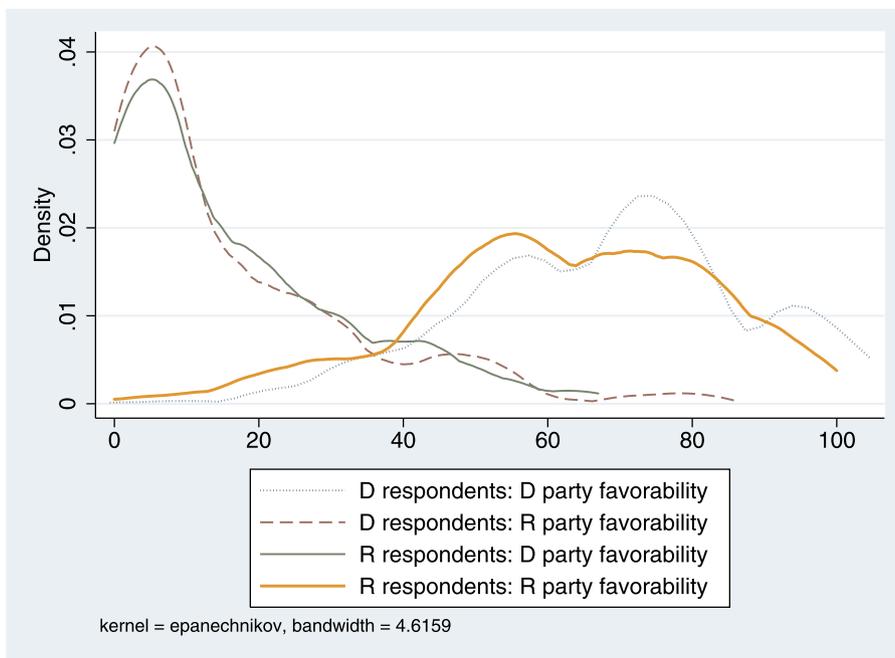
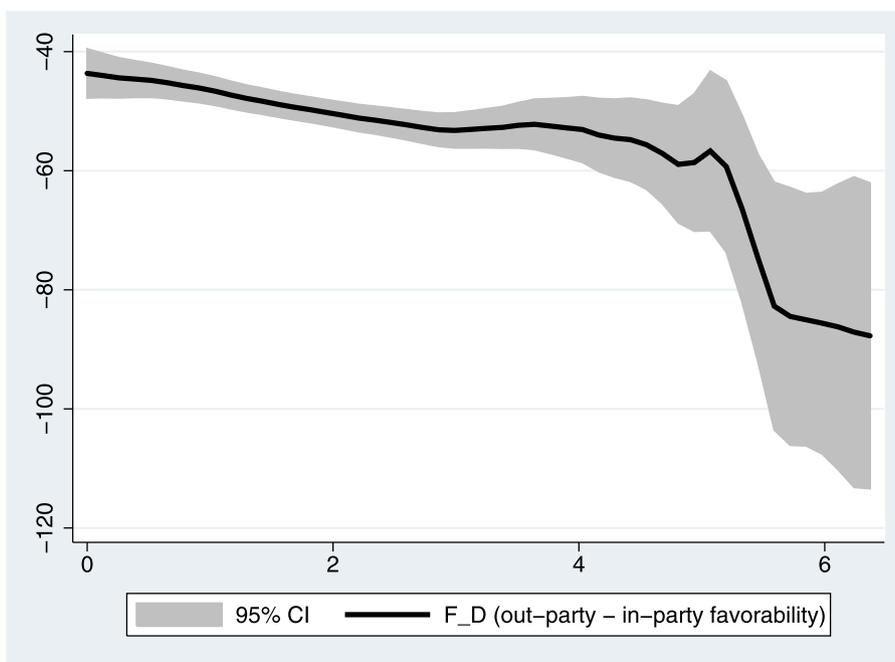**Fig. 2.** Kernel densities of party favorability ratings.

Fig. 3. Smooth local polynomial plot of *OP* (*x*-axis) versus F_D (y-axis).

To test Hypothesis 1.A, I use "Distance" as the dependent variable, which is perceived ideological distance from the out-party relative to that of the in-party.[22] For these analyses, I include as controls fixed effects for "Partisanship" (equal to 0, 1, or 2 for "lean" toward party but not member, "not very strong" member, "strong" member, respectively), since partisan identity is not directly incorporated in the dependent variable.

To test Hypotheses 1.B and 1.C, I use self-reported ideological extremism ("Extremism," measured on a 0–1 scale) and Partisanship as dependent variables.[23] For the former, I exclude ideology controls and Distance as it is a function of self-reported ideology, and include Partisanship fixed effects. For the latter, I exclude party strength on the right-hand-side (though do include a dummy for party, "Democrat"), but include Distance as a control variable, and ideology controls (non-linear functions of both self-reported ideology and issue-preference ideology estimated by Tausanovitch and Warshaw (2013)).

To test Hypotheses 2.A-2.D (and the latter parts of Hypotheses 1.A-

---

[22] This is the absolute value of the difference between own ideology and perception of the out-party's ideology, minus the absolute value of this difference with respect to the in-party, with all ideologies measured on seven-point scales, and differences coded as zeroes for respondent's who answer "don't know" about the party's ideology.

[23] Extremism is the average of the normalized deviation of the mid-point from self-reported ideology on 5- and 7-point scales.

1.C regarding effects on out-party dislike), I use the dependent variable of the difference in party favorabilities, F_D, as dislike is by nature comparative, and include Partisanship and Distance as regressors, and both types of ideology as additional controls. To test Hypotheses 2.B and 2.C, I interact *OP* with measures of media consumption and education, as described further in the next section. To test Hypothesis 2.D, I use a dependent variable reflecting belief in the Obama "Birther" theory, described further in the next section.

I use Tobit models for the analysis of F_D as the dependent variable because it is censored (from −100 to 0; I drop observations with $F_D > 0$ as these likely reflect survey response errors) and OLS for all other models for computational simplicity and so estimated coefficients can be interpreted as average marginal effects. I use robust standard errors and use the survey weights whenever possible.[24]

For the analysis of each hypothesis, I examine pure confidence, *C*, as both a placebo independent variable (dropping *OP*), and as a control variable with *OP* (results are similar when I replace *C* with a measure of pure knowledge, and I cannot use both *C* and knowledge as controls because of their collinearity with *OP*). For each hypothesis, I also examine three additional specifications. The first includes an interaction of *OP* with Democrat. As mentioned in Section 2, there may be asymmetries in behavior of the parties that justifies dislike. There may also be asymmetries in cognitive styles across members of the parties (see, e.g., Jost et al., 2003).

The second additional specification includes *MR* as a control variable, which is an index of a respondent's propensity to engage in motivated reasoning/response. I construct *MR* as follows. Previous literature shows that survey respondents answer questions about economic conditions more favorably when a co-partisan is in office as president and less favorably when the president is from the out-party (Bullock, 2009), and this is also the case for economic questions in the CCES data set.[25] This phenomenon is relatively likely to be driven by motivated reasoning (respondents truly believe co-partisans perform relatively well because they are motivated to believe this) and/or because of expressive, or motivated, responding. Thus, respondent answers to these questions may reflect a general tendency toward motivated behavior, and so I construct a respondent-topic level measure of motivated reasoning/responding as:

$$MR_{it} = \begin{cases} \left(\dfrac{e_{it}}{e_t^*}\right) \times C_{it} & \text{if } e_{it} \text{ is "motivated"}; \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

e_{it} is defined as "motivated" if it is in the direction favorable politically to the respondent, given her party (good economic conditions for Democrats, bad conditions for Republicans). This variable is thus increasing in both the favorability of the stated answer (difference between that answer and truth) and respondent's confidence in the answer, with a positive interaction. It is fairly clear that motivated reasoning would cause Republicans to state lower job growth numbers and higher inflation, and cause Democrats to state larger job growth

numbers. It is less clear what ideal inflation would be for Democrats, and thus to be conservative I use only the two job questions for MR_{it}. There are two other survey questions that are useful for this purpose, on the respondent's perception of recent performance of the national economy (1–5 scale) and the unemployment rate, which I adjust based on the respondent's party so that a higher number represents a more "motivated" answer (and arbitrarily top-code unemployment at 20%). I use the first principal component across these four measures per respondent to estimate a respondent-level measure, MR_i; for the analysis that includes *MR* I use a version of *OP* based only on general knowledge questions, since *MR* incorporates the economic questions. In the Appendix, I provide evidence of validity of *MR*, and the distinction between *OP* and *MR* is demonstrated further in the birther theory analysis.[26] *MR* is certainly not a perfect measure of motivated reasoning, but it is still useful to examine if *OP* estimates are robust to its inclusion.

The third alternative specification is one in which I account for measurement error in *OP* using instrumental variable (IV) estimation. Ortoleva and Snowberg (2015) construct two measures of *OP*, one using just the general knowledge questions and one using just economic questions, and account for measurement error by instrumenting for each measure using the other (estimating two separate IV models). Given this set-up (two noisy measures of a variable, with neither being a priori superior to the other), Gillen et al. (2018) suggest that an alternative procedure is more efficient, which I follow.[27]

5. Results

I standardize variables whose scales are difficult to interpret, *OP* (overprecision), *MR* (motivated reasoning), *C* (confidence); each unit of these variables now equals one standard deviation of the original variable. Table 2 presents results for the analysis of the first set of hypotheses. Panel A presents analysis of Hypothesis 1.A. Three of the five *OP* estimates are insignificant; the *OP* coefficient is significant at 10% in the baseline model and at 5% when *C* is included, and negative in both cases, which is inconsistent withh the hypothesis. These results provide some evidence that, conditional on this large set of controls, extremism *declines* as *OP* increases, i.e., that being more extreme is "more rational" (having more accurately calibrated confidence in one's knowledge). An explanation for this phenomenon could be that extremism entails private, social identity-related benefits, consistent with claims discussed by Kahan et al. (2012, 2013). However, these results are non-robust and far from conclusive; none of the estimates are significant at 1% and in three of the five specifications the estimates are not even significant at 10%.

The results for Hypothesis 1.B are quite similar to those for Hypothesis 1.A, which is not surprising given that both dependent variables (Extremism and Distance) are direct functions of the respondent's self-reported ideology. Estimated *OP* effects are substantially stronger and more significant for Hypothesis 1.C, regarding the Partisanship dependent variable. The results support the hypothesis: that *OP* increases strength of partisanship. The baseline model indicates that a one standard deviation increase in *OP* predicts an increase in expected level of party strength of 0.14; the IV estimate is 0.438. The estimate for the placebo variable, *C*, is

---

[24] I do not cluster standard errors (OS clustered standard errors by age; for my analysis, this type of clustering yields smaller standard errors but is not clearly justified, so it is more conservative to use non-clustered robust standard errors). For some methods it is more convenient (or only possible) to treat the weights as analytic weights; for analysis in which both weight options are available, results are quite similar either way.

[25] Democrat respondents consistently described and predicted more favorable economic conditions than Republicans; the mean (without top-coding) responses to the question on job changes in the last year were − 4.2 (million) for Democrats and − 6.1 for Republicans; on predicted job changes, 0.89 and − 2.44, respectively; on inflation last year, 8.31% and 10.19% (resp.) and on predicted inflation, 5.34% and 8.05%. The signs of each of the differences are the same (though magnitudes smaller) when restricting the sample to the most highly educated group.

[26] I (arbitrarily) top-code unemployment guesses at 20%. I examined several variants on this definition of *MR*, including using all four economic questions, using negative values for responses unfavorable to the co-party, using just the economic knowledge questions, and using just the survey questions on economic performance without confidence measures, and different types of top-coding, and obtained generally similar results, though validation results are much weaker without the economic knowledge questions.

[27] This procedure is to combine the measures by duplicating and stacking the data set, and creating two variables, one equal to the first measure for the first half of observations and the second measure for the second half, and the second variable the reverse (second measure for first half, first measure for second), and then using one as IV for the other (and clustering standard errors by the initial observation identifier).

Table 2
Analysis of Hypotheses 1.A, 1.B, and 1.C.

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Hypothesis 1.A (Dependent variable = Distance)						
OP	−0.068* (0.039)		−0.122** (0.054)	−0.100 (0.061)	−0.053 (0.036)	−0.239 (0.146)
C		−0.003 (0.018)	0.036 (0.025)			
OP × Dem				0.052 (0.078)		
MR					−0.030 (0.045)	
Panel B: Hypothesis 1.B (Dependent variable = Extremism)						
OP	−0.025* (0.015)		−0.046** (0.020)	−0.016 (0.023)	0.001 (0.014)	−0.069 (0.051)
C		−0.000 (0.007)	0.015* (0.009)			
OP × Dem				−0.014 (0.028)		
MR					0.012 (0.014)	
Panel C: Hypothesis 1.C (Dependent variable = Partisanship)						
OP	0.140*** (0.040)		0.178*** (0.053)	0.069 (0.062)	0.095** (0.037)	0.438*** (0.168)
Distance	0.104** (0.051)	0.093* (0.052)	0.107** (0.051)	0.102** (0.051)	0.112** (0.052)	0.134** (0.053)
C		0.032 (0.020)	−0.026 (0.026)			
OP × Dem				0.115 (0.082)		
MR					0.022 (0.043)	
Model	OLS	OLS	OLS	OLS	OLS	IV
N	739	739	739	739	736	1478

All models estimated using survey weights, with economic, demographic, and political knowledge controls, and robust standard errors (clustered by respondent for IV model, which uses duplicated data set). Panels A and B models include partisanship fixed effects, and panels A and C models include self-reported ideology and issue-based ideology controls. All control variables described in text. *, **, *** denote 10%, 5%, 1% significance.

near zero and insignificant. Including MR reduces the OP effect to 0.095, but it remains significant at 5%. The interaction coefficient is insignificant, failing to support effects varying by party.[28]

Table 3 presents results for analysis of Hypothesis 2.A. Results for OP are all significant at 5% or 1% and supportive of the hypothesis. The baseline model implies a one standard deviation increase in OP predicts about a five point decline in relative out-party favorability, and a 16 point decline for the IV model. The placebo effect (C) is much smaller, and disappears when included with OP. The results also indicate that a one unit increase in Partisanship predicts an approximately 7.5 point decrease in out-party dislike (5 points in IV model). Taking the products of these estimates with the baseline and IV OP coefficients from Panel C of Table 2, the indirect effects of a standard deviation increase in OP on dislike range from 1.05 to 2.2 points. The combined (direct plus indirect) IV effect of 18.2 favorability points (per OP standard deviation) is 71% of one standard deviation in F_D. Again, there is no significant variation in effects across parties.[29]

In Table 4, I report results for interactions of four dummy variables with OP to examine heterogeneity in OP's direct effects on F_D, and to address Hypotheses 2.B and 2.C. The dummies are indicators for being above the median in age, education, income, and media exposure.[30] Only the age interaction term is significant; it is positive and about the same magnitude as the OP coefficient, indicating that OP has a near-zero overall effect for respondents whose age is above the median. The (unreported) non-interacted high age dummy coefficient is approximately −15 (over twice the OP coefficient, and just over the IV OP coefficient from Table 4). These results together imply that over-precision drives dislike for younger people only, and that transitioning to the older age group has the same effect on dislike as moving about half-way up the OP distribution. This age effect is consistent with Boxell et al. (2017). The other interactions, which more directly address the hypotheses, are insignificant.

These results provide fairly strong support for Hypothesis 2.C (that overprecision effects do not increase with cognitive ability, proxied crudely by education), but fail to support Hypothesis 2.B (that overprecision effects increase with media exposure). The media interaction may be attenuated due to counter-acting effects (e.g., media exposure could be positively correlated with an aspect of cognition that reduces F_D, nullifying the positive effect from exposure to more out-party information). Moreover, the media variable is admittedly a quite noisy measure of media exposure. It is also possible that quantity of media exposure does not matter as much as type of media exposure (which I do not have data on)—or that overprecision direct effects occur mainly through mechanisms other than media.

Table 3 suggests that OP directly causes dislike, but does not show specifically how OP does this, i.e., which beliefs, if any, that OP might affect that in turn drive dislike. To address this issue (Hypothesis 2.D), I conduct an auxiliary analysis in which I examine OP effects on a variable capturing belief in the Obama "birther" theory (the theory that he was not born in the US, Birther, equal to 1, 2, ..., 5 (strongest belief in the theory)). This variable measures beliefs on a factual, though perhaps ambiguous, issue. The Caltech module includes questions on several other factual issues but I restrict attention to the birther question because responses vary by party most substantially (by a wide margin), suggesting its importance for partisan affect. I conduct this analysis separately for Republicans and Democrats as multiple coefficients are expected to vary by party. I examine the effects of OP and MR on Birther, and also the effects of Birther on F_D.

Table 5 presents these results. They imply that higher OP predicts belief in the birther theory for Republicans, and this effect is robust to including the MR control. The results also imply that birther belief contributes to hostility, as Birther predicts lower F_D for Republicans. Results also show that MR drives disbelief in the birther theory for Democrats; the sign of the MR effect for Republicans is as expected but it is not quite significant. OP has a marginally significant positive effect on Birther for Democrats; this might seem surprising, but is reasonable given the prevalence of the theory, and also supports the difference between measured unmotivated and motivated bias. Democrats who were simply more prone to believe what they hear might have been more likely to believe the theory. These results support OP having effects on factual issues that could affect partisan feelings, and support OP effects being distinct from MR effects. The results also hint at differences in various effects across the parties, which would require a larger sample to carefully explore.

---

[28] The simplest way to reconcile the conflicting results here—the results supporting the opposite of Hypotheses 1.A and 1.B, and results supporting Hypothesis 1.C—is to simply discount the former, since they are less robust. However, the two types of results could both be valid, since the dependent variables are distinct, but I refrain from further speculation here.

[29] In results reported in a referee appendix, I show that results are also robust to controlling for religion and religiosity, and find evidence that religion being "very important" increases strength of partisan identity, suggestive (significant at just 10% in a limited number of specifications) evidence that this type of religiosity increases relative out-party favorability, and that for Republicans being a member of a non-Christian religion decreases relative out-party favorability.

[30] I use OS's measure of media exposure, which is the first principal component of four variables, each assessing use of a different type of media in the past 24 h (blog, TV, radio, newspaper). The sample is smaller for the model with the income interaction because I drop respondents who answered "prefer not to say" (which was not an option for other questions).

Table 3
Analysis of Hypothesis 2.A (Dependent Variable = F_D).

	OLS	OLS	OLS	OLS	OLS	IV
OP	−5.037***		−4.879***	−4.073**	−2.919**	−16.035**
	(1.241)		(1.724)	(1.700)	(1.140)	(6.923)
Partisanship	−7.524***	−8.118***	−7.533***	−7.468***	−7.788***	−5.058**
	(1.394)	(1.412)	(1.395)	(1.387)	(1.400)	(2.025)
Distance	−9.612***	−9.157***	−9.598***	−9.597***	−9.550***	−10.380***
	(1.648)	(1.677)	(1.656)	(1.652)	(1.660)	(1.814)
C		−1.667***	−0.106			
		(0.589)	(0.801)			
OP × Dem				−1.581		
				(2.469)		
MR					0.348	
					(1.205)	
N	739	739	739	739	736	1478

All models estimated using survey weights, with economic, demographic, political knowledge, self-reported ideology, issue-based ideology controls, and robust standard errors (clustered by respondent for IV model, which uses duplicated data set). *, **, *** denote 10%, 5%, 1% significance.

Table 4
Analysis of Hypotheses 2.B and 2.C (Dependent Variable = F_D).

	(1)	(2)	(3)	(4)
OP	−6.835***	−5.274***	−4.478**	−4.697**
	(1.532)	(1.505)	(1.750)	(2.092)
$OP \times I(> \text{age median})$	6.006***			
	(2.034)			
$OP \times I(> \text{education median})$		2.268		
		(2.294)		
$OP \times I(> \text{income median})$			−0.825	
			(2.241)	
$OP \times I(> \text{media median})$				0.192
				(2.475)
N	739	739	651	739

All models estimated using survey weights, with non-interacted respective indicator variable, Partisanship, economic, demographic, political knowledge, self-reported ideology, issue-based ideology controls, and robust standard errors. *, **, *** denote 10%, 5%, 1% significance.

Table 5
Analysis of Hypothesis 2.D.

	Republicans				Democrats			
	Dep var: Birther		Dep var: F_D		Dep var: Birther		Dep var: F_D	
OP	0.381***	0.252**	−4.014**	−2.852*	0.080	0.119*	−3.482*	−3.665**
	(0.098)	(0.100)	(1.673)	(1.703)	(0.077)	(0.068)	(1.773)	(1.763)
Partisansh.	0.156	0.195	−5.899***	−5.424**	0.137*	0.147*	−8.974***	−9.289***
	(0.118)	(0.119)	(2.133)	(2.140)	(0.080)	(0.078)	(1.908)	(1.883)
Distance	−0.122	−0.186	−13.567***	−13.938***	−0.281***	−0.280***	−9.809***	−9.165***
	(0.130)	(0.143)	(2.690)	(2.557)	(0.091)	(0.085)	(2.237)	(2.260)
MR		0.133				−0.302***		
		(0.087)				(0.098)		
Birther				−3.048**				2.295
				(1.321)				(1.977)
N	338	337	338	338	401	399	401	401

All models estimated using OLS with survey weights and economic, demographic, political knowledge, self-reported ideology, issue-based ideology, party, Partisanship controls (described in text), and robust standard errors. *, **, *** denote 10%, 5%, 1% significance.

6. Concluding remarks

Dislike causes bias: people tend to be biased against things they dislike because they dislike them, political parties included. This paper examines and provides evidence in support of causality in the other direction: bias being the source of dislike. I find that the bias of over-precision does not predict ideological extremism or a measure of partisan social distance, but that this bias does predict: (1) stronger partisan identity (which in turn predicts out-party dislike), (2) a direct increase in out-party dislike, holding fixed ideology and partisanship, and (3) beliefs about a factual issue, which in turn predict dislike. The results augment the better established literature arguing that dislike is caused in part by beliefs that may be biased by motivated forces. Dislike is likely also caused by beliefs that are biased more incidentally.

Again, these results do not at all rule out other explanations for partyism, which almost surely has many causes. Moreover, even if one questions the distinction that I draw between motivated and un-motivated bias (in general, or regarding the interpretation of the particular estimates reported in this paper), and prefers to interpret this paper's results as reflecting general cognitive bias effects, these results are, I would argue, still an important contribution to the literature on this phenomenon. The empirical relationship established between bias and partisan hostility clarifies the notion that competing partisans normatively should not dislike each other as much as they do—that the

dislike is driven in part by perceptions that are objectively false, and dislike could be abated if each side could understand the other (and itself) better. Public awareness of partyism being caused by cognitive errors could even help mitigate partyism, by increasing both intrinsic and social incentives to do so.

There are several aspects of this paper that could be improved in future work. First, there is the lack of experimental variation. One way to make progress in this regard would be to measure respondent-level overprecision, then present respondents with suggestive but noisy information indicating the out-party (in-party) is a bad (good) actor, and measure effects of this information on affect. If respondents who are generally more overconfident in beliefs demonstrate stronger reactions to the noisy information, this would support Hypothesis 2.A. Another possibility would be to present respondents with information on their own overprecision and see if this moderates overprecision and consequently partisan hostility. The second issue that could be improved upon is the measure of overprecision—although it is based on questions standard in this literature, these questions are not tailored to the context of politics or inter-personal beliefs. It is possible that modifying these questions could improve the overprecision measure for this context. Perhaps some people are particularly overconfident in their judgments of other people, rather than knowledge in general, and if so, we would expect this interpersonal overconfidence to be most relevant to partyism. Similarly, it may be possible to obtain a better measure of motivated reasoning as a distinct factor. Third, the analysis of heterogeneity, by party and other factors, was limited, mainly due to sample size; increasing this is of course costly, but feasible and potentially worthwhile.

Appendix A. Appendix to *"Unmotivated Bias" and Partisan Hostility: Empirical Evidence*

A1. Model

To show how overprecision can increase partyism, consider the following simplified version of the model and analysis of Stone (2016). That paper does not directly show that overprecision causes partyism, but implies that overprecision in conjunction with two other biases would increase a general tendency toward negative inferences about the opposition party's character.

There are two players, L (left) and R (right). Each player can be thought of as a politician representing her/his party's median voter. Each player i chooses a policy, $x_i \in \mathbb{R}$, simultaneously. There are two versions of the model, one in which policies are chosen unilaterally, and one in which they are chosen jointly. In the unilateral case, payoffs are:

$$u_L(x_L) = -x_L - (1/s_L)(x_L - \tau_L)^2;$$
$$u_R(x_R) = x_R - (1/s_R)(x_R - \tau_R)^2.$$

Each payoff consists of two terms. The first is a direct private payoff. For L, this decreases in x_L, and for R, increases in x_R. The second term is a subjective social welfare loss function, which is minimized when $x_i = \tau_i$, which is i's taste for the level of x_i that is best for the social good. Thus, each player i faces a trade-off between private benefits (large x_i for $i = R$ and small x_i for $i = L$) and social benefits (matching x_i to τ_i). The policy choice can be interpreted broadly as an action representing "leftist" or "rightist" government; leftists, on average, privately benefit from "leftist" ("bigger") government due to a larger safety net, bureaucracy, etc, and vice versa for rightists (on average benefitting from lower taxes, less regulation, etc).

The parameter $s_i \in (0, \infty)$ inversely measures i's social-mindedness, or can be interpreted to directly measure i's "selfishness." As s_i grows, i puts more weight on i's private payoff in choosing x_i. It is assumed that hostility toward the other player increases as the posterior expectation of that player's s parameter increases.

Straightforward payoff maximization implies the optimal policy choices are:

$$x_L^* = -s_L + \tau_L,$$
$$x_R^* = s_R + \tau_R. \tag{A.1}$$

The extremism of i's action is increasing in s_i, and in the degree to which τ_i is aligned with i's private interests, for $i = L$ and $i = R$.

Each player observes the policy choices plus noise: $\hat{x}_i = x_i + \epsilon^i$, with $\epsilon^i \sim N(0, \sigma_\epsilon^2)$. The true distribution of s_i, $i \in \{L, R\}$, is truncated Normal, with lower bound zero, $\mu_s > 0$, and σ_s^2. Due to the truncation, $E(s_i)$ is equal to μ_s plus a normalizing factor; however, it simplifies the analysis, and does not affect the nature of the results, to treat this distribution as (non-truncated) $N(\mu_s, \sigma_s^2)$. For each i, $\tau_{-i} \sim N(\mu_{\tau_{-i}}, \sigma_\tau^2)$.

It is assumed that $\mu_{\tau_L} < \mu_{\tau_R}$, that is tastes are correlated with private interests. This could be due to either selection (the L party is more likely to select members who truly believe larger government is socially optimal) and/or motivated reasoning (privately benefitting from larger government may make some people truly believe that larger government is also socially optimal).

As noted above, Stone (2016) studies two other biases, these are the false consensus bias and limited strategic thinking. The false consensus bias is the bias toward over-estimating consensus in tastes. In this setting, it would imply that $\mu_{\tau_L^R} = \mu_{\tau_R} - b$ and $\mu_{\tau_R^L} = \mu_{\tau_L} + b$ for some $b > 0$.

It is then straightforward to show that the false consensus bias causes each player i to have an upwardly-biased estimate of s_{-i} after observing $\hat{x}_{-i}$, in expectation. Moreover, this bias is increasing in the perceived precision of priors on τ_{-i} and in the precision of the observed signal. This result is formalized as follows, with σ_p^i denoting i's perception of the value of the σ for parameter p. (The proof follows directly from analysis and results in Stone (2016).)

Proposition A.1. *If policies are chosen unilaterally, $E_i(s_{-i}|\hat{x}_{-i}) > s_{-i}$ in expectation, and $E_i(s_{-i}|\hat{x}_{-i}) - s_{-i}$ is increasing in $1/\sigma_\tau^i$ and in $1/\sigma_\epsilon^i$, and decreasing in $1/\sigma_s^i$, for $i \in \{L, R\}$.*

Note that while overprecision about tastes and new information exacerbate growth in negative inferences, overprecision in the prior on s_{-i} reduces this growth. However, if priors on s_{-i} are overly precise and the prior mean is biased upward due to a prior bias against the "out-group" as is also plausible (as also discussed by Stone (2016)), then this would be another way that overprecision exacerbates hostility since overprecision would reduce updating toward the true, lower (on average) value of s_{-i}.

For the second bias, limited strategic thinking, the strategic version of the model (payoffs) must be used, which is:

$$u_L(x_L, x_R) = -0.5(x_L + x_R) - (1/s_L)(0.5(x_L + x_R))^2;$$
$$u_R(x_L, x_R) = 0.5(x_L + x_R) - (1/s_R)(0.5(x_L + x_R))^2.$$

L and R now jointly determine a policy that affects both their payoffs, $0.5(x_L + x_R)$. Again, L privately benefits from a smaller policy, and R from a

larger policy. Social payoffs are also calculated the same way with the simplification that $\tau_L = \tau_R = 0$, with this being common knowledge.

There is no Nash equilibrium in this game, and if the x_i's were bounded, L would choose the smallest possible x_L, and R would choose the largest value, for all realizations of the s_i's, and so there would be no inferences about these parameters.

A standard behavioral alternative to Nash equilibrium is the level-k model. In this model, level-0 players choose non-strategically, level-1 players best-respond to level-0 opponents, etc. At least one player's beliefs about the other's strategic thinking are incorrect, and (maximally) over-precise (completely certain), and this is true for both players if they are at the same level of strategic thinking. For example, if both players are level-2, they are both sure that the other player is level-1, and thus they are both wrong, and both certain about this incorrect belief.

The level-k solution method is easy to implement for this model. Natural assumptions for level-0 play are either the mid-point of the policy space, or the value of τ_i for each i, both of which are equal to 0. Level-k choices can then easily be shown to be $x_R = s_R + (k-1)\mu_s$ and $x_L = -s_L - (k-1)\mu_s$. Level-2 and level-3 choices for R are $x_R = s_R + \mu_s$ and $x_R = s_R + 2\mu_s$.

Clearly, under-estimating an opponent's level of strategic thinking will lead to overestimation of their s value. In fact, if both players are at the same level, they both update about s_{-i} with a signal that, as interpreted, is biased upward by μ_s. That is, a level-k L thinks that R is level-$(k-1)$, and plays $s_R + (k-2)\mu_s$. Thus, L subtracts off $(k-2)\mu_s$ from $\hat{x}_R$, leaving an "adjusted signal" of $s_R + \mu_s$ on average. More generally, player i's posterior mean for s_{-i} is a weighted average of the prior, μ_s, and the s_{-i} plus μ_s. This weighted average is always greater than the prior, and certainly greater than the true value s_{-i} in expectation (over the ϵ's), and again over-precision is part of the culprit. (A player who holds correct, and thus not over-precise, beliefs about the opponent's level does not update about s_{-i} upward on average.) Again, the effect is exacerbated by overprecision in the signal, but not in the prior. This result is summarized as follows.

Proposition A.2. *If the policy is chosen jointly and the players are level-k thinkers at the same level, $E_i(s_{-i}|\hat{x}_{-i}) > \mu_s$ and $E_i(s_{-i}|\hat{x}_{-i}) - \mu_s$ is increasing in $1/\sigma_\epsilon^i$, and decreasing in $1/\sigma_s^i$, for $i \in \{L, R\}$.*

It is simplest to characterize results for players at the same level. If the levels of thinking were different by two or more, then one player would under-estimate the other's s_i on average, and the other would over-estimate the other's s_i, but the over-estimation magnitude would be larger than that of under-estimation, so there would still be upwardly biased inferences of s_i on average (i.e., increased hostility on average).

A2. Validity of OP

I assess the validity of *OP* in a few ways. In Table A.1 I present results from analysis of the effects of *C* and *OS*'s and my measures of *OP* (OP_{OS} and *OP*, respectively) on the extremism outcomes studied by OS, using OS's models and sample definition. The outcomes include Extremism, purged issue-based extremism (residuals from regressions of estimated ideology based on issue-preferences from Hill and Tausanovitch (2015) on economic and demographic controls), and three "Strong Partisan" dummies (Strong Partisan 1 codes independents (in addition to self-reported strong members of a party) as ones; Strong Partisan 2 codes independents as zeroes; Strong Partisan 3 drops independents). *C* and OP_{OS} have similar, positive effects on Extremism; there are no significant effects on purged issue-based extremism (the coefficient for *OP* is negative); and all three measures have similar (positive) effects on party strength. Table A.2 reports these results for the sample I use in the analysis of the paper (which excludes independents, so only one definition of Strong Partisan is used); all effects for the ideology dependent variables are insignificant, while the party strength effects remain significant. These results indicate that *OP* has some predictive power for extremism outcomes, which is comparable to, though less than, that of pure confidence measures. Table A.3 presents mean values of *C*, OP_{OS} and *OP* for the various levels of education; while *C* and OP_{OS} are uncorrelated with education, *OP* consistently declines as education increases (but there is still substantial variation in *OP* within education groups), consistent with literature showing a negative correlation between education/cognitive ability and cognitive bias (Benjamin et al., 2013; Stango et al., 2016). I also examine the correlations of *C*, *K*, and *OP* with the political knowledge index. We would expect (if the variables are properly measured) that both *C* and *K* would be positively correlated with political interest and knowledge, and that *OP* would be uncorrelated or negatively correlated. The correlations are consistent with these expectations (0.19, 0.28 and -0.07, respectively). Overall, these results (predictiveness of Strong Partisan defined in various ways, and correlations with education and political knowledge) support the validity of *OP*.

A3. Validity of MR

To assess the validity of *MR*, Table A.4 presents results for *MR* included in the extremism regressions reported in Table A.1, replacing *OP/C* with measures based only on general knowledge questions to avoid collinearity (results are similar when the original *OP/C* variables are used). We would

Table A1
OS extremism outcomes.

	Extremism (self-reported)	Purged issue-based extremism	Strong Partisan 1	Strong Partisan 2	Strong Partisan 3
OP_{OS}	0.0406***	0.0698	0.0506**	0.0624***	0.0673***
	(0.0129)	(0.0440)	(0.0226)	(0.0201)	(0.0229)
N	988	993	993	993	855
C	0.0388***	0.0594	0.0509**	0.0611***	0.0657***
	(0.0130)	(0.0438)	(0.0226)	(0.0202)	(0.0232)
N	988	993	993	993	855
OP	0.0073	-0.0666	0.0689***	0.0619***	0.0693***
	(0.0137)	(0.0405)	(0.0222)	(0.0200)	(0.0230)
N	988	993	993	993	855

All models estimated by OLS and use specifications from OS (economic and demographic controls, age and media fixed effects) and robust standard errors. *, **, *** denote 10%, 5%, 1% significance.

Table A2

OS extremism outcomes for estimation sample.

	Self-reported extremism	Purged issue-based extremism	Strong Partisan 3
OP_{OS}	0.0215	0.0637	0.0511**
	(0.0151)	(0.0573)	(0.0244)
N	739	739	739
C	0.0210	0.0489	0.0519**
	(0.0152)	(0.0580)	(0.0250)
N	739	739	739
OP	−0.0088	−0.0708	0.0757***
	(0.0183)	(0.0540)	(0.0249)
N	739	739	739

All models estimated by OLS on sample used for main analysis ($F_D \leq 0$ and Distance not missing) and use specifications from OS (economic and demographic controls, age and media fixed effects) and robust standard errors. *, **, *** denote 10%, 5%, 1% significance.

Table A3

OP and *C* summary statistics by level of education.

Education		OP_{OS}	C	OP
No HS	Mean	2.133	1.798	2.556
	SD	1.109	1.108	1.173
	N	18	18	18
High school grad	Mean	1.881	1.628	2.083
	SD	0.938	0.919	0.973
	N	230	230	230
Some college	Mean	2.077	1.84	2.07
	SD	0.954	0.943	0.909
	N	177	177	177
2-year	Mean	1.992	1.69	1.924
	SD	1.053	0.981	0.922
	N	55	55	55
4-year	Mean	2.013	1.797	1.748
	SD	0.941	0.964	0.803
	N	169	169	169
Post-grad	Mean	2.142	1.988	1.702
	SD	0.853	0.891	0.714
	N	90	90	90
Total	Mean	2.004	1.77	1.956
	SD	0.947	0.946	0.91
	N	739	739	739

Summary statistics for sample used for main analysis ($F_D \leq 0$ and Distance not missing).

Table A4

OS extremism outcomes for estimation sample with *MR* controls.

	Self-reported extremism	Purged issue-based extremism	Strong Partisan
OP_{OS}	0.0131	−0.0827	0.0331
	(0.0143)	(0.0525)	(0.0237)
MR	0.0448***	0.3031***	0.0068
	(0.0097)	(0.0454)	(0.0175)
N	736	732	736
C	0.0088	−0.0524*	0.0188
	(0.0086)	(0.0310)	(0.0148)
MR	0.0447***	0.3037***	0.0069
	(0.0097)	(0.0455)	(0.0176)
N	736	732	736
OP	−0.0004	0.0267	0.0704***
	(0.0151)	(0.0496)	(0.0230)
MR	0.0632***	−0.0582	0.0077
	(0.0133)	(0.0519)	(0.0234)
N	736	732	736

All models estimated by OLS on sample used for main analysis ($F_D \leq 0$ and Distance not missing), *OP/C* based only on general knowledge questions, and use specifications from OS (economic and demographic controls, age and media fixed effects) and robust standard errors. *, **, *** denote 10%, 5%, 1% significance.

expect that motivated reasoning would be correlated with ideological extremism, and this is indeed the case, supporting the validity of estimated *MR*. The *OP* effects on Strong Partisan are robust to including *MR*, supporting the lack of redundancy between *MR* and *OP*.[31]

References

Abramowitz, A., Webster, S., 2015. All politics is national: the rise of negative partisanship and the nationalization of us house and senate elections in the 21st century. Annual Meeting of the Midwest Political Science Association Conference. pp. 16–19.

Ahler, D.J., 2014. Self-fulfilling misperceptions of public polarization. J. Polit. 76 (03), 607–620.

Andreoni, J., Mylovanov, T., 2012. Diverging opinions. Am. Econ. J. 4 (1), 209–232.

Andris, C., Lee, D., Hamilton, M.J., Martino, M., Gunning, C.E., Selden, J.A., 2015. The rise of partisanship and super-cooperators in the US house of representatives. PLoS One 10 (4), e0123507.

Azzimonti, M., 2018. Partisan conflict and private investment. J. Monet. Econ. 93, 114–131.

Barber, M., McCarty, N., 2015. Causes and consequences of polarization. Political Negotiation: A Handbook. 37. pp. 39–43.

Bénabou, R., Tirole, J., 2016. Mindful economics: the production, consumption, and value of beliefs. J. Econ. Persp. 30 (3), 141–164.

Benjamin, D.J., Brown, S.A., Shapiro, J.M., 2013. Who is 'behavioral'? Cognitive ability and anomalous preferences. J. Eur. Econ. Assoc 11 (6), 1231–1255.

Boxell, L., Gentzkow, M., Shapiro, J.M., 2017. Greater internet use is not associated with faster growth in political polarization among us demographic groups. Proc. Natl Acad. Sci. 114 (40), 10612–10617. https://doi.org/10.1073/pnas.1706588114.

Bullock, J.G., 2009. Partisan bias and the Bayesian ideal in the study of public opinion. J. Polit. 71 (03), 1109–1124.

Caplan, B., 2001. Rational irrationality and the microfoundations of political failure. Public Choice 107 (3–4), 311–331.

Caplan, B., 2002. Systematically biased beliefs about economics: robust evidence of judgemental anomalies from the survey of americans and economists on the economy. Econ. J. 112 (479), 433–458.

Chambers, J.R., Melnyk, D., 2006. Why do i hate thee? conflict misperceptions and intergroup mistrust. Pers. Soc. Psychol. Bull. 32 (10), 1295–1311.

Chambers, J.R., Schlenker, B.R., Collisson, B., 2013. Ideology and prejudice: the role of value conflicts. Psychol. Sci. 24 (2), 140–149.

Cheng, I.-H., Hsiaw, A., 2017. Distrust in experts and the origins of disagreement. In: SSNR working paper,, . https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2864563.

Congleton, R.D., 2001. Rational ignorance, rational voter expectations, and public policy: a discrete informational foundation for fiscal illusion. Public Choice 107 (1–2), 35–64.

DeVault, J.M., 2013. Political polarization, congressional redistricting, and trade liberalization. Public Choice 157 (1–2), 207–221.

Enders, A.M., Armaly, M.T., 2018. The differential effects of actual and perceived polarization. Polit. Behav. 1–25. https://doi.org/10.1007/s11109-018-9476-2.

Esponda, I., Pouzo, D., 2016. Retrospective Voting and Party Polarization. Technical Report. UC Berkeley. Working Paper

Fernbach, P.M., Rogers, T., Fox, C.R., Sloman, S.A., 2013. Political extremism is supported by an illusion of understanding. Psychol. Sci. 24 (6), 939–946.

Fryer, R.G., 2013. Updating Beliefs with Ambiguous Evidence: Implications for Polarization. Technical Report. National Bureau of Economic Research.

Gentzkow, M., Shapiro, J.M., Taddy, M., 2016. *Measuring polarization in high-dimensional data: Method and application to congressional speech* (No. w22423). National Bureau of Economic Research.

Gillen, B., Snowberg, E., Yariv, L., 2018. Experimenting with measurement error: techniques with applications to the Caltech cohort study. J. Polit. Econ forthcoming.

Graham, J., Haidt, J., Nosek, B.A., 2009. Liberals and conservatives rely on different sets of moral foundations. J. Pers. Soc. Psychol. 96 (5), 1029–1046.

Greene, J.D., 2014. Moral Tribes: Emotion, Reason, and the Gap Between Us and Them. Penguin.

Grubb, M.D., 2015. Overconfident consumers in the marketplace. J. Econ. Persp. 29 (4), 9–35.

Haidt, J., 2012. The Righteous Mind: Why Good People Are Divided by Politics and Religion. Vintage.

Hetherington, M.J., Rudolph, T.J., 2015. Why Washington Won't Work: Polarization, Political Trust, and the Governing Crisis. 104 University of Chicago Press.

Hill, S.J., Tausanovitch, C., 2015. A disconnect in representation? comparison of trends in congressional and public polarization. J. Polit. 77 (04), 10581075.

Hoppe, E.I., Kusterer, D.J., 2011. Behavioral biases and cognitive reflection. Econ. Lett. 110 (2), 97–100.

Huber, G.A., Malhotra, N., 2017. Political homophily in social relationships: evidence from online dating behavior. J. Polit. 79 (1), 269–283.

Iyengar, S., Krupenkin, M., 2018. The strengthening of partisan affect. Polit. Psychol. 39, 201–218.

Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., Westwood, S.J., 2019. The Origins

and Consequences of Affective Polarization in the United States. Annu. Rev. Polit. Sci. 22, 1.

Iyengar, S., Sood, G., Lelkes, Y., 2012. Affect, not ideology a social identity perspective on polarization. Public Opin. Q. 76 (3), 405–431.

Iyengar, S., Westwood, S.J., 2015. Fear and loathing across party lines: new evidence on group polarization. Am. J. Pol. Sci. 59 (3), 690–707.

Jha, A., Boudreaux, C.J., Banerjee, V., 2018. Political leanings and social capital. J. Behav. Exp. Econ. 72, 95–105.

Jost, J.T., Glaser, J., Kruglanski, A.W., Sulloway, F.J., 2003. Political conservatism as motivated social cognition. Psychol. Bull. 129 (3), 339.

Kahan, D., et al., 2012. Why we are poles apart on climate change. Nature 488 (7411), 255.

Kahan, D.M., 2015. The politically motivated reasoning paradigm, part 1: What politically motivated reasoning is and how to measure it. Emerging trends in the social and behavioral sciences: An interdisciplinary, searchable, and linkable resource, 1–16.

Kahan, D.M., Peters, E., Dawson, E.C., Slovic, P., 2013. Motivated Numeracy and Enlightened Self-government. Yale Law School. Public Law Working Paper

Laibson, D., Zeckhauser, R., 1998. Amos Tversky and the ascent of behavioral economics. J. Risk Uncertain. 16 (1), 7–47.

Lelkes, Y., 2016. Mass polarization: Manifestations and measurements. Public Opin. Q. 80 (S1), 392–410.

Lelkes, Y., Sood, G., Iyengar, S., 2017. The hostile audience: the effect of access to broadband internet on partisan affect. Am. J. Pol. Sci. 61 (1), 5–20.

de Leon, F.L.L., Rizzi, R., 2016. Does forced voting result in political polarization? Public Choice 166 (1–2), 143–160.

Levendusky, M., Malhotra, N., 2016. Does media coverage of partisan polarization affect political attitudes? Polit. Commun. 33 (2), 283–301.

Levy, R., 2014. Soothing politics. J. Public Econ. 120, 126–133.

Martin, G.J., Yurukoglu, A., 2017. Bias in cable news: persuasion and polarization. Am. Econ. Rev. 107 (9), 2565–2599.

Mason, L., 2015. "I disrespectfully agree": the differential effects of partisan sorting on social and issue polarization. Am. J. Pol. Sci. 59 (1), 128–145.

Mason, L., 2018. Uncivil Agreement: How Politics Became Our Identity. University of Chicago Press.

Moore, D.A., Dev, A.S., 2017. Individual differences in overconfidence. Encyclopedia of Personality and Individual Differences. Springer.

Moore, D.A., Healy, P.J., 2008. The trouble with overconfidence. Psychol. Rev 115 (2), 502–517.

Moore, D.A., Swift, S.A., 2010. The three faces of overconfidence in organizations. Soc. Psychol. Org. 147–184.

Moore, D.A., Tenney, E.R., Haran, U., 2015. Overprecision in judgment. The Wiley Blackwell Handbook of Judgment and Decision Making. pp. 182–209.

Moore, D.A., Tenney, E.R., Haran, U., 2016. Overprecision in judgment. Handbook of Judgment and Decision Making. Wiley, New York, pp. 182–212.

Neisser, U., 1971. Cognitive Psychology. Appleton-century-crofts, New York. 1967

Noori, M., 2016. Cognitive reflection as a predictor of susceptibility to behavioral anomalies. Judg. Decis. Mak. 11 (1), 114.

Ortoleva, P., Snowberg, E., 2015. Are conservatives overconfident? Eur. J. Polit. Econ. 40, 333–344.

Ortoleva, P., Snowberg, E., 2015. Overconfidence in political behavior. Am. Econ. Rev. 105 (2), 504–535.

Pessoa, L., 2008. On the relationship between emotion and cognition. Nat. Rev. Neurosci. 9 (2), 148–158.

Potrafke, N., 2018. Government ideology and economic policy-making in the United States—a survey. Public Choice 174 (1–2), 145–207.

Rempel, J.K., Sutherland, S., 2016. Hate: Theory and Implications for Intimate Relationships. The Psychology of Love and Hate in Intimate Relationships. Springer, pp. 105–129.

Robinson, R.J., Keltner, D., Ward, A., Ross, L., 1995. Actual versus assumed differences in construal: "naive realism" in intergroup perception and conflict. J. Pers. Soc. Psychol 68 (3), 404.

Rogowski, J.C., Sutherland, J.L., 2016. How ideology fuels affective polarization. Polit. Behav. 38 (2), 485–508.

Rollwage, M., Dolan, R.J., Fleming, S.M., 2018. Metacognitive failure as a feature of those holding radical beliefs. Curr. Biol. 28 (24), 4014–4021.

Rozin, P., Royzman, E.B., 2001. Negativity bias, negativity dominance, and contagion. Pers. Soc. Psychol. Rev. 5 (4), 296–320.

Schnellenbach, J., Schubert, C., 2015. Behavioral political economy: a survey. Eur. J. Polit. Econ. 40, 395–417.

Selinger, J.S., 2016. Embracing Dissent: Political Violence and Party Development in the United States. University of Pennsylvania Press.

Stango, V., Yoong, J., Zinman, J., 2016. *We are all behavioral, more or less: Measuring the prevalence, heterogeneity and importance of multiple behavioral factors*. In: Working

[31] I (arbitrarily) top-code unemployment guesses at 20%. I examined several variants on this definition of *MR*, including using all four economic questions, using negative values for responses unfavorable to the co-party, using just the economic knowledge questions, and using just the survey questions on economic performance without confidence measures, and different types of top-coding, and obtained generally similar results, though validation results are much weaker without the economic knowledge questions.

paper, . Available at https://dartmouth.edu/jzinman/Papers/WeAreAllBehavioralDraft1.pdf.

Stone, D.F., 2016. Just a Big Misunderstanding? Bias and Bayesian Affective Polarization. SSRN Working Paper

Sunstein, C.R., 2015. Partyism. University of Chicago Legal Forum. pp. 1.

Tajfel, H.E., 1978. Differentiation Between Social Groups: Studies in the Social Psychology of Intergroup Relations. Academic Press.

Tausanovitch, C., Warshaw, C., 2013. Measuring constituent policy preferences in congress, state legislatures, and cities. J. Polit. 75 (02), 330–342.

Tetlock, P., Gardner, D., 2015. Superforecasting: The Art and Science of Prediction. Random House.

Westfall, J., Van Boven, L., Chambers, J.R., Judd, C.M., 2015. Perceiving political polarization in the united states party identity strength and attitude extremity exacerbate the perceived partisan divide. Persp. Psychol. Sci. 10 (2), 145–158.

Westwood, S.J., Iyengar, S., Walgrave, S., Leonisio, R., Miller, L., Strijbis, O., 2018. The tie that divides: cross-national evidence of the primacy of partyism. Eur. J. Polit. Res. 57 (2), 333–354.