

Overconfidence in Political Behavior*

Pietro Ortoleva
Columbia University

pietro.ortoleva@columbia.edu
www.columbia.edu/~po2205/

Erik Snowberg
California Institute
of Technology and NBER
snowberg@caltech.edu
www.hss.caltech.edu/~snowberg/

July 1, 2014

Abstract

This paper studies, theoretically and empirically, the role of overconfidence in political behavior. Our model of overconfidence in beliefs predicts that overconfidence leads to ideological extremeness, increased voter turnout, and increased strength of partisan identification. Moreover, the model makes many nuanced predictions about the patterns of ideology in society, and over a person's lifetime. These predictions are tested using unique data that measure the overconfidence, and standard political characteristics, of a nationwide sample of over 3,000 adults. Our numerous predictions find strong support in these data. In particular, we document that overconfidence is a substantively and statistically important predictor of ideological extremeness and voter turnout.

JEL Classifications: D03, D72, D83, C83

Keywords: behavioral political economy, overconfidence, ideology, extremeness, voting, political parties

*Snowberg gratefully acknowledges the support of NSF grant SES-1156154. We thank Stephen Ansolabehere, Marc Meredith, Chris Tausanovitch, and Christopher Warshaw for sharing their data, and Sergio Montero and Gerelt Tserenjigmid for excellent research assistance. The authors are indebted to John Aldrich, Scott Ashworth, Larry Bartels, Roland Benabou, Jon Bendor, Adam Berinsky, Ethan Bueno de Mesquita, John Bullock, Steve Callander, Pedro Dal Bó, Ben Gillen, Faruk Gul, Horacio Larreguy, Gabe Lenz, Alessandro Lizzeri, John Matsusaka, Andrea Mattozzi, Antonio Merlo, Massimo Morelli, Steve Morris, Gerard Padró-i-Miquel, Eric Oliver, Wolfgang Pesendorfer, Matthew Rabin, Ken Shotts, Holger Sieg, Theda Skocpol, Mike Ting, Francesco Trebbi, Leeat Yariv, and Eric Zitzewitz for useful discussions. We also thank seminar participants at the AEA, the University of British Columbia, the University of Chicago, Columbia, Duke, Harvard, the IV Workshop on Institutions at CRENoS, the University of Maryland, MPSA, the NBER, the Nanyang Technological University, NYU, HEC Paris, the University of Pennsylvania, Princeton, The Prioriat Workshop, USC, SPSA, and Washington University, St. Louis for thoughtful feedback.

1 Introduction

Without heterogeneity in *ideology*—preferences and opinions over political actions—there would be little need for the institutions studied by political economists.¹ However, the sources of ideology have received scant attention: since Marx, political economists have largely viewed ideology as driven by wealth or income—despite the fact that these variables explain little of the variation in ideology (Meltzer and Richard, 1981; Frank, 2004; Acemoglu and Robinson, 2006; Gelman, 2009).

This paper proposes a complementary theory in which differences in ideology are also due to imperfect information processing. This theory predicts that *overconfidence* in one’s own beliefs leads to ideological extremeness, increased voter turnout, and stronger identification with political parties. Our predictions find strong support in a unique dataset that measures the overconfidence, and standard political characteristics, of a nationwide sample of over 3,000 adults. In particular, we find that overconfidence is the most reliable predictor of ideological extremeness, and an important predictor of voter turnout in our data.

By adopting a behavioral basis for ideology, we help answer puzzling questions such as why politicians and voters are becoming more polarized despite the increased availability of information (McCarty et al., 2006), or why political rumors and misinformation, such as “Global warming is a hoax”, are so persistent (Berinsky, 2012).² Moreover, as behavioral findings deepen our understanding of market institutions (Bertrand, 2009; Baker and Wur-gler, 2013), a behavioral basis for ideology promises greater understanding of the design and consequences of political institutions (Callander, 2007; Bisin et al., 2011).

In our model, overconfidence and ideology arise due to imperfect information processing. Citizens passively learn about a state variable through their experiences (signals). However, to varying degrees, citizens underestimate how correlated these experiences are, and thus,

¹Without heterogeneity, institutions will still be useful for coordination (Weingast, 1997). Traditionally, opinions have not been considered part of ideology, however, recent work has provided compelling arguments that ideology should include some beliefs (McMurray, 2013a,b).

²In early 2013, 37% of U.S. voters agreed with this statement (Public Policy Polling, 2013). Only 41% believe global warming is caused by human activity, compared with 97% of climate scientists (Yale Project on Climate Change Communication, 2013). Similar levels of agreement with other political rumors or “conspiracy theories” are regularly found among voters (Berinsky, 2012; Public Policy Polling, 2013).

have different levels of overconfidence about their information. This underestimation—which we call *correlational neglect*—may have many sources. For example, citizens may choose to get information from a biased media outlet, but fail to fully account for the bias. Indeed, unbeknownst to most users, Google presents different news sources for the same search depending on a user’s location.³ Alternatively, exchanging information on a social network could lead to correlational neglect if citizens fail to understand that much of the information comes from people similar to themselves, if they fail to recognize the influence of their own previous reports on others’ current reports (DeGroot, 1974; DeMarzo et al., 2003), or if they fail to account for the presence of rational herds (Eyster and Rabin, 2010). Recent laboratory experiments find strong evidence of correlational neglect (Enke and Zimmermann, 2013).

Our primary theoretical result is that overconfidence and ideological extremeness are connected. This follows an uncomplicated logic. For example, consider a citizen who notes the number of people in her neighborhood who are unemployed, and uses this information to deduce the state of the national economy. Suppose further that she lives in a neighborhood with high unemployment. If the citizen believes that the employment status of each person is relatively uncorrelated, she will think she has a lot of information about the state of the national economy—she will be overconfident—and favor generous aid to the unemployed and loose monetary policy. If, instead, she realizes that local unemployment has a common cause—say, a factory shutting down—then she will understand that she has comparatively little information about the national economic situation, and believe that although the situation is bad, it is not likely to be dire, and will support more moderate policies.

Our data—from the 2010 Cooperative Congressional Election Survey (CCES)—strongly supports this prediction. A one-standard-deviation change in overconfidence is related to 12–28% of a standard-deviation change in ideological extremeness, depending on the specification. This relationship is as large as, and more stable than, the relationship between extremeness and economic variables. Indeed, the range of correlations for each economic

³See: <http://vimeo.com/51181384>. Actively seeking information that matches ones’ prior, or ignoring information counter to it, is termed *confirmation bias*. Although as Levy and Razin (2014) note, these two different structures of confirmation bias will lead to different patterns of information aggregation in elections, our model encompasses both these common forms of confirmation bias.

variable include points that are statistically indistinguishable from zero, suggesting that overconfidence is an important and distinct predictor of ideological extremeness.

The size and complexity of this data allows for the testing of more subtle predictions. For example, if citizens become more conservative when they have more experiences or signals, through aging or media exposure, then overconfidence should be correlated with conservatism. Moreover, extremism should be more correlated with overconfidence for conservatives than liberals. These results find robust support in the data.

To extend this model to voter turnout, we posit an expressive voting model in which the expressive value of voting is increasing with a citizen's belief that one party's policy is better for her (Matsusaka, 1995; Degan and Merlo, 2011; Degan, 2013). Similarly, strength of partisan identification is modeled as the probability a citizen places on her favored party's policy being better for her.

As more overconfident citizens are more likely to believe that one or the other party is likely to have the right policy for them, they are more likely to turn out to vote. This is true even conditional on ideology. The opposite conditional statement also holds: more ideologically extreme citizens are more likely to vote, conditional on overconfidence. Thus, our model matches the well-known empirical regularity that more ideologically extreme citizens are more likely to vote. Similar predictions hold for strength of partisan identification.

This second set of predictions are, once again, robustly supported by the data. Using verified voter turnout data, we document that a one-standard-deviation change in overconfidence is associated with 7–19% increase in voter turnout. This is a more important predictor of turnout in our data than income, education, race, gender, or church attendance.

Finally, we theoretically analyze how our results would be altered by citizens choosing how much (costly) information to acquire, or communicating their ideology to each other. Both of these extensions strengthen our primary results. When citizens can acquire more information, it will be more overconfident citizens that do so. This occurs because more overconfident citizens neglect correlation to a greater degree, and hence believe that additional signals are more valuable. When citizens can communicate, more-overconfident citizens will attribute

differences in ideology to anything other than their own information being incorrect, and hence update less than less-overconfident citizens. Both of these possibilities accentuate the correlation between ideological extremeness and overconfidence.

The remainder of this section provides more details on the behavioral phenomena of overconfidence, connects our work to the literature, and previews the structure of the paper.

1.1 What is Overconfidence?

Overconfidence describes related phenomena in which a person thinks some aspect of his or her's, usually performance or information, is better than it actually is. These phenomena are the subject of a large literature in psychology, economics, and finance, having been first documented in Alpert and Raiffa (1969/1982). This literature has documented overconfidence in a wide range of contexts, and among people from a wide range of backgrounds.

Moore and Healy (2007, 2008) divide overconfidence into three, often conflated, categories: over-estimation, over-placement, and over-precision. Over-estimation is when people believe that their performance on a task is better than it actually is. Over-placement is when people incorrectly believe that they perform better than others—as in the classic statement that, “93% of drivers believe that they are better than average.”⁴

In this paper we focus on over-precision: the belief that one's information is more precise than it actually is. There are two reasons for this focus. First, while over-estimation and over-placement often suffer from reversals,⁵ this does not seem to be the case for over-precision. In other words, it appears that (almost) everyone exhibits over-precision (almost) all the time (Moore and Healy, 2007, 2008). Second, over-precision has a very natural interpretation in political contexts: it is the result of people believing that their own experiences are more informative about policy than they actually are. Despite our narrower focus, we continue to use the term overconfidence.

Overconfidence is usually a modeling fundamental. By contrast, as noted above, we

⁴Interestingly, this may be perfectly rational; see Benoît and Dubra (2011).

⁵That is, people tend to perceive their performance as better than it actually is when a task is easy, and worse when the task is difficult (Erev et al., 1994).

derive it as a consequence of correlational neglect. This allows for the derivation of additional predictions on the evolution of overconfidence and extremism with age and media exposure.

1.2 Literature

This work contributes to the emerging literature on behavioral political economy, which applies findings from behavioral economics to understand the causes and consequences of political behavior.⁶ This approach promises to allow political economists to integrate the insights of a half-century of psychology-based political behavior studies.

A particular appeal of applying behavioral insights to political economy is that many of the feedback mechanisms that have led scholars to doubt the importance of behavioral phenomena in markets do not seem to exist in politics. Specifically, as an individual's political choice is unlikely to be pivotal, citizens who make poor political choices do not suffer worse consequences than those who make good political choices. Moreover, this lack of direct feedback drastically reduces a citizen's ability to learn of her bias. This is in stark contrast to markets, where poor choices directly impact the decision-maker, which some scholars argue will eliminate behavioral biases. Furthermore, behavioral traits that may be detrimental in markets may, in some cases, be useful in facilitating collective action (Benabou and Tirole, 2002, 2006; Benabou, 2008).

Two papers focus on the normative implications of correlational neglect in political economy. Glaeser and Sunstein (2009) studies "credulous Bayesian" information transmission in groups. This follows models in which correlational neglect is related to network structure (DeGroot, 1974; DeMarzo et al., 2003; Golub and Jackson, 2010; Chandrasekhar et al., 2012). It notes that this bias may lead to group polarization, overconfidence in beliefs, and worse aggregate decision-making. Levy and Razin (2014), in contrast, shows that correlational neglect may lead to better information aggregation in elections. Our work has a different focus: we present a general model of correlational neglect, and derive and test positive results

⁶This literature is small, and includes Matsusaka (1995); Bendor et al. (2003, 2011); Callander and Wilson (2006, 2008); Bisin et al. (2011); Degan and Merlo (2011); Lizzeri and Yariv (2012); Passarelli and Tabellini (2013).

to understand the inter-relationships between overconfidence, ideology, extremism, turnout, party identification, media exposure, and age.

This paper is related to a number of additional literatures. First and foremost, the study of ideology, voting, and partisan identification are the subject of massive literatures in political science. Second, overconfidence is the focus of a large literature in behavioral economics and finance (see, for example, Odean, 1998; Daniel et al., 1998; Camerer and Lovo, 1999; Santos-Pinto and Sobel, 2005). Third, there are a small number of papers that study the role of beliefs on preferences for redistribution (Piketty, 1995; Alesina and Angeletos, 2005). Fourth, our modeling technique comes from the small literature utilizing the normal learning model.⁷ Finally, our model of turnout follows those in which voters are either regret- or choice-avoidant (Matsusaka, 1995; Degan and Merlo, 2011; Degan, 2013).⁸

1.3 Structure

This paper is unconventionally structured: it rotates between theoretical and empirical results. This allows for the data to inform the theoretical analysis, and clarifies the role that assumptions play in results. Section 2 introduces the theoretical structure and data. Our analysis begins in Section 3.1 with an examination of how overconfidence and ideology evolve with the number of signals. This preliminary check shows that implications of our model that differ from a fully Bayesian benchmark find strong support in our data. Section 3.2 examines our primary result: the correlation between overconfidence and ideological extremeness. In addition, the restrictions from the previous subsection allows for predictions about ideology and overconfidence. The final two subsections of Section 3 examine additional, more subtle, predictions about the relationships between overconfidence, ideology, and extremeness.

Section 4 adds the expressive voting structure that allows our model to generate predictions for turnout and partisan identification. Section 5 examines theoretically how our results

⁷Although the literature is not large, it cannot be completely reviewed here. Early papers include Zechman (1979), Achen (1992). For a recent review, see the introduction of Bullock (2009). In this literature, our model is closest to Blomberg and Harrington (2000), although like all fully Bayesian models, this model is inconsistent with the data here, as discussed in Section 3.1.

⁸For a discussion of how our results relate to other models of voter turnout, see Appendix D.

would change if citizens could acquire additional information. Finally, Section 6 discusses issues related to identification, and directions for future work.

2 Framework and Data

This section presents our model, and formally defines correlational neglect and overconfidence. This is followed by a discussion of our data, and how we use it to construct measures of overconfidence, ideology, voter turnout, and partisan identification.

2.1 Theoretical Framework

There is a unit measure of citizens $i \in [0, 1]$. Each citizen i has a utility over political actions that depends on a state of the world. A citizen’s belief about the state are determined by her experiences, and ideology encompasses both beliefs about the state and preferences.

Utilities: Each citizen i has a standard quadratic-loss utility over actions $a_i \in \mathbb{R}$, which depends on a state $x \in \mathbb{R}$, and a preference bias b_i

$$U(a_i, b_i | x) = -(a_i - b_i - x)^2.$$

Throughout this paper a_i is the policy implemented by government. A citizen’s preference bias is an i.i.d. draw from a normal distribution with mean 0 and precision τ_b . We write this as $b_i \sim \mathcal{N}[0, \tau_b]$. The state x is a single draw from $\mathcal{N}[0, \tau]$.

With uncertainty about the state, it is straightforward to show that the policy preferred by citizen i will be $a_i^* = b_i + \mathbb{E}_i[x]$, where \mathbb{E}_i is the expectation taken over citizen i ’s beliefs. We define this quantity as the citizen’s *ideology*,

$$\mathcal{I}_i = b_i + \mathbb{E}_i[x], \tag{1}$$

and, as the expectation of x is zero, ideological extremeness as $\mathcal{E}_i = |\mathcal{I}_i|$.

Experiences, Beliefs, and Correlational Neglect: The core of the model is the process by which citizens form beliefs about the state. In our model, each citizen is well-calibrated

about the informativeness of individual experiences, but underestimates how correlated her experiences are. This will lead to varying degrees of overconfidence in the population.

Each citizen starts with the correct prior $\mathcal{N}[0, \tau]$ about the state. Citizens have multiple experiences over time, which are signals about the state, $e_{it} = x + \varepsilon_{it}$, $t \in \{1, 2, \dots, n_i\}$, with e_{it} independent of b_i , that is $e_{it} \perp b_i$. Each $\varepsilon_{it} \sim \mathcal{N}[0, 1]$, and the signals are correlated, with $\text{Corr}[\varepsilon_{it}, \varepsilon_{it'}] = \rho$.⁹ However, citizen i underestimates this correlation: she believes $\text{Corr}[\varepsilon_{it}, \varepsilon_{it'}] = \rho_i \in [0, \rho)$.

Definition. *A citizen suffers from **correlational neglect** when $\rho_i < \rho$.*

The magnitude of correlational neglect varies by citizen, and is an i.i.d. draw from F_{ρ_i} with support $[0, \rho)$, and $\rho_i \perp (e_{it}, b_i)$. As $\rho_i < \rho$ for all i , all citizens in our model are correlational neglecters.

For tractability, we assume n_i , the number of signals received by citizen i , is exogenous. Section 5.1 relaxes this assumption, and shows that more overconfident citizens value additional information more highly, and that endogenizing the acquisition of information strengthens our results.

Overconfidence: As our data measures overconfidence, our theoretical results are in terms of this variable. Denote the precision of citizen i 's posterior belief as $\kappa_i + \tau$, which we refer to as the citizen's *confidence*. Additionally, denote by $\kappa + \tau$ the posterior belief the citizen would have if she had accurate beliefs about the correlation between signals.

⁹Formally, ε_i is distributed according to a mean-zero multinomial normal with covariance matrix

$$\Sigma_{\varepsilon_i} = \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \cdots & 1 \end{pmatrix}. \text{ However, citizen } i \text{ believes that } \Sigma_{\varepsilon_i} = \begin{pmatrix} 1 & \rho_i & \cdots & \rho_i \\ \rho_i & 1 & \cdots & \rho_i \\ \vdots & \vdots & \ddots & \vdots \\ \rho_i & \rho_i & \cdots & 1 \end{pmatrix}.$$

Each ε_{it} has unit variance, so $\text{Corr}[\varepsilon_{it}, \varepsilon_{it'}] = \text{Cov}[\varepsilon_{it}, \varepsilon_{it'}] = \rho$.

Alternatively, we could model the state in a multi-dimensional space with multi-dimensional errors over time, and citizens either underestimate the amount of correlation between dimensions, or across time, or both. This does not add to the testable predictions of the model, see Appendix D.

Definition. Overconfidence is the difference between a citizen’s confidence, and how confident she would be if she were properly calibrated, $\kappa_i - \kappa$. Given two citizens i and j , we say that i is **more overconfident** than j if $\kappa_i \geq \kappa_j > 0$.¹⁰

We often refer to κ_i as a citizen’s level of overconfidence.

Before discussing data, we briefly comment on the interpretation of x . Research that uses a similar utility function sees this state variable as informative of optimal policy. However, this may lead to unappealing normative implications. In particular, it suggests that the optimal policy could be found by studying the distribution of public opinion, and that this policy may be more extreme than the median of fairly extreme groups. To keep the same structure, but eliminate such conclusions, we could simply add an aggregate bias to the signaling structure.¹¹ As this additional uncertainty affects every citizen in the same way, it does not affect the aggregate conclusions in this paper. However, it would affect how citizens learn from each other, a subject we examine in Section 5.2.

2.2 Data

Our data comes from the Harvard and Caltech modules of the 2010 Cooperative Congressional Election Study (CCES) (Alvarez, 2010; Ansolabehere, 2010a,b). This data is unique (as far as we know) in that it allows a survey-based measure of overconfidence in beliefs as well as political characteristics.

The CCES is an annual cooperative survey. Participating institutions purchase a module of at least 1,000 respondents, who are asked 10–15 minutes of customized questions. In addition, every respondent across all modules is asked the same battery of basic economic and political questions. The complete survey is administered online by Knowledge Networks. Each module uses a matched-random sampling technique to achieve a representative sample, with over-sampling of certain groups (Ansolabehere, 2012; Ansolabehere and Rivers, 2013).

¹⁰All results hold defining overconfidence as κ_i/κ .

¹¹ This would be expressed formally as: $\varepsilon_{it} \sim \mathcal{N}[\pi_0, 1]$ with $\pi_0 \sim \mathcal{N}[0, 1]$ identical for all citizens.

2.2.1 Overconfidence

The most important feature of this data, for our purposes, is that it allows for a measure of overconfidence. This measure is constructed from four subjective questions about respondent confidence in their guesses about four factual quantities, adjusting for a respondent’s accuracy on the factual question. This is similar to the standard psychology measure in that it elicits confidence and controls for knowledge. However, it differs in that we cannot say for certain whether a given respondent is overconfident, just that their confidence, conditional on knowledge, is higher or lower than another respondent. Therefore, we use previous research, which shows that (almost) everyone exhibits over-precision (almost) all the time (Moore and Healy, 2007, 2008), to argue that this is a measure of overconfidence.¹²

The factual and confidence questions were asked as part of another set of studies (Ansolabehere et al., 2011, 2013). Respondents were asked their assessment of the current unemployment and inflation rate, and what the unemployment and inflation rate would be a year from the date of the survey. Respondents were then asked their confidence about their answer to each factual question on a qualitative, six-point scale.

Confidence reflects both knowledge and overconfidence, so subtracting knowledge from confidence leaves overconfidence.¹³ 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.¹⁴ This allows the regression to pick the points to deduct for each level of accuracy, such that knowledge absorbs as much variation as possible.

¹²Psychological studies typically elicit a large (up to 150) number of 90% confidence intervals and count the percent of times that the actual answer falls within a subject’s confidence interval. This number, subtracted from 90, is used as a measure of overconfidence. Our measure has advantages over the typical psychology approach—see Appendix B, which also contains all survey questions.

¹³Theoretically, we need to control for the precision a citizen would have if they were properly calibrated. As we do not observe this, we control for accuracy, which is, in our theory, correlated.

¹⁴That is, we use a semi-nonparametric sieve method to control for knowledge (Chen, 2007). Ideally one would impose a monotonic control function, however, doing so is methodologically opaque, see Athey and Haile (2007); Henderson et al. (2009). In keeping with the treatment of these factual questions in Ansolabehere et al. (2011, 2013), we topcode responses to the unemployment and inflation questions at 25, limiting a respondent’s inaccuracy.

Table 1: Overconfidence is correlated with gender and age, but not education or income.

Dependent Variable:	Overconfidence			Confidence		
Gender (Male)	0.45*** (.078)	0.44*** (.080)	0.43*** (.079)	0.48*** (.077)	0.47*** (.078)	0.46*** (.0022)
Age (in years)	0.012*** (.0023)	0.013*** (.0024)	0.012*** (.0023)	0.013*** (.0022)	0.014*** (.0023)	0.013*** (.0022)
Education	$F = 1.12$ $p = 0.36$			$F = 2.03$ $p = 0.08$		
Income	$F = 1.33$ $p = 0.21$			$F = 1.82$ $p = 0.05$		
N	2,927					

Notes: ***, **, * denote statistical significance at the 1%, 5% and 10% level with standard errors, clustered by age (73 clusters), in parentheses. All specifications estimated using WLS with CCES sampling weights.

Each of the resultant overconfidence measures are measured with error, as some respondents with little knowledge will randomly provide accurate answers. Thus, we use the first principal component of the four measures.¹⁵ Finally, to standardize regression coefficients, we subtract the minimum level of overconfidence, and divide by the standard deviation.

In keeping with previous research, overconfidence is strongly correlated with a respondent's gender, as shown in Table 1 (see, for example, Lundeberg et al., 1994). Section 3.1 predicts that overconfidence is correlated with age. This is also clear in Table 1. This predicted relationship leads us to cluster standard errors by age.¹⁶ Additionally, as the CCES over-samples certain groups, such as voters, we estimate specifications using WLS and the supplied sample weights (Ansolabehere, 2012).

However, overconfidence is uncorrelated with education or income. Note that these latter controls are ordered categorical variables, so we provide F -tests on the five and fifteen

¹⁵Consistent with each measure consisting of an underlying dimension plus i.i.d. measurement error, the first principal component weights each of the four questions approximately equally. Also consistent with this structure, our results are substantively similar using any one of the four questions in isolation. So, for example, they hold if we use only variables pertaining to present conditions, or only to future predictions.

¹⁶Age also has a greater intraclass correlation than state of residence, education, or income, making age the most conservative choice. The intraclass correlation is small for all of these variables, thus clustering on any one of them produces similar results, which are also similar to heteroskedastic-consistent standard errors. For consistency, we continue to cluster by age when age is on the right-hand side. Classical standard errors are approximately 25% smaller.

dummy variables that, respectively, represent these categories. For comparison, we construct a confidence measure from the first principal component of confidence scores. Education and income are related to this measure, providing some confirmation that actual knowledge has been purged from the overconfidence measure.¹⁷

While the data we use to elicit overconfidence is quite similar to that used in psychology, there are some differences. First, we use questions about economic measures—unemployment, inflation—as opposed to general knowledge questions—for example, “When was Shakespeare born?” Second, these questions elicit confidence directly, while studies in psychology typically elicit confidence intervals. To understand whether our slightly different approach provides similar results, we added four general knowledge questions—eliciting confidence with an interval—to the 2011 CCES. The 2011 CCES also included the confidence questions from the 2010 version. The main finding is reassuring: the results we can examine in the (more limited) 2011 CCES hold using general knowledge-based measures of overconfidence. These results can be found in Section 6.1, and more about using surveys to measure overconfidence can be found in Appendix B.

2.2.2 Dependent Variables

The predictions in this paper concern three types of dependent variables: ideology, voter turnout, and strength of partisan identification.

Ideology: This study uses one main and two alternative measures of ideology. The main measure is scaled ideology from Tausanovitch and Warshaw (2011), which they generously provided to us. This measure is generated using item response theory (IRT) to scale responses to eighteen issue questions asked on the CCES—for example, questions about abortion and gun control. A similar process generates the Nominate Scores used to evaluate the ideology of members of Congress (Poole and Rosenthal, 1985).¹⁸

¹⁷All our results—theoretical and empirical—hold using confidence rather than overconfidence.

¹⁸There are many ways to aggregate these individual issues into ideology. For example, one could aggregate groups of related issues into different ideological dimensions. To eliminate concerns about specification searching, we prefer to use a measure generated by other scholars, see Appendix D.

Our alternative measures of ideology are direct self-reports. The CCES twice asks respondents to report their ideology: from extremely liberal to extremely conservative. The first elicitation is when the respondent agrees to participate in surveys (on a five point scale), and the second when taking the survey (on a seven point scale). We normalize each of these measures to the interval $[-1, 1]$, and average them. Those that report they “don’t know” are either dropped from the sample, or treated as moderates (0). Results are presented for both cases. These self-reported measures are imperfectly correlated with scaled ideology (0.42).¹⁹

Voter Turnout: Turnout is ascertained from the voting rolls of the state in which a respondent lives. Voter rolls vary in quality between states, but rather than trying to control for this directly, we include state fixed effects in most of our specifications.²⁰

Partisan Identification: At the time of the survey, respondents were asked whether they identify with the Republican or Democratic Party, or neither. If they report one of the political parties—for example the Democrats—they are then asked if they are a “Strong Democrat” or “Not so Strong Democrat”. Those who report they were neither Republican or Democrat were asked if they lean to one party or the other, and are allowed to say that they do not lean toward either party. Those who report they are strong Democrats or Republicans are coded as strong partisan identifiers. Independents—those who do not lean toward either party—are coded as either strong party identifiers, weak party identifiers, or are left out of the analysis. Results are presented for all three resultant measures.

¹⁹In improving and refining the paper in accordance with referee suggestions, we eliminated a number of specifications found in the working paper (Ortoleva and Snowberg, 2013a). This working paper includes more specifications with alternative ideology measures, extremism measures that are constructed directly from ideology without first controlling for economic variables, demographic controls, and unweighted specifications. The results in all cases are substantially similar.

²⁰The state of Virginia did not make their rolls available, so the 60 respondents from Virginia are dropped from turnout regressions (see Ansolabehere and Hersh, 2010). Classifying as non-voters the 42 respondents who were found to have voted in the primary but not the general election does not change the results.

2.2.3 Controls

Economic Controls: Political economy theories generally view ideology as a function of wealth or income. Therefore, we include controls for all the wealth and income related variables in the CCES. The CCES provides these controls as categories: for example, rather than providing years of education, it groups education into categories such as “Finished High School”. Thus, we introduce a dummy variable for each category of each economic control. We also include a category for missing data for each variable. These controls are: income (16 categories), education (6 categories), stock ownership (3 categories), home ownership (4 categories), union / union member in household (8 categories), state (52 categories, including DC and missing).

Number of Signals: The CCES contains two sets of questions that are reasonable proxies for the number of signals: media exposure, and age. The CCES contains four questions that ask whether or not a respondent received news from a specific media channel: blogs, TV, radio, and newspapers. We take the first principal component of these four yes/no questions to create a more continuous index of media exposure. This principal component also de-emphasizes TV, as nearly all respondents report getting political information from this channel. Age is calculated as 2010—the year of the survey—minus birth year.

All controls are entered categorically. This strategy is both too conservative and not conservative enough. Not conservative enough because there are likely other relevant unobserved factors, and too conservative, as entering these variables categorically allows the implied control function to be non-monotonic, as opposed to the theoretical monotonic relationship. There are 16 categories of media exposure, and 73 categories (years) for age.

3 Ideology, Extremeness, and Overconfidence

We begin our analysis by contrasting our model with fully Bayesian models, and then continue to our primary results.

3.1 Media, Age, Overconfidence, and Ideology

Our model is partially Bayesian—citizens update using Bayes’ rule, but with an incorrect likelihood function. Obvious alternative models are *fully Bayesian*, in which citizens are correct about the link between their experiences and the state x , and citizens eventually learn x . In this subsection we show that such alternative models are incompatible with an important feature of the data: ideological extremeness increases with the number of experiences (signals).

In particular, consider any fully Bayesian model in which citizens’ priors have full support, and they receive n private signals. Then there exists some n^* such that if $n > n^*$ the variance of citizens’ posterior means is non-increasing in n . Note the fact that citizens’ posterior means may initially diverge is due to the fact that if citizens have a common prior mean, the first signal(s) will cause divergence. After these initial signals, the variance of posterior means must weakly decrease as more information is revealed and citizens learn x . Of course, if all citizens’ prior means are x , then posteriors need never diverge or converge.

This implies that the population variance of ideology, conditional on n , $\text{Var}[\mathcal{I}_i|n]$ is non-increasing in n in fully Bayesian models. In contrast, in our model:

Proposition 1.

1. *Overconfidence is increasing with the number of experiences (signals) n .*
2. *The mean ideology in the population, conditional on n , $\mathbb{E}[\mathcal{I}_i|n]$, is increasing in n if and only if $x > 0$, and decreasing in n iff $x < 0$.*
3. *If ρ is large enough, $\text{Var}[\mathcal{I}_i|n]$ is increasing in n .*

Proof. All proofs are in Appendix A. ■

To build intuition for Proposition 1, consider the extreme case in which $\rho_i = 0$ and $\rho = 1$; that is, when experiences are perfectly correlated, but citizen i believes that they are independent. In this case, each experience is identical, so it will make the citizen more confident without increasing her information—leading to the first part of the proposition. Moreover, as the mean of the ideology distribution will tend toward x as n increases, if $x > 0$, then ideology will increase with n .

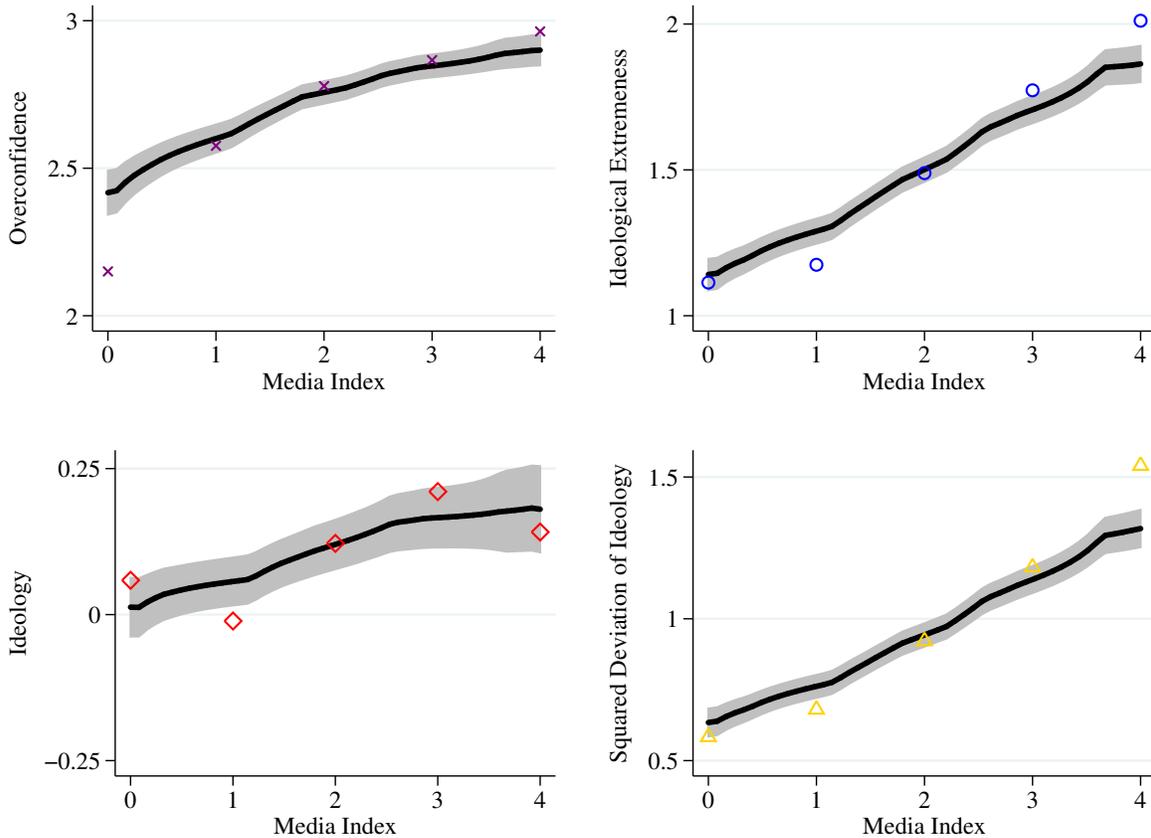
The final part of Proposition 1 implies that under certain parameter values our model is compatible with patterns that are incompatible with fully Bayesian models: citizens’ beliefs could become more polarized with more signals. To understand the difference, note that when ρ is large in our model, then each additional signal contains very little new information, yet some citizens believe it does. New signals will thus push their beliefs towards a biased view of x , rendering subjects more polarized. As described above, this would not be the case in a fully Bayesian model, as citizens beliefs will converge towards x —or, at least, not diverge.

Before turning to the empirical examination of this proposition, we note that this proposition provides a potential answer to the first puzzle posed in the introduction: why politicians and voters are becoming more polarized, despite the increased availability of information through the internet (McCarty et al., 2006). The third part of the proposition suggests that an increase in the number of signals can actually increase ideological extremeness, and thus, polarization. This occurs because additional signals are correlated, and thus provide limited additional information. Citizens neglect this correlation, and thus update “too much”, which increases polarization. Note that this occurs even if media exposure is not more polarized, as seems to be the case (Gentzkow and Shapiro, 2011).

3.1.1 Empirical Analysis

We examine the patterns suggested by Proposition 1, using media exposure as a measure of n , in Figure 1. The visual patterns in Figure 1 are found to be statistically robust in Table 2. Each panel of Figure 1 shows a smoothed, non-parametric fit with 95% confidence intervals, and averages for each value of the media index. The first panel shows that, in accordance with Proposition 1, overconfidence increases with media exposure. The second panel shows that ideological extremeness increases with media exposure. The third and fourth panels show that this is due to both a rightward shift associated with more media exposure, and an increase in ideological dispersion. This, along with the second and third parts of Proposition 1, implies certain restrictions on the parameters of the model. In particular:

Figure 1: Media Exposure, Overconfidence, and Ideology



Notes: Each point is the average for those respondents with a specific index of media exposure intensity (that is, blogs, tv, radio, and newspaper) rounded to the nearest integer. Trendiness, in black, and 95% confidence intervals, in gray, use an Epanechnikov kernel with a bandwidth of 0.8.

Implication 1. $x > 0$.²¹

Implication 2. ρ is large enough so that $\text{Var}[\mathcal{I}_i|n]$ is increasing in n .

The second implication can be examined in other datasets: specifically, we replicate the fourth panel of Figure 1 using data from the American National Election Survey, see Appendix D.

While the fourth panel of Figure 1 shows that the data is inconsistent with any fully Bayesian model, neither it, nor the third panel, is a test of our model. Specifically, the second and third part of Proposition 1 allows for either an increasing or decreasing relationship, depending on parameter values. Moreover, in proving this result we have assumed that media exposure is exogenous, which is unlikely to hold. We address this issue in two ways.

²¹Note that this does not imply that conservative citizens are “correct”—see the end of Section 2.1.

Table 2: Regressions support the visual patterns in Figure 1.

Dependent Variable:	Overconfidence		Ideology		Squared Deviation	
Media Index	0.21*** (.041)	0.15*** (.038)	0.061*** (.022)	0.057** (.022)	0.29*** (.028)	0.19*** (.028)
Economic Controls		Y		Y		Y
Age (73 Categories)		Y		Y		Y
R^2	0.043	0.23	0.037	0.17	0.085	0.26
Age	0.20*** (.037)	0.15*** (.038)	0.12*** (.030)	0.066** (.030)	0.16*** (.030)	0.14*** (.033)
Economic Controls		Y		Y		Y
Media (16 Categories)		Y		Y		Y
R^2	0.040	0.19	0.014	0.16	0.027	0.24
N	2,927		2,868		2,868	

Notes: ***, **, * denote statistical significance at the 1%, 5% and 10% level with standard errors, clustered by age (73 clusters), in parentheses. All specifications estimated using WLS with CCES sampling weights. When age is the main dependent variable, it is divided by its standard deviation to standardize coefficients.

First, we show theoretically in Section 5.1 that endogenizing media exposure strengthens our results. Intuitively, this occurs because more overconfident citizens neglect correlation to a greater degree, and thus believe that additional signals are more valuable. They will thus consume more media, becoming more overconfident, and more extreme.

Second, we also examine our results using another proxy for the number of signals: age. Age is not a choice, nor is it likely to be affected by one's overconfidence or ideology. Moreover, as we have already made parametric restrictions on the basis of Figure 1, Proposition 1 now gives testable predictions for the relationships between age, overconfidence, ideology, and extremeness. In particular, the patterns with respect to age should be the same as those with respect to media. These predictions are tested, and shown to hold, in Table 2. Moreover, the regressions show that the patterns with respect to age are robust to controlling for media, and vice-versa.

To summarize, the patterns in the data are inconsistent with any fully Bayesian model, but are consistent with our theory. Moreover, to draw additional, testable predictions, we

use the implications of these patterns as assumptions in some of what follows. When doing so, we state it explicitly.

3.2 Ideological Extremeness

Our primary result is:

Proposition 2. *Overconfidence and ideological extremeness are positively correlated. This is true conditional on n , and independent of n if ρ is large enough.*²²

To build the intuition of this result, it is useful to recast our model as one in which citizens receive only a single signal, but overestimate its precision. Specifically, we can model each citizen as if they have a single experience $e_i = x + \varepsilon_i$, where $\varepsilon_i \sim \mathcal{N}[0, \kappa]$, $\forall i$. However, citizens overestimate the precision of this signal: that is, they believe that $\varepsilon_i \sim \mathcal{N}[0, \kappa_i]$, where $\kappa_i \geq \kappa$. If we properly define e_i , κ and κ_i , then this “model” will give the same results when there is no heterogeneity in n .

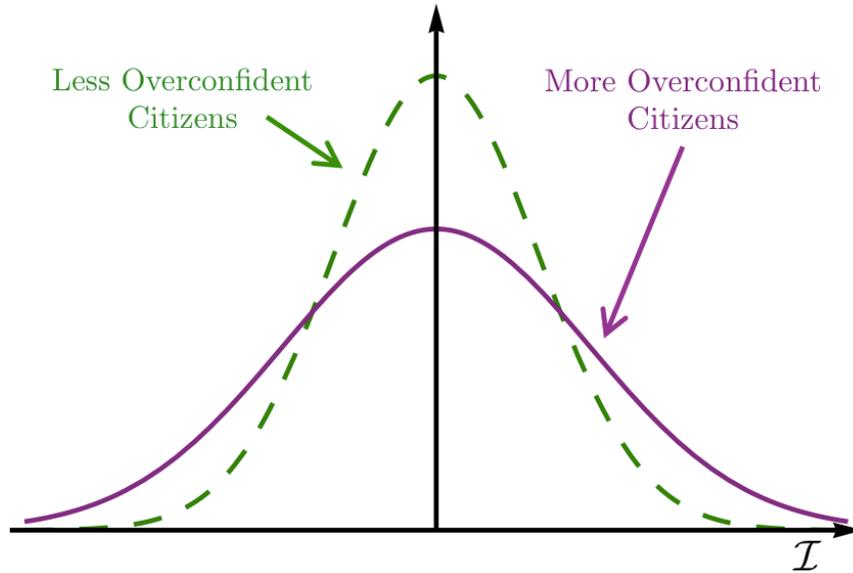
Lemma 3. *Define $e_i \equiv \frac{1}{n_i} \sum_{t=1}^{n_i} e_{it}$. Then $\kappa = \frac{n_i}{1 + (n_i - 1)\rho}$, and $\kappa_i = \frac{n_i}{1 + (n_i - 1)\rho_i}$.*

Fix n and consider two citizens with the same preference bias $b = 0$ and the same experience $e \geq 0$, but two different levels of overconfidence κ_1 and κ_2 , with $\kappa_1 > \kappa_2$. Using the definition of ideology in (1) and Bayes’ rule: $\mathcal{I}_i = b_i + \mathbb{E}_i[x] = \frac{\kappa_i e}{\tau + \kappa_i}$, where \mathbb{E}_i is the expectation over citizen i ’s beliefs. As citizens’ mean beliefs, and hence ideology, are increasing in κ_i , then the more overconfident citizen will have a more extreme ideology. Intuitively, the more overconfident citizen believes her experience is a better signal of the state, and hence updates more, becoming more extreme.

To see that this results in a positive correlation, we examine the entire distribution of ideologies. The logic above implies that the distribution of ideologies for those who are more overconfident will be more spread out than the distribution for those who are less overconfident. Figure 2 shows the distribution of ideologies for two levels of overconfidence

²²Specifically, this holds as long as ρ is large enough that population variance of ideology, conditional on n , $\text{Var}[\mathcal{I}_i|n]$, is increasing in n —see Proposition 1.

Figure 2: Overconfidence and Ideological Extremeness are Correlated



with $x = 0$. In that figure, as one moves further from the ideological center, citizens are more likely to be more overconfident, generating a positive correlation between overconfidence and ideological extremeness. The simplicity of the figure is driven by the assumption that $x = 0$: if $x \neq 0$, the distributions will not be neatly stacked on top of each other, and the relationship will be more complex—but Proposition 2 shows that there is a positive correlation between overconfidence and ideological extremeness for any value of x .²³ Using Implication 1 we can also derive an additional prediction:

Proposition 4. *If $x > 0$ overconfidence and ideology are positively correlated, both independent of, and conditional on, n .*

Proposition 4 follows directly from the discussion above. As the full distribution of e_i is unbiased, $\mathbb{E}[\mathcal{I}_i | \kappa_i] = \frac{\kappa_i x}{1 + \kappa_i}$. When $x > 0$, this is increasing in κ_i .

3.2.1 Empirical Analysis

As political economy theories generally view ideology as a function of wealth or income, we first control for the effect of economic variables—as listed in Section 2.2.3—on ideology. We

²³The proof of Proposition 2 relies on the fact that both overconfidence and extremeness are increasing in correlational neglect. Any distribution of noise that has this property will produce the same results. For more discussion, see Appendix D.

Table 3: Overconfidence is robustly related to ideology and extremeness.

Dependent Variable:	Ideology			Ideological Extremeness Purged of Economic Controls		
	Overconfidence	0.22*** (.028)	0.22*** (.023)	0.20*** (.024)	0.23*** (.028)	0.17*** (.027)
Economic Controls		Y	Y		Y	Y
Number of Signals			Y			Y
R^2	0.047	0.16	0.23	0.054	0.19	0.29
N				2,868		

Notes: ***, **, * denote statistical significance at the 1%, 5% and 10% level with standard errors, clustered by age (73 clusters), in parentheses. All specifications estimated using WLS with CCES sampling weights.

then take the absolute value of the residuals from these regressions as measures of ideological extremeness. All three measures of ideology and ideological extremeness are divided by their standard error to standardize regression coefficients. The first three columns of Table 3 show both an empirical examination of Proposition 4, and the regression, in the second column, used to construct the ideological extremeness measure.

We now examine our primary prediction—Proposition 2—by regressing the resultant measure of ideological extremeness on overconfidence. The relationship between ideological extremeness and overconfidence is statistically very robust—with t-statistics on this novel result between ~ 5 and ~ 10 .²⁴

While we have shown that the relationship between ideological extremeness or ideology and overconfidence is statistically robust, is it substantively important? Table 4 suggests the answer is yes. Specifically, it shows the change in ideological extremeness, and ideology, associated with a one-standard-deviation change in the various economic controls. As the table shows, overconfidence is as predictive of ideological extremeness as income, education, and stock ownership, and more predictive than union membership, home ownership or age. Moreover, as this relationship is more consistent across specifications, it suggests that overconfidence is a separate phenomenon that is not captured by standard controls. A

²⁴The closest empirical result we are aware of appears in Footnote 14 of Kuklinski et al. (2000), which notes a strong correlation (0.34) between strength of partisan identification and confidence in incorrect opinions.

Table 4: Overconfidence is a substantively important predictor of ideology and ideological extremeness.

A one standard deviation change in _____ is associated with:	a _____ standard deviation change in ideology.		a _____ standard deviation change in extremeness.	
	Minimum	Maximum	Minimum	Maximum
Income	6%	8%	< 1%	28%
Education	10%	20%	2%	25%
Union Member	< 1%	4%	< 1%	8%
Home Owner	3%	17%	< 1%	16%
Stock Owner	6%	14%	< 1%	27%
Overconfidence	17%	25%	12%	28%

Notes: The minimum and maximum effect size come from regressions with no other variables, and all other variables, respectively, across the three different measures of ideology and extremeness. These specifications can be found in Appendix D. For the controls, we consider extremeness measures both that have, and have not, been purged of the economic effect on ideology. Effect sizes for categorical variables are based on entering them linearly in regressions.

similar pattern emerges for left-right ideology. It is worth noting that income and education are fairly stable predictors of left-right ideology, although they are not as substantively important as overconfidence.

3.3 Differences between Left and Right

We now use Implication 1 ($x > 0$) to draw additional, subtle, predictions from the model.

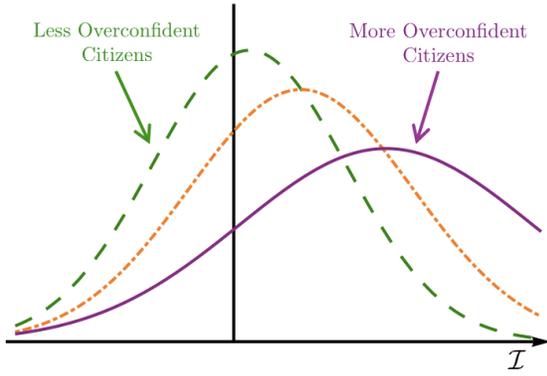
Proposition 5. *If $x > 0$ then $\text{Cov}[\mathcal{E}, \kappa_i | \mathcal{I}_i \geq 0] > \text{Cov}[\mathcal{E}, \kappa_i | \mathcal{I}_i \leq 0]$ both independent of, and conditional on, n .*

Proposition 6. *If $x > 0$ and ρ is large enough then $\text{Cov}[\mathcal{E}, n_i | \mathcal{I}_i \geq 0] > \text{Cov}[\mathcal{E}, n_i | \mathcal{I}_i \leq 0]$.*²⁵

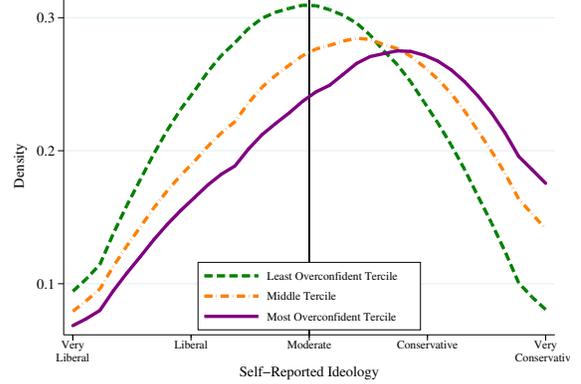
Proposition 5 states that if $x > 0$, then the covariance between overconfidence and extremeness is larger for those to the right-of-center than for those to the left-of-center. The mathematical intuition is illustrated in Figure 3(a), which uses three different levels of κ_i . Moving right from the center, average overconfidence is quickly increasing, along with ideological extremeness. This leads to a large covariance between overconfidence and

²⁵Simulations indicate this holds for all $\rho > \rho_i$. ρ close to one is needed for tractability.

Figure 3: The theoretical structure of Proposition 5, and the data used to examine it.



(a) Theory: When average ideology is increasing in overconfidence.



(b) Data: Distribution of self-reported ideology by tercile of overconfidence. (Smoothed using an Epanechnikov kernel, bandwidth 0.8.)

ideological extremeness. Moving left, ideological extremeness is also increasing, but average overconfidence initially decreases. Eventually, average overconfidence will increase, but this occurs in a region that contains a relatively small measure of citizens. Thus, the covariance to the left will be either small and negative or small and positive, depending on the relative measure of citizens in the regions with positive and negative covariances. Either way, the covariance between overconfidence and ideological extremeness will be smaller for left-of-center citizens than right-of-center citizens.

A similar logic underlies Proposition 6: fixing ρ_i , then by Proposition 1 overconfidence is increasing in n . Moreover, mean ideology increases in n —if Implication 1 holds—as does the variance of ideology—if Implication 2 holds—producing similar patterns to Figure 3(a).

3.3.1 Empirical Analysis

Initial support for the first part of Proposition 5 comes from a comparison of Figure 3(a), generated by theory, and Figure 3(b), generated from the data. A statistical analysis is found in the first panel of Table 5, which finds that ideological extremeness has a substantially higher covariance with overconfidence for those to the right of center than for those to the left of center, in accordance with Proposition 5. Proposition 6 also finds support in the statistical examination of Table 5. The final two columns control for age when testing the

Table 5: There is a greater covariance between extremeness and overconfidence for right-of-center citizens than left of center citizens.

Dependent Variable:	Scaled Extremeness (from Ideology, Purged of Economic Controls)					
	Left of Center	Right of Center	Left of Center	Right of Center	Left of Center	Right of Center
Covariance with Overconfidence	0.063** (.031)	0.28*** (.037)	0.012 (.027)	0.18*** (.031)	-0.014 (.024)	0.11*** (.024)
Difference		0.22*** (.049)		0.16*** (.041)		0.13*** (.034)
Economic Controls			Y	Y	Y	Y
Number of Signals					Y	Y
Covariance with Media Exposure	0.18*** (.031)	0.33*** (.040)	0.095*** (.023)	0.20*** (.037)	0.084*** (.020)	0.16*** (.035)
Difference		0.14*** (.051)		0.11*** (.044)		0.077** (.040)
Economic Controls			Y	Y	Y	Y
Age (73 Categories)					Y	Y
Covariance with Age	0.10*** (.030)	0.21*** (.049)	0.065*** (.024)	0.16*** (.037)	0.067*** (.022)	0.14*** (.036)
Difference		0.11** (.057)		0.098** (.044)		0.075** (.042)
Economic Controls			Y	Y	Y	Y
Media (16 Categories)					Y	Y
N	1,123	1,745	1,123	1,745	1,123	1,745

Notes: ***, **, * denote statistical significance at the 1%, 5% and 10% level with standard errors, clustered by age (73 clusters), in parentheses. The Frisch-Waugh-Lovell Theorem is used to compute conditional covariances. Age is standardized in these regressions. Similar results hold using partial correlations.

proposition using media, and control for media when testing the proposition using age. This emphasizes that although patterns in ideology with respect to media exposure and age may look similar, they are driven by distinct underlying variation.

It should be noted that while specular results hold if $x < 0$, these examinations are still tests of our model. In particular, the pattern in Figure 3(a) is by Implication 1 ($x > 0$), derived from the fact that ideology is increasing with media exposure. Indeed, it may be surprising that the data matches the implication of the theory as closely as it does in Figure

3. Propositions 5 and 6 provide a way to state, and test, this relationship statistically.

These patterns emphasize that our results are not driven by the relationship between overconfidence and ideology. Specifically, if overconfidence lead directly to conservatism, there should be a negative relationship between extremism and overconfidence for those left of center.²⁶ No such pattern exists in Table 5. Moreover, given the relationship between overconfidence and the number of signals in Table 2 (that is, $d\kappa_i/dn_i > 0$), there should be a negative relationship between extremism and the number of signals for those left of center. This, too, does not find support. This implies that the mechanism we have identified applies to both liberals and conservatives. This point is discussed further in Appendix C.

3.4 Covariances and the Number of Signals

Previous subsections have examined various elements of our theory: signals (age and media), correlational neglect (overconfidence), and ideological extremeness in a pairwise fashion. Our final proposition brings these elements together.

Proposition 7. *If ρ is large enough, then $\text{Cov}[\mathcal{E}, \kappa_i - \kappa|n_i]$ is increasing in n_i .*

This proposition holds because the proportion of extremism and overconfidence that is due to signals—as opposed to priors or preference biases—is increasing in n_i . Importantly, as examining this proposition relies on grouping together citizens who all have the same level of media exposure, whatever causes that particular level of exposure is held constant.²⁷

3.4.1 Empirical Analysis

We begin by breaking our sample into quartiles by level of media exposure, and then calculating the covariance between ideological extremeness and overconfidence for each quartile. The patterns in Table 6 are broadly supportive of Proposition 7. However, testing the proposition pushes the limits of our data—most differences are not statistically significant. Indeed, as the standard errors on the differences between quartiles are approximately 0.06, this means

²⁶Formally, this would be modeled as $\mathcal{I}_i = g(\kappa_i)$ with $g' > 0$.

²⁷Note this is not an issue when age is used as a measure of the number of signals.

Table 6: The correlation and covariance between overconfidence and ideological extremeness are increasing in n .

Quartile of Media Exposure:	Lowest	2nd Lowest	2nd Highest	Highest
Corr[$\mathcal{E}, \kappa_i - \kappa n$]	0.13	0.20	0.21	0.30
Cov[$\mathcal{E}, \kappa_i - \kappa n$]	0.13** (.063)	0.18*** (.045)	0.18*** (.037)	0.27*** (.046)
Inter-Quartile Difference	0.048 (.076)	0.0044 (.058)	0.086* (.059)	
Two Quartile Difference		0.052 (.073)	0.091* (.064)	
Three Quartile Difference			0.14** (.078)	
N	512	719	1169	468

Notes: ***, **, * denote statistical significance at the 1%, 5% and 10% level using a one-tailed test for differences, with standard errors, clustered by age (73 clusters), in parentheses. All specifications estimated using WLS with CCES sampling weights. Unequal quartile sizes come from the use of sampling weights and the lumpiness of the media exposure measure.

that differences of covariances must be about 0.1 to obtain statistical significance. This is close to the maximum difference in covariances between sub-groups.

Including controls gives a difference between the first and fourth quartile of 0.12 (s.e. = .064, $p < 0.05$). Dividing the sample quartiles in age produces a difference of 0.13 (s.e. = .095, $p < 0.1$) without controls, and 0.085 (s.e. = .073, $p = 0.12$) with controls.

4 Turnout and Partisan Identification

We now turn to analyze different dependent variables: voter turnout and partisan identification. To analyze these behaviors, we must first specify how citizens make these political choices. Specifically, we posit an expressive voter model in which the expressive value of voting is increasing with a citizen's belief that one party's policy is better for her, and then

move to examine the implications of this model theoretically and empirically.

4.1 Formalization

Turnout and partisan identification will depend on the policy positions adopted by parties. We assume that there are two parties committed to platforms L and R , with $L = -R$.²⁸ Denote by $U_j(b_i|x)$ the utility that a citizen with preference bias b_i receives from the platform of party j when the state is x . Party R 's position will be better for citizen i in state x when $U_R(b_i|x) > U_L(b_i|x)$. As in the above description, we assume citizen i turns out to vote if and only if

$$\left| \text{Prob}_i[U_R(b_i|x) > U_L(b_i|x)] - \frac{1}{2} \right| - c_i > 0. \quad (2)$$

where $c_i \sim F_c$ is the idiosyncratic cost of turning out to vote. We assume F_c strictly increasing on $(0, \frac{1}{2})$, and $c_i \perp (b_i, \rho_i, e_{it})$. Appendix D shows that (2) produces the same comparative statics as the canonical voting model of Riker and Ordeshook (1968) with a large electorate, and regret- or choice-avoidant voters (Matsusaka, 1995; Degan and Merlo, 2011).²⁹

Finally, we model strength of partisan identification using the left-hand side of (2), but with a (possibly different) distribution of costs F'_c .³⁰

4.2 Predictions

This model of turnout gives several predictions:

Proposition 8. *Conditional on n :*

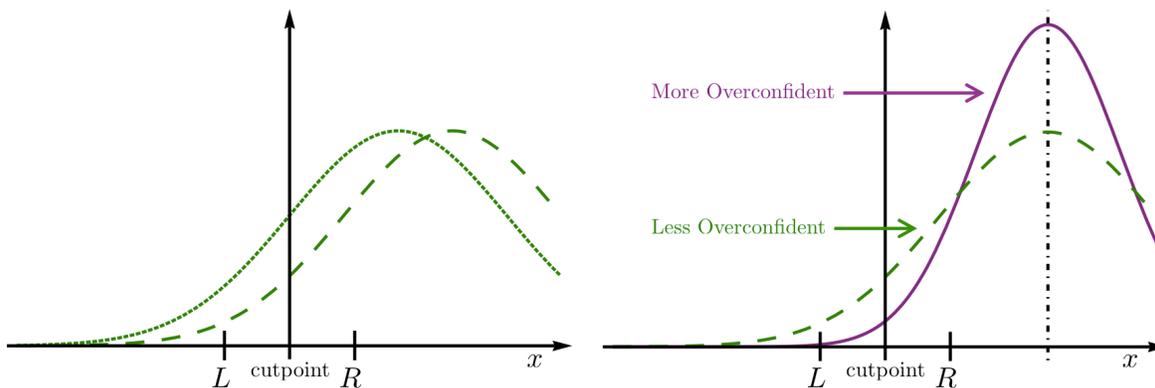
1. *More ideologically extreme citizens are more likely to turn out to vote.*
2. *Conditional on overconfidence, more ideologically extreme citizens are more likely to turn out.*

²⁸Symmetric divergence can be generated from a Calvert (1985) model with policy and office motivated parties that are uncertain about the median voter's ideology due to the random realization of x .

²⁹Note that Riker and Ordeshook (1968) contains both a pivotal and expressive component. In large elections the expressive component dominates. It is straightforward to see that (2) is consistent with regret-avoidance. As discussed in Degan and Merlo (2011), it is also consistent with choice-avoidance if voters will never learn whether their choice is correct or not. For a deeper discussion of these points, see Appendix D.

³⁰We adopt this formulation to simplify and shorten the exposition. Identical predictions are obtained from a more complex model of partisan identification discussed in Appendix D.

Figure 4: Intuition for Proposition 8 and Corollary 9.



(a) More ideologically extreme citizens are more likely to turn out.

(b) More overconfident citizens are more likely to turn out, conditional on ideology

3. *Conditional on ideology, more overconfident citizens are more likely to turn out.*

If ρ is large then these predictions also hold independent of n .

The first part of Proposition 8 is a well-documented empirical regularity: more ideologically extreme citizens are more likely to turn out. The second part of Proposition 8 makes a stronger prediction: more ideologically extreme citizens are more likely to turn out, even controlling for overconfidence. Figure 4(a) helps build intuition. It depicts the posterior of two citizens with the same level of overconfidence, but different ideologies. While both prefer R to L , the more extreme citizen assigns a higher probability to R having the correct policy, and hence is more likely to turn out.

The third part of Proposition 8 describes the role of overconfidence in turnout: more overconfident citizens are more likely to turn out, even controlling for ideology. The intuition is apparent from Figure 4(b), which shows the posterior of two citizens, both with $b = 0$ and the same posterior mean beliefs $\mathbb{E}_i[x]$, but different levels of overconfidence. While both prefer R to L , the more overconfident citizen assigns a higher probability to R having the correct policy, and hence, is more likely to turn out.

Note that this provides an alternative explanation as to why right-leaning and older people are more likely to vote: because they are more overconfident. This contrasts with explanations in the literature that attribute these patterns to increased income changing the

cost or benefits of voting, or even feelings of increased patriotism among those groups.³¹

The final predictions examined in the survey data concern the strength of partisan identification. These results follow directly from Proposition 8, as (2) characterizes both turnout and partisan identification.

Corollary 9. *Conditional on n , strength of partisan identification is increasing in overconfidence, both conditional on, and independent of, ideological extremeness. Moreover, conditional on overconfidence, strength of partisan identification is increasing in ideological extremeness. If ρ is large enough, these results hold independent of n .*

4.3 Empirical Analysis

We examine Proposition 8 using verified voter turnout from the 2010 CCES.³² The results, shown in Table 7, are supportive of the proposition: Columns 3 and 4 show that more ideologically extreme citizens are more likely to vote, even conditional on overconfidence; and more overconfident citizens are more likely to vote, even conditional on ideological extremeness.³³ Moreover, in Columns 5–7, we show that these patterns hold even conditioning on the number of signals, n . However, when the economic controls are added to all of the theoretical controls, in Column 9, the result remains positive, but loses statistical significance.

To get a full accounting of the effect of overconfidence on turnout, we need to first account for the fact that overconfidence also leads to ideological extremeness. Doing so, a one-standard deviation increase in overconfidence is associated with a 15–19% (depending on the specification) increase in turnout—a 7.5–9.5 percentage point increase versus a baseline turnout rate of 51% in the data. This effect is substantively important as it is larger than the effect of income, education, union membership, and over half of the effect size associated with ideological extremeness and age—all known to be important correlates of turnout.

³¹We thank an anonymous referee for pointing this out.

³²One of the advantages of the CCES dataset is that it provides verified voter turnout in addition to self-reported turnout, which is known to be unreliable. Our results also hold, and indeed are stronger, if we use self-reported turnout.

³³Ideological extremeness here is not purged of economic effects, as our theoretical results are stated conditional on extremeness, not extremeness without wealth or income effects. Using the purged measure produces nearly identical results.

Table 7: Turnout is increasing with ideological extremeness and overconfidence, as predicted by Proposition 8.

Dependent Variable:	Turnout Decision								
	0.096*** (.017)	0.60*** (.015)	0.056*** (.017)	0.035** (.015)	0.038** (.016)	0.035** (.016)	0.026* (.016)	0.026* (.016)	0.013 (.015)
Overconfidence									
Ideological Extremeness			0.18*** (.014)	0.16*** (.014)	0.17*** (.016)	0.16*** (.015)	0.14*** (.016)		0.13*** (.015)
Economic Controls		Y		Y				Y	Y
Media (16 Categories)					Y		Y	Y	Y
Age (73 Categories)						Y	Y	Y	Y
R^2	0.037	0.19	0.16	0.26	0.20	0.27	0.29	0.31	0.36
N	2,808								

Notes: ***, **, * denote statistical significance at the 1%, 5% and 10% level with standard errors, clustered by age (73 clusters), in parentheses. All specifications estimated using WLS with CCES sampling weights.

Table 8: Overconfidence is correlated with strength of partisan identification, even controlling for ideological extremeness.

Treatment of Independents:	Weak (0)			Missing (.)			Strong (1)		
	0.052*** (.013)	0.051*** (.013)	0.030** (.013)	0.060*** (.015)	0.063*** (.013)	0.039*** (.013)	0.050*** (.016)	0.058*** (.014)	0.038** (.013)
Overconfidence									
Ideological Extremeness			0.12*** (.015)			0.11*** (.014)			0.066*** (.012)
Economic Controls		Y	Y		Y	Y		Y	Y
Number of Signals			Y			Y			Y
R^2	0.012	0.084	0.18	0.015	0.097	0.19	0.0098	0.073	0.15
N	2,868			2,545			2,868		

Notes: ***, **, * denote statistical significance at the 1%, 5% and 10% level with standard errors, clustered by age (73 clusters), in parentheses. All specifications estimated using WLS with CCES sampling weights.

We now examine partisan identification. As noted in Section 2.2.2, we construct three measures of partisan identification, all of which code someone who identifies as a “Strong Democrat” or “Strong Republican” as a strong partisan identifier (1), and most others as weak partisan identifiers (0). The three measures differ in how they treat those who identify as “independent”. The first matches theory and codes independents as weak partisan identifiers (0). However, it has been suggested that independents may also hold strongly to that identity, so we show that our results are robust to this possibility, by dropping these respondents (.), or coding them as strong partisan identifiers (1).

Table 8 then regresses these measures on overconfidence, ideological extremeness, economic controls, and controls for the number of signals. The results are consistent with theory, no matter which measure is used. Doing the same accounting exercise as above, a one standard-deviation change in overconfidence is associated with a 9–12% increase in the probability a respondent classifies themselves as strongly partisan—a 4.5–6 percentage point increase from a mean rate of 54%, 44% and 49%, respectively, for the three different measures. This is 48–95% of the effect size associated with ideological extremeness. Note again that the controls that strongly affect the results are those that have a theoretical role—ideological extremeness, age, and media exposure.

One other pattern in Table 8 is worth noting: ideological extremeness is a worse predictor of strength of partisan identification when independents are treated as strong partisan identifiers. Intuitively, there are few respondents who hold extremely conservative or liberal views, but identify as independent.

5 Other Sources of Information

In this section we examine how relaxing our assumptions about the sources of information would affect our primary results. First, we consider the case in which citizens choose the number of costly signals to acquire. Second, we consider the effects of citizens sharing their ideologies with each other. In both cases we find that these extensions strengthen our results.

5.1 Endogenous Information Acquisition

In standard models, risk-averse citizens acquire information to reduce their uncertainty about the state x . In these models, more confident citizens would demand fewer signals—in our context, less media. However, in our model more overconfident citizens acquire *more* information. Moreover, endogenizing information acquisition strengthens our primary result.

To formalize, define the cost of acquiring a signal as c , and a citizen’s perceived optimal number of signals as n_i^* . Further, define $V_i(\kappa_i|n_i)$ as the value to citizen i , with overconfidence κ_i , of an additional signal, given that she has already received n_i signals. Then:

Proposition 10.

1. $V_i(\kappa_i|n_i)$ is increasing in κ_i .
2. n_i^* is increasing in the degree of correlational neglect.

The first part of Proposition 10 may seem counter-intuitive: why would someone who believes they are more certain of the state place a greater value on the additional information? The intuition comes from the second part: as more overconfident citizens neglect correlation to a greater degree, they believe additional signals have more information, and thus value.

The second part of Proposition 10 is consistent with the first panel of Figure 1: media exposure is increasing in overconfidence. It also has an additional implication: endogenizing media exposure reinforces the relationship between overconfidence and extremeness found in Proposition 2. This occurs because those with a greater degree of correlational neglect are now more overconfident and more extreme for two reasons: correlational neglect and increased media consumption.

5.2 Communication between Citizens

What if citizens could learn the point of view of citizens outside their network, or receive information from public sources? In this subsection, we show theoretically that this would, interestingly, strengthen the correlation between overconfidence and ideological extremeness.

This occurs because when more overconfident citizens meet someone with a different ideology, they attribute this difference to factors other than the information possessed by the other citizen—as, by construction, they believe that “they know better”. Therefore, more overconfident citizens will tend to update less than less overconfident citizens, making more overconfident citizens relatively more extreme.

We illustrate this pattern in two ways. First we consider citizens with arbitrary preference biases, b_i , who are unaware that other citizens may be overconfident. Second, citizens are aware that others may be overconfident, but there are no preference biases ($b_i = 0, \forall i$). In the first case, citizens will attribute disagreement to the bias of others; in the second, they will attribute it to others’ overconfidence. More overconfident citizens will attribute more of the difference to these other factors.

Throughout this section, we assume that after n private signals, each citizen i meets another, randomly chosen, citizen j and is told her ideology. It is straightforward to extend the analysis to citizens meeting any finite number of other citizens, or observing any finite number of public signals with known precision.³⁴

5.2.1 Unawareness of Overconfidence

As noted above, we begin by assuming citizens are unaware of overconfidence.

Proposition 11. *When citizen i is told the ideology of citizen j , and she believes $\kappa_j = \kappa$:*

1. *The ideology of citizen i after communication is $\alpha_i \mathcal{I}_i + \beta_i \mathcal{I}_j$ for some $\alpha_i, \beta_i \in \mathbb{R}_{++}$, where α_i is increasing in κ_i and β_i is decreasing in κ_i .*
2. *If $\mathcal{I}_j \neq (\mathcal{I}_i - b_i) \frac{\kappa}{\kappa + \tau}$, then i ’s mean belief about the extremeness of j ’s preferences is increasing in i ’s level of overconfidence $\frac{d\mathbb{E}_i[b_j]}{d\kappa_i} > 0$.*

When i meets j , she knows that the difference in their ideologies may have two sources: different preference biases and different information. The more overconfident citizen i is, the more confident she is that she and j received similar signals. Thus, she believes their

³⁴Matching with like-minded individuals is encompassed by correlational neglect. If there is uncertainty about the distribution of overconfidence in the population, or the mean preference bias in the population, our results extend to public signals regarding the distribution of ideology.

difference in ideologies is due to differences in preference biases. In turn, this leads i to only slightly update her beliefs.

This intuition also characterizes how overconfident citizens would update in the face of media reports contradicting their point of view. As long as there is some chance that the media is biased, more overconfident citizens will attribute the contradiction to media bias, and, hence, update less.

5.2.2 No Preference Biases

Next, we consider the case in which citizens are (correctly) aware of the fact that others are overconfident. For simplicity, we assume that all citizens have no preference bias ($b_i = 0, \forall i$), and that this is common knowledge. Define F_{κ_i} as the distribution of posterior precisions in the population, and $\underline{\kappa} = \inf\{\kappa | F_{\kappa_i}(\kappa) > 0\}$, then:

Proposition 12. *Suppose $b_i = 0, \forall i$. When citizen i is told the ideology of citizen j :*

1. *The ideology of citizen i after communication is $\gamma_i \mathcal{I}_i + \delta_i \mathcal{I}_j$ for some $\gamma_i, \delta_i \in \mathbb{R}_{++}$, where γ_i is increasing in κ_i and δ_i is decreasing in κ_i .*
2. *$\mathbb{E}_i[\kappa_j]$ is increasing in κ_i if i and j are on opposite sides of the aisle, ($\mathcal{I}_i * \mathcal{I}_j < 0$) or if j is more ideological extreme than i ($\mathcal{E}_j > \mathcal{E}_i$).*
3. *$\mathbb{E}_i[\kappa_j]$ is decreasing in κ_i if i and j are on the same side of the aisle ($\mathcal{I}_i * \mathcal{I}_j > 0$), and $\mathcal{E}_i > \frac{\tau + \underline{\kappa}}{\underline{\kappa}} \mathcal{E}_j$.*

Proposition 12 has a similar form, and intuition, to Proposition 11. When a citizen meets someone with a different ideology, she can attribute the difference to either differences in information, or in how the other citizen processes information. Following the logic above, more overconfident citizens attribute more of the difference to other citizens' overconfidence.

However, the other parts of Proposition 12 are more nuanced. In particular, if the other citizen is more extreme, or is on the other side of the aisle, the first citizen attributes this to overconfidence. But when the other citizen is on the same side of the aisle but is less extreme, the first citizen believes that the other under-interprets her information, that is, she "lacks the courage of her convictions".

Proposition 11 and 12 both imply that communication causes more overconfident citizens to have *relatively* more dispersed ideologies. This leads to a greater correlation between overconfidence and ideological extremeness.

Finally, these results allow us to briefly consider a puzzle presented in the introduction: why political rumors and misinformation are so persistent. Our model suggests a possible answer: it is very difficult to persuade overconfident citizens that their prior is incorrect, as they tend to attribute contradictory information to others' biases.

6 Discussion: Identification and Future Directions

We conclude with a summary of our major results, and then turn to a discussion of identification and directions for future research.

This paper introduces a model of correlational neglect leading to overconfidence, and draws implications for political behavior. In particular, the model predicts that overconfidence and extremism are positively correlated, that both overconfidence and ideological extremism are independently correlated with voter turnout, that overconfidence is increasing with the number of signals—that is, age and media exposure—and that, moreover, the correlation between ideology and overconfidence is increasing in the number of signals. Taking into account the findings of Section 3.1, the model makes additional predictions: ideology and overconfidence are positively correlated, the covariance between extremism and overconfidence is greater for those right-of-center than left-of-center, and that the covariance between media exposure and the number of signals is greater for those right of center than left of center. These implications are examined using unique survey data. All find support in this data, most at very high levels of statistical significance, and when controlling for the number of signals and all available economic factors.

Table 9: A general knowledge-based measure of overconfidence produces the same results.

Dependent Variable:	Self-Reported Ideological Extremeness (“Don’t Know” treated as centrist)			
Overconfidence (Economy)	0.16*** (.047)	0.14*** (.040)		
Overconfidence (General Knowledge)			0.17*** (.042)	0.13** (.050)
Economic Controls		Y		Y
Number of Signals		Y		Y
R^2	0.024	0.30	0.029	0.29
N			989	

Notes: ***, **, * denote statistical significance at the 1%, 5% and 10% level with standard errors, clustered by age (69 clusters), in parentheses.

6.1 Identification

For our results to be identified, correlational neglect must be something akin to a personality trait: set early in life, with changes unrelated to political conditions. While this is plausible, it is not testable with our data. However, we can gain deeper insight by considering two classes of threats to identification: reverse causality and third-factor causation.

One might object that the factual questions used to measure overconfidence are inherently ideological, and thus extremeness causes confident responses. While Ansolabehere et al. (2011) do not find partisan differences in factual answers, we have also examined other ways of eliciting overconfidence. Specifically, we were allowed to place a few questions on the 2011 CCES that would measure overconfidence on general knowledge items, such as the year of Shakespeare’s birth and the population of Spain. Moreover, confidence was elicited using a confidence interval, similar to the method used in the psychology literature. While the 2011 survey is limited in other ways—it was much shorter and smaller, only allowed for self-reported ideology, and did not contain voter turnout data—we can use it to examine the central relationship between extremeness and overconfidence.

Table 9 shows that the results are substantively unchanged in the 2011 data, and that the results using the different measures of overconfidence are statistically indistinguishable. We believe this should eliminate concerns that the correlation between extremeness and

overconfidence is driven by the questions we use to measure overconfidence.

However, there may be “something else” causing both ideology and overconfidence: for example, particular patterns of brain or social development. To our knowledge, the literature does not suggest any obvious third factors that would explain all of our empirical findings. If such a third factor is found, it would clearly be very important. Even if that occurs, we believe our results will still provide useful insights into the relationship between overconfidence and political characteristics. Indeed, as correlational neglect *is* a third factor in this sense, it may turn out that this “something else” is a set of mechanisms underlying correlational neglect.

6.2 Future Directions

This returns us to the introduction, where we noted two puzzles, and suggest that a behavioral basis for ideology promises to deepen our understanding of political institutions. The first puzzle was why political polarization has seemed to increase with an increase in access to information. As noted in Section 3.1, our theory provides a potential answer: if additional signals are correlated such that the increase in the number of signals is greater than the increase in information, this will lead to greater polarization. The second puzzle concerned the durability of political rumors and misinformation. As noted at the end of Section 5.2, a related mechanism may be responsible: overconfident citizens will tend to attribute contradictory information to the biases of others rather than to their own misinformation.

Understanding how these patterns interact with institutions must be left to future work, however, we illustrate the usefulness of our findings by sketching a model of primaries with overconfident voters (Ortoleva and Snowberg, 2013b). Two parties have primaries to nominate candidates for executive office. Between the primaries and the general election, nature will send each voter a signal of the state. It is well known that primary voters are more ideologically extreme than the general electorate. Based on the evidence presented above, these voters are also more overconfident. Thus, although primary voters know the ideology of the median voter at the time of the primary, they expect nature’s signals to agree with their beliefs, drawing the median voter toward their ideology. This implies that primary

voters will select candidates on opposite sides of the median voter, expecting nature's signal to pull the median voter towards the more extreme position of the primary voters. Moreover, the losing candidates' partisans will think the median voter ignored "the truth". We believe this sketch provides some insight into the nomination of, and partisan reactions to the defeat of, John Kerry in 2004, and Mitt Romney in 2012.

References

- Acemoglu, Daron and James A. Robinson**, *Economic Origins of Dictatorship and Democracy*, New York: Cambridge University Press, 2006.
- Achen, Christopher H.**, “Social Psychology, Demographic Variables, and Linear Regression: Breaking the Iron Triangle in Voting Research,” *Political Behavior*, 1992, 14 (3), 195–211.
- Alesina, Alberto and George-Marios Angeletos**, “Fairness and Redistribution,” *The American Economic Review*, 2005, 95 (4), 960–980.
- Alpert, M. and H. Raiffa**, “A Progress Report on the Training of Probability Assessors,” in Daniel Kahneman, Paul Slovic, and Amos Tversky, eds., *Judgement Under Uncertainty: Heuristics and Biases*, Cambridge and New York: Cambridge University Press, 1969/1982, pp. 294–305.
- Alvarez, R. Michael**, “Cooperative Congressional Election Study, 2010: Caltech Content,” 2010. [Computer File] Release 1: February 26, 2012. Pasadena, CA: California Institute of Technology [producer] <http://cces.gov.harvard.edu>.
- Ansolabehere, Stephen**, “Cooperative Congressional Election Study, 2010: Harvard/MIT Content,” 2010. [Computer File] Release 1: February 26, 2012. Cambridge, MA: Harvard University [producer] <http://cces.gov.harvard.edu>.
- , “Cooperative Congressional Election Study, 2010: Common Content,” 2010. [Computer File] Release 1: February 26, 2012. Cambridge, MA: Harvard University [producer] <http://cces.gov.harvard.edu>.
- , “Guide to the 2010 Cooperative Congressional Election Survey,” 2012. Available at <http://hdl.handle.net/1902.1/17705>.
- **and Douglas Rivers**, “Cooperative Survey Research,” *Annual Review of Political Science*, 2013, 16, 307–329.
- **and Eitan Hersh**, “The Quality of Voter Registration Records: A State-by-State Analysis,” 2010. Harvard University, *mimeo*.
- , **Marc Meredith, and Erik Snowberg**, “Macro-Economic Voting: Local Information and Micro-Perceptions of the Macro-Economy,” 2011. California Institute of Technology, *mimeo*.
- , – , **and** – , “Asking About Numbers: Why and How,” *Political Analysis*, 2013, 21 (1), 48–69.
- Athey, Susan and Philip A. Haile**, “Nonparametric Approaches to Auctions,” in James J. Heckman and Edward Leamer, eds., *Handbook of Econometrics, Volume 6A*, Elsevier, 2007, pp. 3847–3965.

- Baker, Malcolm and Jeffrey Wurgler**, “Behavioral Corporate Finance: An Updated Survey,” in George M. Constantinides, Milton Harris, and Rene M. Stulz, eds., *Handbook of the Economics of Finance, Volume 2A: Corporate Finance*, North-Holland, 2013, pp. 357–424.
- Benabou, Roland**, “Ideology,” *Journal of the European Economic Association*, 2008, 6 (2–3), 321–352.
- **and Jean Tirole**, “Self-Confidence and Personal Motivation,” *The Quarterly Journal of Economics*, 2002, 117 (3), 871–915.
- **and –**, “Belief in a Just World and Redistributive Politics,” *The Quarterly Journal of Economics*, 2006, 121 (2), 699–746.
- Bendor, Jonathan, Daniel Diermeier, and Michael M. Ting**, “A Behavioral Model of Turnout,” *American Political Science Review*, 2003, 97 (2), 261–280.
- , – , **David Siegel, and Michael M. Ting**, *A Behavioral Theory of Elections*, Princeton, NJ: Princeton University Press, 2011.
- Benoît, Jean-Pierre and Juan Dubra**, “Apparent Overconfidence,” *Econometrica*, 2011, 79 (5), 1591–1625.
- Berinsky, Adam**, “Rumors, Truths, and Reality: A Study of Political Misinformation,” 2012. Massachusetts Institute of Technology, *mimeo*.
- Bertrand, Marianne**, “CEOs,” *Annual Review of Economics*, 2009, 1 (1), 121–150.
- Bisin, Alberto, Alessandro Lizzeri, and Leeat Yariv**, “Government Policy with Time Inconsistent Voters,” 2011. California Institute of Technology, *mimeo*.
- Blomberg, S. Brock and Joseph E. Harrington**, “A Theory of Rigid Extremists and Flexible Moderates with an Application to the US Congress,” *The American Economic Review*, 2000, 90 (3), 605–620.
- Bullock, John G.**, “Partisan Bias and the Bayesian Ideal in the Study of Public Opinion,” *The Journal of Politics*, 2009, 71 (3), 1109–1124.
- Callander, Steven**, “Bandwagons and Momentum in Sequential Voting,” *The Review of Economic Studies*, 2007, 74 (3), 653–684.
- **and Catherine H. Wilson**, “Context-dependent Voting,” *Quarterly Journal of Political Science*, 2006, 1 (3), 227–254.
- **and –**, “Context-dependent Voting and Political Ambiguity,” *Journal of Public Economics*, 2008, 92 (3), 565–581.
- Calvert, Randall L.**, “Robustness of the Multidimensional Voting Model: Candidate Motivations, Uncertainty, and Convergence,” *American Journal of Political Science*, 1985, 29 (1), 69–95.

- Camerer, Colin F. and Dan Lovallo**, “Overconfidence and Excess Entry: An Experimental Approach,” *The American Economic Review*, 1999, 89 (1), 306–318.
- Chandrasekhar, Arun G., Horacio Larreguy, and Juan Pablo Xandri**, “Testing Models of Social Learning on Networks: Evidence from a Framed Field Experiment,” 2012. Massachusetts Institute of Technology, *mimeo*.
- Chen, Xiaohong**, “Large Sample Sieve Estimation of Semi-Nonparametric Models,” in James J. Heckman and Edward Leamer, eds., *Handbook of Econometrics, Vol 6B*, Elsevier, 2007, pp. 5549–5632.
- Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam**, “Investor Psychology and Security Market Under- and Overreactions,” *The Journal of Finance*, 1998, 53 (6), 1839–1885.
- Degan, Arianna**, “Civic Duty and Political Advertising,” *Economic Theory*, 2013, 52 (2), 531–564.
- **and Antonio Merlo**, “A Structural Model of Turnout and Voting in Multiple Elections,” *Journal of the European Economic Association*, 2011, 9 (2), 209–245.
- DeGroot, Morris H.**, “Reaching a Consensus,” *Journal of the American Statistical Association*, 1974, 69 (345), 118–121.
- DeMarzo, Peter M., Dimitri Vayanos, and Jeffrey Zwiebel**, “Persuasion Bias, Social Influence, and Unidimensional Opinions,” *The Quarterly Journal of Economics*, 2003, 118 (3), 909–968.
- Enke, Benjamin and Florian Zimmermann**, “Correlation Neglect in Belief Formation,” 2013. IZA Discussion Paper Number 7372.
- Erev, Ido, Thomas S. Wallsten, and David V. Budescu**, “Simultaneous Over- and Underconfidence: The Role of Error in Judgment Processes,” *Psychological Review*, 1994, 101 (3), 519–527.
- Eyster, Erik and Matthew Rabin**, “Naive Herding in Rich-Information Settings,” *American Economic Journal: Microeconomics*, 2010, 2 (4), 221–243.
- Frank, Thomas**, *What’s the Matter with Kansas?: How Conservatives won the Heart of America*, New York: Metropolitan Books, 2004.
- Gelman, Andrew**, *Red State, Blue State, Rich State, Poor State: Why Americans Vote the Way They Do*, Princeton, New Jersey: Princeton University Press, 2009.
- Gentzkow, Matthew and Jesse M. Shapiro**, “Ideological Segregation Online and Offline,” *The Quarterly Journal of Economics*, 2011, 126 (4), 1799–1839.
- Glaeser, Edward L. and Cass R. Sunstein**, “Extremism and Social Learning,” *Journal of Legal Analysis*, Winter 2009, 1 (1), 263–324.

- Golub, Benjamin and Matthew O. Jackson**, “Naive Learning in Social Networks and the Wisdom of Crowds,” *American Economic Journal: Microeconomics*, 2010, 2 (1), 112–149.
- Henderson, Daniel J, John A List, Daniel L Millimet, Christopher F Parmeter, and Michael K Price**, “Imposing Monotonicity Nonparametrically in First-Price Auctions,” 2009. IDEAS, *mimeo*.
- Kuklinski, James H., Paul J. Quirk, Jennifer Jerit, David Schwieder, and Robert F. Rich**, “Misinformation and the Currency of Democratic Citizenship,” *The Journal of Politics*, 2000, 62 (3), 790–816.
- Levy, Gilat and Ronny Razin**, “Correlation Neglect, Voting Behaviour and Polarization,” 2013. London School of Economics, *mimeo*.
- Lizzeri, Alessandro and Leeat Yariv**, “Collective Self-Control,” 2012. California Institute of Technology, *mimeo*.
- Lundeberg, Mary A., Paul W. Fox, and Judith Punčohař**, “Highly Confident but Wrong: Gender Differences and Similarities in Confidence Judgments,” *Journal of Educational Psychology*, 1994, 86 (1), 114–121.
- Matusaka, John G.**, “Explaining Voter Turnout Patterns: An Information Theory,” *Public Choice*, 1995, 84 (1), 91–117.
- McCarty, Nolan M., Keith T. Poole, and Howard Rosenthal**, *Polarized America: The Dance of Ideology and Unequal Riches*, Cambridge, MA: MIT Press, 2006.
- McMurray, Joseph C.**, “Ideology as Opinion: A Spatial Model of Common-value Elections,” 2013. Brigham Young University, *mimeo*.
- , “Why the Political World is Flat: An Endogenous Left-Right Spectrum in Multidimensional Political Competition,” 2013. Brigham Young University, *mimeo*.
- Meltzer, Allan H. and Scott F. Richard**, “A Rational Theory of the Size of Government,” *Journal of Political Economy*, 1981, 89 (5), 914–927.
- Moore, Don A. and Paul J. Healy**, “The Trouble with Overconfidence,” 2007. Carnegie Mellon University, *mimeo*.
- and —, “The Trouble with Overconfidence,” *Psychological Review*, 2008, 115 (2), 502–517.
- Odean, Terrance**, “Volume, Volatility, Price, and Profit When All Traders are Above Average,” *The Journal of Finance*, 1998, 53 (6), 1887–1934.
- Ortoleva, Pietro and Erik Snowberg**, “Overconfidence in Political Behavior,” 2013. NBER Working Paper Series # 19250.

- and –, “They’ll See! Primaries with Overconfident Voters,” 2013. California Institute of Technology, *mimeo*.
- Passarelli, Francesco and Tabellini, Guido**, “Emotions and Political Unrest,” 2013. *CESifo Working Paper Series*, #4165.
- Piketty, Thomas**, “Social Mobility and Redistributive Politics,” *The Quarterly Journal of Economics*, August 1995, 110 (3), 551–584.
- Poole, Keith T. and Howard Rosenthal**, “A Spatial Model for Legislative Roll Call Analysis,” *American Journal of Political Science*, 1985, 29 (2), 357–384.
- Public Policy Polling**, “Democrats and Republicans Differ on Conspiracy Theory Beliefs,” April 2, 2013. available at http://www.publicpolicypolling.com/pdf/2011/PPP_Release_National_ConspiracyTheories_040213.pdf.
- Riker, William H. and Peter C. Ordeshook**, “A Theory of the Calculus of Voting,” *The American Political Science Review*, 1968, 62 (1), 25–42.
- Santos-Pinto, Luis and Joel Sobel**, “A Model of Positive Self-Image in Subjective Assessments,” *The American Economic Review*, December 2005, 95 (5), 1386–1402.
- Tausanovitch, Chris and Christopher Warshaw**, “Testing Theories of Representation in Congress, State Legislatures and Cities Using a ‘Super Survey’,” *The Journal of Politics*, 2011, 75 (2), 330–342.
- Weingast, Barry R.**, “The Political Foundations of Democracy and the Rule of Law,” *American Political Science Review*, June 1997, 91 (2), 245–263.
- Yale Project on Climate Change Communication**, “Scientific and Public Perspectives on Climate Change,” June 3, 2013. available at http://environment.yale.edu/climate-communication/files/ClimateNote_Consensus_Gap_May2013_FINAL6.pdf.
- Zechman, Martin J.**, “Dynamic Models of the Voter’s Decision Calculus: Incorporating Retrospective Considerations into Rational-Choice Models of Individual Voting Behavior,” *Public Choice*, 1979, 34 (3), 297–315.

Appendix A Proofs—For Online Publication

Schmidt’s Lemma. *If $f(y)$ and $g(y)$ are monotone functions, and $\text{sign}(f'(y)) = \text{sign}(g'(y))$, then $\text{Cov}[f(y), g(y)] > 0$.*

Proof. See Schmidt (2003). ■

Proof of Lemma 3: The posterior likelihood in the model is proportional to

$$\begin{aligned}
 \mathcal{L}(x|e_i) &\propto \mathcal{L}(e_i|x)\mathcal{L}_0(x) \\
 &\propto \exp \left\{ -\frac{1}{2} \begin{pmatrix} x - e_{i1} \\ x - e_{i2} \\ \vdots \\ x - e_{in_i} \end{pmatrix}^T \begin{pmatrix} 1 & \rho_i & \cdots & \rho_i \\ \rho_i & 1 & \cdots & \rho_i \\ \vdots & \vdots & \ddots & \vdots \\ \rho_i & \rho_i & \cdots & 1 \end{pmatrix} \begin{pmatrix} x - e_{i1} \\ x - e_{i2} \\ \vdots \\ x - e_{in_i} \end{pmatrix} \right\} \exp \left\{ -\frac{1}{2}x^2\tau \right\} \\
 &= \exp \left\{ -\frac{1}{2} \left(\frac{nx^2 - 2x \sum_{t=1}^{n_i} e_{it}}{1 + (n_i - 1)\rho_i} + C \right) \right\} \exp \left\{ -\frac{1}{2}x^2\tau \right\} \\
 &\propto \exp \left\{ -\frac{1}{2} \frac{n_i + \tau(1 + (n_i - 1)\rho_i)}{1 + (n_i - 1)\rho_i} \left(x - \frac{\sum_{t=1}^{n_i} e_{it}}{n + \tau(1 + (n_i - 1)\rho_i)} \right)^2 \right\}
 \end{aligned}$$

where C is constant with respect to x . Thus, defining $e_i = \frac{1}{n_i} \sum_{t=1}^{n_i} e_{it}$, the posterior belief of a citizen is distributed according to

$$\mathcal{N} \left[\frac{n_i e_i}{n_i + \tau(1 + (n_i - 1)\rho_i)}, \frac{n_i + \tau(1 + (n_i - 1)\rho_i)}{1 + (n_i - 1)\rho_i} \right]. \quad (3)$$

Substituting $\rho_i = \frac{n_i - \kappa_i}{(n_i - 1)\kappa_i}$ the posterior is given by $\mathcal{N} \left[\frac{\kappa_i e_i}{\kappa_i + \tau}, \kappa_i + \tau \right]$, which is the same as the posterior that a citizen would have if they received a single signal $e_i = x + \varepsilon_i$, where the citizen believes $\varepsilon_i \sim \mathcal{N}[0, \kappa_i]$. Finally, note that $\mathbb{E}[e_i] = x$, and

$$\text{Var}[e_i] = \left(\frac{1}{n_i} \right)^2 \sum_{t=1}^{n_i} \text{Var}[\varepsilon_{it}] + 2 \left(\frac{1}{n_i} \right)^2 \frac{n_i(n_i - 1)}{2} \text{Cov}[\varepsilon_{it}, \varepsilon_{it'}] = \frac{1}{n_i} + \frac{n_i - 1}{n_i} \rho.$$

Thus, $e_i \sim \mathcal{N} \left[x, \frac{n_i}{1 + (n_i - 1)\rho} \right] \equiv \mathcal{N}[x, \kappa]$. ■

Proof of Proposition 1: A citizen’s overconfidence after n_i signals is given by:

$$\frac{n_i + \tau(1 + (n_i - 1)\rho_i)}{1 + (n_i - 1)\rho_i} - \frac{n_i + \tau(1 + (n_i - 1)\rho)}{1 + (n_i - 1)\rho} > 0 \iff \rho_i < \rho$$

The difference in overconfidence between the citizen with $n_i + 1$ and n_i signals is given by

$$\frac{n_i(\rho - \rho_i)(2 + (n_i - 1)(\rho + \rho_i - \rho\rho_i))}{(1 + (n_i - 1)\rho_i)(1 + n_i\rho_i)(1 + (n_i - 1)\rho)(1 + n\rho)} > 0$$

because $0 < \rho_i < \rho < 1$ and $n_i \geq 2$.

The second and third part follow from Lemma 3: using (3) a citizen's mean belief is distributed according to

$$\mathcal{N}\left[\frac{n_i x}{n_i + \tau(1 + (n_i - 1)\rho_i)}, \frac{(n_i + \tau(1 + (n_i - 1)\rho_i))^2}{n_i(1 + (n_i - 1)\rho)}\right]$$

and the difference between the mean of that distribution at $n_i + 1$ and n_i is

$$\frac{x\tau(1 - \rho_i)}{(1 + n_i + \tau(1 + n_i\rho_i))(n_i + \tau(1 + (n_i - 1)\rho_i))}$$

which is positive iff x is positive, and negative iff x is negative.

For the third part, we need to show $\text{Var}[\mathcal{I}_i | n + 1] - \text{Var}[\mathcal{I}_i | n] > 0$. That is:

$$\frac{(n_i + 1)(1 + n_i\rho)}{(n_i + 1 + \tau(1 + n_i\rho_i))^2} - \frac{n_i(1 + (n_i - 1)\rho)}{(n_i + \tau(1 + (n_i - 1)\rho_i))^2} > 0 \quad (4)$$

The LHS of (4) is 0 when:

$$\rho = \frac{n_i(n_i + 1)(1 + \rho_i\tau)^2 - (1 - \rho_i)^2\tau^2}{n_i((n_i + 1)(1 + 2\tau) + \tau^2(2 + \rho_i(n_i - 1)(2 - \rho_i)))}$$

The derivative of the LHS of (4) with respect to ρ is

$$\frac{n_i((n_i + 1)(1 + 2\tau) + \tau^2(2 + \rho_i(n_i - 1)(2 - \rho_i)))}{(n_i + 1 + \tau(1 + n_i\rho_i))^2(n_i + \tau(1 + (n_i - 1)\rho_i))^2} \geq 0$$

Therefore, as long as $\rho \geq \frac{n_i(n_i + 1)(1 + \rho_i\tau)^2 - (1 - \rho_i)^2\tau^2}{n_i((n_i + 1)(1 + 2\tau) + \tau^2(2 + \rho_i(n_i - 1)(2 - \rho_i)))}$, (4) is satisfied. This will hold for all n if $\rho \geq \frac{1 + \rho_i\tau}{1 + \tau(2 - \rho_i)}$, and, thus, for all ρ_i if ρ is large. \blacksquare

Proof of Proposition 2: Fix $n_i = n$. Using (3) and the distribution of e_i we have that the posterior distribution of the mean of beliefs of citizens with a given κ_i is given by

$$\mathbb{E}_i[x|\kappa_i] \sim \mathcal{N}\left[\frac{\kappa_i x}{\tau + \kappa_i}, \kappa \left(\frac{\tau + \kappa_i}{\kappa_i}\right)^2\right],$$

where $\Phi[\cdot]$ denotes the c.d.f. of the standard normal distribution. Using the fact that residuals from an OLS regression are orthogonal to regressors we have:

$$\text{Cov}[\mathcal{E} - \mathbb{E}[\mathcal{E}|\kappa_i], \kappa_i] = 0 \quad \Rightarrow \quad \text{Cov}[\mathcal{E}, \kappa_i] = \text{Cov}[\mathbb{E}[\mathcal{E}|\kappa_i], \kappa_i] \quad (5)$$

Note that $\mathcal{I}_i|\kappa_i = b_i + \mathbb{E}_i[x|\kappa_i]$ is a sum of two independent random normal variables. Thus,

$$\mathcal{I}|\kappa_i \sim \mathcal{N}\left[\mu, \frac{1}{\sigma^2}\right]. \quad (6)$$

$$\text{where } \mu = \frac{\kappa_i x}{\tau + \kappa_i} \quad \text{and} \quad \sigma^2 = \frac{1}{\kappa} \left(\frac{\kappa_i}{\tau + \kappa_i}\right)^2 + \frac{1}{\tau_b} \quad (7)$$

as $\mathcal{E} = |\mathcal{I}|$, $\mathcal{E}|\kappa_i$ is distributed according to a folded normal distribution, so

$$\mathbb{E}[\mathcal{E}|\kappa_i] = 2\sigma \phi\left[\frac{\mu}{\sigma}\right] + \mu \left(1 - 2\Phi\left[-\frac{\mu}{\sigma}\right]\right) \quad (8)$$

where $\phi[\cdot]$ is the standard normal p.d.f. Taking the derivative of (8) yields:

$$\frac{d\mathbb{E}[\mathcal{E}|\kappa_i]}{d\kappa_i} = \frac{\partial\mathbb{E}[\mathcal{E}|\kappa_i]}{\partial\sigma} \cdot \frac{d\sigma}{d\kappa_i} + \frac{\partial\mathbb{E}[\mathcal{E}|\kappa_i]}{\partial\mu} \cdot \frac{d\mu}{d\kappa_i} \quad (9)$$

$$\frac{\partial\mathbb{E}[\mathcal{E}|\kappa_i]}{\partial\sigma} = 2\phi\left[\frac{\mu}{\sigma}\right] > 0 \quad \text{and} \quad \frac{\partial\mathbb{E}[\mathcal{E}|\kappa_i]}{\partial\mu} = 1 - 2\Phi\left[-\frac{\mu}{\sigma}\right] \quad (10)$$

$$\frac{d\sigma}{d\kappa_i} = \frac{1}{2\sigma\sqrt{\kappa}} \frac{d\sigma^2}{d\kappa_i} = \frac{1}{2\sigma\sqrt{\kappa}} \frac{\tau}{(\tau + \kappa_i)^2} > 0 \quad \text{and} \quad \frac{d\mu}{d\kappa_i} = x \frac{\tau}{(\tau + \kappa_i)^2}$$

so clearly the first term in (9) is positive. Note that $\frac{\partial\mathbb{E}[\mathcal{E}|\kappa_i]}{\partial\mu}$ is positive iff $x > 0$, so the second term in (9) is ≥ 0 . This implies that $\mathbb{E}[\mathcal{E}|\kappa_i]$ is increasing in κ_i . Schmidt's Lemma then implies $\text{Cov}[\mathbb{E}[\mathcal{E}|\kappa_i], \kappa_i] > 0$, and so $\text{Corr}[\mathcal{E}, \kappa_i] > 0$.

For the logic in the previous paragraph to hold in the case when $n_i \sim F_n$, $\text{Var}[\mathcal{I}|n]$ must be increasing in n , which is exactly the condition in Implication 2. ■

Proof of Proposition 4: Following the logic of (5) we have $\text{Cov}[\mathcal{I}, \kappa_i] = \text{Cov}[\mathbb{E}[\mathcal{I}|\kappa_i], \kappa_i]$.

$\mathbb{E}[\mathcal{I}|\kappa_i] = \frac{\kappa_i x}{\tau + \kappa_i}$. As this is increasing in κ_i when $x > 0$, Schmidt's Lemma implies that $\text{Corr}[\mathcal{I}, \kappa_i] > 0$. ■

Proof of Proposition 5: Fix $n_i = n$, and assume that $b_i = 0$ for all i . (We shall relax this assumption later.) By the definition of covariance, and as $\mathcal{E} = |\mathcal{I}_i|$

$$\begin{aligned}\text{Cov}[\mathcal{E}_i, \kappa_i | \mathcal{I}_i \geq 0] &= \mathbb{E}[\mathcal{I}_i \kappa_i | \mathcal{I}_i \geq 0] - \mathbb{E}[\mathcal{I}_i | \mathcal{I}_i \geq 0] \mathbb{E}[\kappa_i | \mathcal{I}_i \geq 0] \\ \text{Cov}[\mathcal{E}_i, \kappa_i | \mathcal{I}_i \leq 0] &= \mathbb{E}[-\mathcal{I}_i \kappa_i | \mathcal{I}_i \leq 0] - \mathbb{E}[-\mathcal{I}_i | \mathcal{I}_i \leq 0] \mathbb{E}[\kappa_i | \mathcal{I}_i \leq 0] \\ &= -\mathbb{E}[\mathcal{I}_i \kappa_i | \mathcal{I}_i \leq 0] + \mathbb{E}[\mathcal{I}_i | \mathcal{I}_i \leq 0] \mathbb{E}[\kappa_i | \mathcal{I}_i \leq 0].\end{aligned}$$

So

$$\text{Cov}[\mathcal{E}, \kappa_i | \mathcal{I}_i \geq 0] > \text{Cov}[\mathcal{E}, \kappa_i | \mathcal{I}_i \leq 0] \quad (11)$$

holds if and only if

$$\begin{aligned}\mathbb{E}[\mathcal{I}_i \kappa_i | \mathcal{I}_i \geq 0] - \mathbb{E}[\mathcal{I}_i | \mathcal{I}_i \geq 0] \mathbb{E}[\kappa_i | \mathcal{I}_i \geq 0] + \mathbb{E}[\mathcal{I}_i \kappa_i | \mathcal{I}_i \leq 0] - \mathbb{E}[\mathcal{I}_i | \mathcal{I}_i \leq 0] \mathbb{E}[\kappa_i | \mathcal{I}_i \leq 0] &> 0 \\ \mathbb{E}[\mathcal{I}_i \kappa_i | \mathcal{I}_i \geq 0] - \mathbb{E}[\mathcal{I}_i | \mathcal{I}_i \geq 0] \mathbb{E}[\kappa_i] + \mathbb{E}[\mathcal{I}_i \kappa_i | \mathcal{I}_i \leq 0] - \mathbb{E}[\mathcal{I}_i | \mathcal{I}_i \leq 0] \mathbb{E}[\kappa_i] &> 0 \\ \mathbb{E}[\mathcal{I}_i(\kappa_i - \mathbb{E}[\kappa_i]) | \mathcal{I}_i \geq 0] + \mathbb{E}[\mathcal{I}_i(\kappa_i - \mathbb{E}[\kappa_i]) | \mathcal{I}_i \leq 0] &> 0\end{aligned} \quad (12)$$

Where the second line follows as $\mathbb{E}[\kappa_i | \mathcal{I}_i \geq 0] = \mathbb{E}[\kappa_i | \mathcal{I}_i \leq 0] = \mathbb{E}[\kappa_i]$ due to the fact that the $\text{sign}(\mathcal{I}_i) = \text{sign}(e_i)$ and $e_i \perp \kappa_i$. It then follows that (11) holds iff

$$\begin{aligned}\int_{\underline{\kappa}}^{\infty} \int_0^{\infty} \frac{e_i}{\Pr[\mathcal{I}_i \geq 0]} \frac{\kappa_i}{\kappa_i + \tau} (\kappa_i - \mathbb{E}[\kappa_i]) dF_{e_i} dF_{\kappa} + \int_{\underline{\kappa}}^{\infty} \int_{-\infty}^0 \frac{e_i}{\Pr[\mathcal{I}_i \leq 0]} \frac{\kappa_i}{\kappa_i + \tau} (\kappa_i - \mathbb{E}[\kappa_i]) dF_{e_i} dF_{\kappa} &> 0 \\ \iff \int_{\underline{\kappa}}^{\infty} \frac{\kappa_i}{\kappa_i + \tau} (\kappa_i - \mathbb{E}[\kappa_i]) \left[\int_0^{\infty} \frac{e_i}{\Pr[\mathcal{I}_i \geq 0]} dF_{e_i} + \int_{-\infty}^0 \frac{e_i}{\Pr[\mathcal{I}_i \leq 0]} dF_{e_i} \right] dF_{\kappa} &> 0 \\ \iff \int_{\underline{\kappa}}^{\infty} \frac{\kappa_i}{\kappa_i + \tau} (\kappa_i - \mathbb{E}[\kappa_i]) \left[\mathbb{E}[e_i | e_i \geq 0] + \mathbb{E}[e_i | e_i \leq 0] \right] dF_{\kappa} &> 0 \\ \iff \left[\mathbb{E}[e_i | e_i \geq 0] + \mathbb{E}[e_i | e_i \leq 0] \right] \int_{\underline{\kappa}}^{\infty} \frac{\kappa_i}{\kappa_i + \tau} (\kappa_i - \mathbb{E}[\kappa_i]) dF_{\kappa} &> 0 \\ \iff \left[\mathbb{E}[e_i | e_i \geq 0] + \mathbb{E}[e_i | e_i \leq 0] \right] \text{Cov} \left[\frac{\kappa_i}{\kappa_i + \tau}, \kappa_i \right] &> 0\end{aligned}$$

where the last line follows by inverting the steps used to get to (12). As $\frac{\kappa_i}{\kappa_i + \tau}$ is increasing in κ_i , Schmidt's Lemma gives that $\text{Cov} \left[\frac{\kappa_i}{\kappa_i + \tau}, \kappa_i \right] > 0$. Thus, (11) holds iff

$$\mathbb{E}[e_i|e_i \geq 0] + \mathbb{E}[e_i|e_i \leq 0] > 0. \quad (13)$$

$e_i \sim \mathcal{N}[x, \kappa]$ implies

$$\begin{aligned} x &= \Phi[-x\sqrt{\kappa}]\mathbb{E}[e_i|e_i \leq 0] + (1 - \Phi[-x\sqrt{\kappa}])\mathbb{E}[e_i|e_i \geq 0] \\ \Rightarrow \mathbb{E}[e_i|e_i \geq 0] &= \frac{x - \Phi[-x\sqrt{\kappa}]\mathbb{E}[e_i|e_i \leq 0]}{1 - \Phi[-x\sqrt{\kappa}]} \end{aligned}$$

where $\Phi[\cdot]$ is the standard normal c.d.f. Thus, (13) can be re-written as

$$\frac{(1 - 2\Phi[-x\sqrt{\kappa}])\mathbb{E}[e_i|e_i \leq 0] + x}{1 - \Phi[-x\sqrt{\kappa}]} > 0, \quad (14)$$

which holds as long as the numerator is positive. Note that

$$\mathbb{E}[|e_i|] = x - 2\Phi[-x\sqrt{\kappa}]\mathbb{E}[e_i|e_i \leq 0],$$

and, from using the expectation of the folded normal we have that

$$\mathbb{E}[e_i|e_i \leq 0] = \frac{x\Phi[-x\sqrt{\kappa}] - \frac{1}{\sqrt{\kappa}}\phi[x\sqrt{\kappa}]}{\Phi[-x\sqrt{\kappa}]},$$

where $\phi[\cdot]$ is the standard normal p.d.f. Thus, the numerator of (14) can be re-written as

$$\frac{1}{\Phi[-x\sqrt{\kappa}]} \left(2x\Phi[-x\sqrt{\kappa}] - \frac{\phi[x\sqrt{\kappa}]}{\sqrt{\kappa}} - 2x\Phi[-x\sqrt{\kappa}]^2 + \frac{2\phi[x\sqrt{\kappa}]\Phi[-x\sqrt{\kappa}]}{\sqrt{\kappa}} \right) = \frac{Z(x\sqrt{\kappa})}{\Phi[-x\sqrt{\kappa}]} > 0,$$

which holds iff $Z(x\sqrt{\kappa}) > 0$ for $x \in (0, \infty)$ and $\kappa \in [1, \infty)$. Integration by parts gives:

$$\Phi[-x\sqrt{\kappa}] = \frac{1}{2} - \phi[x\sqrt{\kappa}] \left(x + \frac{x^3}{3} + \frac{x^5}{5 \cdot 3} + \frac{x^7}{7 \cdot 5 \cdot 3} + \dots \right) = \frac{1}{2} - \phi[x\sqrt{\kappa}]q(x)$$

which can be applied to $Z(x\sqrt{\kappa})$ to yield

$$Z(x\sqrt{\kappa}) = \frac{x}{2} - \frac{2}{\sqrt{\kappa}}\phi[x\sqrt{\kappa}]^2 q(x) - 2x\phi[x\sqrt{\kappa}]^2 q(x)^2 = \frac{x}{2} - \frac{q(x)e^{-\kappa x^2}}{\pi} \left(\frac{1}{\sqrt{\kappa}} - xq(x) \right).$$

Note that as $\kappa \in [1, \infty)$, then for a fixed x , $Z(x\sqrt{\kappa})$ is minimized at $\kappa = 1$. Thus, it is sufficient to show that $Z(x\sqrt{1}) = Z(x) > 0, \forall x \in (0, \infty)$. We will now show there is a single inflection point of $Z(x)$, that is $Z''(x^*) = 0$ for a unique value of $x^* \in [0, \infty)$, and then use

this to show that $Z(x) > 0, \forall x \in (0, \infty)$.

Using the fact that $\phi[x]q(x) = \Phi[x] - \frac{1}{2}$, we can re-write

$$\begin{aligned} Z(x) &= \frac{x}{2} - 2\phi[x]\left(\Phi[x] - \frac{1}{2}\right) - 2x\left(\Phi[x] - \frac{1}{2}\right)^2 \\ Z'(x) &= \frac{1}{2} - 2\phi[x]^2 - 2x\phi[x]\left(\Phi[x] - \frac{1}{2}\right) - 2\left(\Phi[x] - \frac{1}{2}\right)^2 \\ Z''(x) &= 2(x^2 - 3)\phi[x]\left(\Phi[x] - \frac{1}{2}\right) + 2x\phi[x]^2 \\ Z'''(x) &= 2\phi[x]\left(x(5 - x^2)\left(\Phi[x] - \frac{1}{2}\right) - (2 + x^2)\phi[x]\right) = 2\phi[x]g(x) \end{aligned}$$

We need bounds on $\Phi[x] - \frac{1}{2}$. As $x > 0$ we have

$$\Phi[x] - \frac{1}{2} = \int_0^x \phi[y] dy > \int_0^x \phi[x] dy = x\phi[x] \quad (15)$$

which implies

$$Z''(x) < 2x(x^2 - 2)\phi[x]^2 < 0 \text{ if } x \in (0, \sqrt{2}]$$

and it is clear that $Z''(x) > 0, \forall x \in [\sqrt{3}, \infty)$. Together this implies that $Z''(x) = 0$ somewhere in $[\sqrt{2}, \sqrt{3}]$. Note from the statement of $Z'''(x)$ above that $\text{sign}(Z'''(x)) = \text{sign}(g(x))$.

Further, applying (15) we have that

$$g''(x) = 10\phi[x] - 8x^2\phi[x] - 6x\left(\Phi[x] - \frac{1}{2}\right) < -14\phi[x]\left(x^2 - \frac{5}{7}\right) < 0 \text{ on } [\sqrt{2}, \sqrt{3}].$$

As $g(x)$ is concave on $[\sqrt{2}, \sqrt{3}]$, this implies that

$$\inf_{x \in [\sqrt{2}, \sqrt{3}]} g(x) = \min \left\{ g(\sqrt{2}), g(\sqrt{3}) \right\}$$

and using (15) we have

$$\begin{aligned} g(\sqrt{2}) &= 3\sqrt{2}\left(\Phi[\sqrt{2}] - \frac{1}{2}\right) - 4\phi[\sqrt{2}] > 2\phi[\sqrt{2}] > 0 \\ g(\sqrt{3}) &= 2\sqrt{3}\left(\Phi[\sqrt{3}] - \frac{1}{2}\right) - 5\phi[\sqrt{3}] > \phi[\sqrt{3}] > 0. \end{aligned}$$

As both are positive, this implies that $Z'''(x) > 0, \forall x \in [\sqrt{2}, \sqrt{3}]$, and this combined with $Z''(x) < 0, \forall x \in [0, \sqrt{2}]$ and $Z''(x) > 0, \forall x \in [\sqrt{3}, \infty)$ implies there is a unique $x^* \in (0, \infty)$ such that $Z''(x^*) = 0$.

There are now two cases: $Z'(x^*) \geq 0$ or $Z'(x^*) < 0$. First, consider $Z'(x^*) \geq 0$. Note that $Z(0) = 0$, $Z'(0) = \frac{\pi-2}{2\pi} > 0$, so if $Z'(x^*) \geq 0$, this implies that $Z'(x) > 0, \forall x \in [0, \infty)$, and thus $Z(x) > 0, \forall x \in (0, \infty)$, as desired.

If, on the other hand, $Z'(x^*) < 0$, then the fact that $\lim_{x \rightarrow \infty} Z'(x) \rightarrow 0$ implies $Z'(x) < 0, \forall x \in [x^*, \infty)$ as $Z''(x) > 0, \forall x \in (x^*, \infty)$. This, coupled with the fact that $\lim_{x \rightarrow \infty} Z(x) \rightarrow 0$, implies that $Z(x^*) > 0$ and $Z(x) > 0, \forall x \in [x^*, \infty)$. In addition, $Z''(x) < 0, \forall x \in [0, x^*)$ implies that $\inf_{x \in [0, x^*]} Z(x) = \min \{Z(0), Z(x^*)\} = Z(0) = 0$. Thus, $Z(x) > 0, \forall x \in (0, \infty)$, as desired.

Thus (13) holds, and so too does (11), when $b_i = 0$ for any n . The fact that the covariance of a random variable that is a sum of random variables is linear in the covariances of the individual random variables means that this will also hold when $b_i \neq 0$ as b_i is independent of κ_i , and independently of n . ■

Proof of Proposition 6: Set $\rho = 1$, which implies that $e_i \perp n_i$. We will use this fact to allow us to map this proposition onto the proof of Proposition 5.

Fix n in that proof to $n' > \max n$. Then each n here maps to a unique ρ_i in the proof of Proposition 5. As nothing in that proof depends on the distribution of ρ_i , the result follows immediately. As this proposition holds for $\rho = 1$, it will continue to hold as long as ρ is close to one. ■

Proof of Proposition 7: Using (5), we have $\text{Cov}[\mathcal{E}, \kappa_i - \kappa | n_i] = \text{Cov}[\kappa_i - \kappa, \mathbb{E}[\mathcal{E} | \kappa_i] | n_i]$. Note that $\kappa_i - \kappa$ and $\mathbb{E}[\mathcal{E} | \kappa_i]$ are both functions of n_i and ρ_i . Then, as ρ_i is distributed according to F_{ρ_i} on $[\underline{\rho}, \bar{\rho}]$ (with $\underline{\rho} \geq 0$, and $\bar{\rho} < \rho$), we have:

$$\begin{aligned}
\frac{d\text{Cov}[\kappa_i - \kappa, \mathbb{E}[\mathcal{E}|\kappa_i]|n_i]}{dn_i} &= \frac{d}{dn_i} \int_{\underline{\rho}}^{\bar{\rho}} ((\kappa_i - \kappa) \mathbb{E}[\mathcal{E}|\kappa_i]|n_i) dF_{\rho_i} - \mathbb{E}[\mathbb{E}[\mathcal{E}|\kappa_i]|n_i] \frac{d}{dn_i} \int_{\underline{\rho}}^{\bar{\rho}} (\kappa_i - \kappa |n_i) dF_{\rho_i} \\
&\quad - \mathbb{E}[\kappa_i - \kappa |n_i] \frac{d}{dn_i} \int_{\underline{\rho}}^{\bar{\rho}} (\mathbb{E}[\mathcal{E}|\kappa_i]|n_i) dF_{\rho_i} \\
&= \int_{\underline{\rho}}^{\bar{\rho}} \left(\frac{d(\kappa_i - \kappa)}{dn_i} \mathbb{E}[\mathcal{E}|\kappa_i] + (\kappa_i - \kappa) \frac{d\mathbb{E}[\mathcal{E}|\kappa_i]}{dn_i} \Big| n_i \right) dF_{\rho_i} \\
&\quad - \mathbb{E}[\mathbb{E}[\mathcal{E}|\kappa_i]|n_i] \int_{\underline{\rho}}^{\bar{\rho}} \left(\frac{d(\kappa_i - \kappa)}{dn_i} \Big| n_i \right) dF_{\rho_i} \\
&\quad - \mathbb{E}[\kappa_i - \kappa |n_i] \int_{\underline{\rho}}^{\bar{\rho}} \left(\frac{d\mathbb{E}[\mathcal{E}|\kappa_i]}{dn_i} \Big| n_i \right) dF_{\rho_i} \\
&= \int_{\underline{\rho}}^{\bar{\rho}} \left(\frac{d(\kappa_i - \kappa)}{dn_i} \mathbb{E}[\mathcal{E}|\kappa_i] \Big| n_i \right) dF_{\rho_i} - \mathbb{E}[\mathbb{E}[\mathcal{E}|\kappa_i]|n_i] \int_{\underline{\rho}}^{\bar{\rho}} \left(\frac{d(\kappa_i - \kappa)}{dn_i} \Big| n_i \right) dF_{\rho_i} \\
&\quad + \int_{\underline{\rho}}^{\bar{\rho}} \left((\kappa_i - \kappa) \frac{d\mathbb{E}[\mathcal{E}|\kappa_i]}{dn_i} \Big| n_i \right) dF_{\rho_i} - \mathbb{E}[\kappa_i - \kappa |n_i] \int_{\underline{\rho}}^{\bar{\rho}} \left(\frac{d\mathbb{E}[\mathcal{E}|\kappa_i]}{dn_i} \Big| n_i \right) dF_{\rho_i} \\
&= \text{Cov} \left[\frac{d(\kappa_i - \kappa)}{dn_i}, \mathbb{E}[\mathcal{E}|\kappa_i] \Big| n_i \right] + \text{Cov} \left[\kappa_i - \kappa, \frac{d\mathbb{E}[\mathcal{E}|\kappa_i]}{dn_i} \Big| n_i \right]
\end{aligned}$$

Note that the proof of Proposition 2 gives that $\mathbb{E}[\mathcal{E}|\kappa_i]|n_i$ is increasing in κ_i , which is decreasing in ρ_i , so $\mathbb{E}[\mathcal{E}|\kappa_i]|n_i$ is decreasing in ρ_i . Using the definition of $\kappa_i - \kappa$ from Lemma 3 we also have that

$$\frac{d^2(\kappa_i - \kappa)}{dn_i d\rho_i} = \frac{1 - \rho_i + n_i(\rho_i - 2)}{(1 + (n_i - 1)\rho_i)^3} < 0$$

So $\mathbb{E}[\mathcal{E}|\kappa_i]$ and $\frac{d(\kappa_i - \kappa)}{dn_i}$ are decreasing in ρ_i . By Schmidt's Lemma, $\text{Cov} \left[\frac{d(\kappa_i - \kappa)}{dn_i}, \mathbb{E}[\mathcal{E}|\kappa_i] \Big| n_i \right] > 0$.

For the second covariance above: from the definition of κ_i and κ in Lemma 3, we have that $\kappa_i - \kappa$ is decreasing in ρ_i . Substituting these definitions into (7), we have

$$\mu = \frac{nx}{n(1 + \rho_i\tau) + \tau(1 - \rho_i)} \quad \text{and} \quad \sigma^2 = \frac{\mu^2}{\kappa x^2} + \frac{1}{\tau_b}$$

we have that $\mathcal{I}|\kappa_i, n_i \sim \mathcal{N}[\mu, \frac{1}{\sigma^2}]$, and $\mathcal{E}|\kappa_i, n_i$ is a folded normal with mean given by (8), which is a function of μ and σ . We thus can write:

$$\begin{aligned}
\frac{d\mathbb{E}[\mathcal{E}|\kappa_i]}{dn_i} &= \frac{\partial\mathbb{E}[\mathcal{E}|\kappa_i]}{\partial\sigma} \cdot \frac{d\sigma}{dn_i} + \frac{\partial\mathbb{E}[\mathcal{E}|\kappa_i]}{\partial\mu} \cdot \frac{d\mu}{dn_i} \\
\frac{d^2\mathbb{E}[\mathcal{E}|\kappa_i]}{dn_i d\rho_i} &= \frac{d}{d\rho_i} \left(\frac{\partial\mathbb{E}[\mathcal{E}|\kappa_i]}{\partial\sigma} \right) \cdot \frac{d\sigma}{dn_i} + \frac{\partial\mathbb{E}[\mathcal{E}|\kappa_i]}{\partial\sigma} \cdot \frac{d^2\sigma}{dn_i d\rho_i} \\
&\quad + \frac{d}{d\rho_i} \left(\frac{\partial\mathbb{E}[\mathcal{E}|\kappa_i]}{\partial\mu} \right) \cdot \frac{d\mu}{dn_i} + \frac{\partial\mathbb{E}[\mathcal{E}|\kappa_i]}{\partial\mu} \cdot \frac{d^2\mu}{dn_i d\rho_i}
\end{aligned} \tag{16}$$

We will now show that (16) is negative. Starting with the second term note that when $\rho = 1$:

$$\begin{aligned}
\frac{d^2\sigma^2}{dn_i d\rho_i} &= 2\sigma \frac{d^2\sigma}{dn_i d\rho_i} + 2 \frac{d\sigma}{d\rho_i} \cdot \frac{d\sigma}{dn_i}, \text{ thus,} \\
\frac{d^2\sigma}{dn_i d\rho_i} &= \frac{1}{2\sigma} \left(\frac{d^2\sigma^2}{dn_i d\rho_i} - 2 \frac{d\sigma}{d\rho_i} \cdot \frac{d\sigma}{dn_i} \right) \\
&= \frac{1}{2\sigma} \left(-\frac{2n\tau(2\tau(n-1)(1-\rho_i) + n(1+\tau))}{(n(1+\rho_i\tau) + \tau(1-\rho_i))^4} + 2 \frac{n^2\tau(n-1) \cdot n\tau(1-\rho_i)}{\sigma^2(n(1+\rho_i\tau) + \tau(1-\rho_i))^6} \right) \\
&< \frac{n\tau}{\sigma} \left(\frac{-n^2(2\tau(n-1)(1-\rho_i) + n(1+\tau)) + n^2\tau(n-1)(1-\rho_i)}{\sigma^2(n(1+\rho_i\tau) + \tau(1-\rho_i))^6} \right) \\
&= -\frac{n\tau}{\sigma} \left(\frac{n^2(\tau(n-1)(1-\rho_i) + n(1+\tau))}{\sigma^2(n(1+\rho_i\tau) + \tau(1-\rho_i))^6} \right) < 0
\end{aligned}$$

where the inequality comes from setting $1/\tau_b = 0$ in the definition of σ^2 . Further, from (10)

we have $\frac{\partial\mathbb{E}[\mathcal{E}|\kappa_i]}{\partial\sigma} > 0$, so the second term of (16) is negative. Next for the fourth term:

$$\frac{d^2\mu}{dn_i d\rho_i} = -\frac{\tau(n-1)(1-\rho_i) + n(1+\tau)}{(n(1+\rho_i\tau) + \tau(1-\rho_i))^3} x$$

and from (10) we have that $\frac{\partial\mathbb{E}[\mathcal{E}|\kappa_i]}{\partial\mu} > 0$ if $x > 0$, and negative if $x < 0$. Therefore the fourth term of (16) is negative (or zero when $x = 0$).

We now show that the first and third terms of (16) are, together, negative. First, we examine the first term, defining $c = \mu/x$, and using the expression for $\frac{\partial\mathbb{E}[\mathcal{E}|\kappa_i]}{\partial\sigma}$ from (10):

$$\begin{aligned}
\frac{d\sigma}{d\rho_i} &= \frac{d\sigma}{d\mu} \cdot \frac{d\mu}{d\rho_i} = \frac{\mu}{\kappa\sigma x^2} \cdot \frac{d\mu}{d\rho_i} \\
\frac{d}{d\rho_i} \left(\frac{\partial\mathbb{E}[\mathcal{E}|\kappa_i]}{\partial\sigma} \right) &= \frac{\partial^2\mathbb{E}[\mathcal{E}|\kappa_i]}{(\partial\sigma)^2} \cdot \frac{d\sigma}{d\rho_i} + \frac{\partial^2\mathbb{E}[\mathcal{E}|\kappa_i]}{\partial\sigma\partial\mu} \cdot \frac{d\mu}{d\rho_i} = 2\phi \left[\frac{\mu}{\sigma} \right] \cdot \frac{\mu}{\sigma^2} \cdot \left(\frac{\mu}{\sigma} \frac{d\sigma}{d\rho_i} - \frac{d\mu}{d\rho_i} \right) \\
&= -2\phi \left[\frac{\mu}{\sigma} \right] \cdot \frac{\mu}{\sigma^2} \cdot \frac{d\mu}{d\rho_i} \cdot \left(1 - \frac{\mu^2}{\sigma^2 x^2} \right) = -2\phi \left[\frac{\mu}{\sigma} \right] \cdot \frac{\mu}{\sigma^2} \cdot \frac{d\mu}{d\rho_i} \cdot \left(\frac{\kappa}{c^2\tau_b + \kappa} \right)
\end{aligned}$$

and now the third term using the expression for $\frac{\partial \mathbb{E}[\mathcal{E}|\kappa_i]}{\partial \mu}$ from (10):

$$\begin{aligned} \frac{d}{d\rho_i} \left(\frac{\partial \mathbb{E}[\mathcal{E}|\kappa_i]}{\partial \mu} \right) &= \frac{\partial^2 \mathbb{E}[\mathcal{E}|\kappa_i]}{(\partial \mu)^2} \cdot \frac{d\mu}{d\rho_i} + \frac{\partial^2 \mathbb{E}[\mathcal{E}|\kappa_i]}{\partial \sigma \partial \mu} \cdot \frac{d\sigma}{d\rho_i} = 2\phi \left[\frac{\mu}{\sigma} \right] \cdot \frac{1}{\sigma} \cdot \left(\frac{d\mu}{d\rho_i} - \frac{\mu}{\sigma} \frac{d\sigma}{d\rho_i} \right) \\ &= 2\phi \left[\frac{\mu}{\sigma} \right] \cdot \frac{1}{\sigma} \cdot \frac{d\mu}{d\rho_i} \cdot \left(\frac{\kappa}{c^2 \tau_b + \kappa} \right). \end{aligned}$$

Deriving a few more quantities:

$$\begin{aligned} \frac{d\sigma}{dn_i} &= \frac{d\sigma}{d\mu} \cdot \frac{d\mu}{dn_i} = \frac{\mu}{\sigma x^2} \cdot \frac{d\mu}{dn_i} \\ \frac{d\mu}{d\rho_i} &= -\frac{n\tau(n-1)x}{(n(1+\rho_i\tau) + \tau(1-\rho_i))^2} = -\frac{c^2\tau(n-1)x}{n} \\ \frac{d\mu}{dn_i} &= \frac{\tau(1-\rho_i)x}{(n(1+\rho_i\tau) + \tau(1-\rho_i))^2} = \frac{c^2\tau(1-\rho_i)x}{n^2}, \end{aligned}$$

which we plug in to show that the first and third term together are:

$$\begin{aligned} \frac{d}{d\rho_i} \left(\frac{\partial \mathbb{E}[\mathcal{E}|\kappa_i]}{\partial \sigma} \right) \cdot \frac{d\sigma}{dn_i} + \frac{d}{d\rho_i} \left(\frac{\partial \mathbb{E}[\mathcal{E}|\kappa_i]}{\partial \mu} \right) \cdot \frac{d\mu}{dn_i} &= 2\phi \left[\frac{\mu}{\sigma} \right] \cdot \frac{1}{\sigma} \cdot \frac{d\mu}{d\rho_i} \cdot \left(\frac{\kappa}{c^2 \tau_b + \kappa} \right) \cdot \left(\frac{d\mu}{dn_i} - \frac{\mu}{\sigma} \frac{d\sigma}{dn_i} \right) \\ &= 2\phi \left[\frac{\mu}{\sigma} \right] \cdot \frac{1}{\sigma} \cdot \left(\frac{\kappa}{c^2 \tau_b + \kappa} \right)^2 \cdot \frac{d\mu}{d\rho_i} \cdot \frac{d\mu}{dn_i} \\ &= -2\phi \left[\frac{\mu}{\sigma} \right] \cdot \left(\frac{\kappa}{c^2 \tau_b + \kappa} \right)^2 \cdot \frac{c^4 \tau^2 x^2 (n-1)(1-\rho_i)}{\sigma n^3} < 0 \end{aligned}$$

This implies that $\frac{d^2 \mathbb{E}[\mathcal{E}|\kappa_i]}{dn_i d\rho_i} < 0$. Thus, both $\kappa_i - \kappa$ and $\frac{d\mathbb{E}[\mathcal{E}|\kappa_i]}{dn_i}$ are decreasing in ρ_i , and so Schmidt's Lemma gives that $\text{Cov} \left[\kappa_i - \kappa, \frac{d\mathbb{E}[\mathcal{E}|\kappa_i]}{dn_i} \middle| n_i \right] > 0$. This implies that $\frac{d}{dn_i} \text{Cov}[\kappa_i - \kappa, \mathbb{E}[\mathcal{E}|\kappa_i]|n_i] > 0$, as desired. Note that $\rho = 1$ is needed only to ensure the negativity of the second term, so if $\rho < 1$, the overall expression will still be negative. (Simulations indicate that any parameter values compatible with Implication 2 are sufficient to guarantee the negativity of the second term.) ■

Proof of Proposition 8 and Corollary 9: Fix $n_i = n$ and consider an individual i with ideology \mathcal{I} , overconfidence κ_i , and preference bias b_i . Suppose, without loss of generality that $\mathcal{I}_i > 0$. Note that $\mathbb{E}_i[x] = \mathcal{I}_i - b_i$. This means that we have $U_R(b_i|x) > U_L(b_i|x)$ if and only if $x > -b_i$. Thus, $\text{Prob}_i[U_R(b_i|x) > U_L(b_i|x)] = \text{Prob}_i[x > -b_i] = 1 - \text{Prob}_i[x < -b_i]$. By construction this is equal to

$$1 - \Phi\left[(-b_i - (\mathcal{I}_i - b_i))\sqrt{\tau + \kappa_i}\right] = \Phi\left[\mathcal{I}_i\sqrt{\tau + \kappa_i}\right]. \quad (17)$$

As $\mathcal{I}_i > 0$, $\mathcal{I}_i\sqrt{\tau + \kappa_i}$ must be strictly increasing in κ_i conditional on \mathcal{I}_i , and in \mathcal{I}_i conditional on κ_i . The same must therefore hold for $\Phi[\mathcal{I}_i\sqrt{\tau + \kappa_i}]$, and hence for $\text{Prob}_i[U_R(b_i|x) > U_L(b_i|x)]$. Note that specular results hold conditional on $\mathcal{I}_i < 0$. Thus, we can replace \mathcal{I}_i with $\mathcal{E}_i = |\mathcal{I}|$ in (17).

Finally, $F_c(\cdot)$ and $F'_c(\cdot)$ are c.d.f.s and thus increasing in their arguments. This, together with the previous argument gives the second and third parts of the proposition and corollary, conditional on n . This, combined with Proposition 2 gives the first part of Proposition 8 and Corollary 9.

Consider $n \sim F_n$. Suppose $\rho = 1$, so $\kappa = 1$ for all citizens. Then, more overconfident citizens, those with greater $\kappa_i - \kappa$, will be more confident (greater κ_i). This gives that, conditional on ideology, more overconfident citizens are more likely to turn out. Moreover, those with the same level of overconfidence will have the same confidence, so more ideological citizens will be more likely to turn out, conditional on overconfidence. This, combined with Proposition 2 gives the first part of Proposition 8 and Corollary 9 independent of n . These would continue to hold for ρ large enough, as in Implication 2. ■

Proof of Proposition 10: The utility of taking action a after receiving n signals is:

$$-\int (a - b_i - x)^2 dF_i[x] = -(a - b_i - \mathbb{E}_i[x])^2 - \text{Var}_i[x|n]$$

so the value of an additional signal is:

$$V_i(\kappa_i|n) = \text{Var}_i[x|n] - \text{Var}_i[x|n+1] = \frac{1 + (n-1)\rho_i}{n + \tau(1 + (n-1)\rho_i)} - \frac{1 + n\rho_i}{n + 1 + \tau(1 + n\rho_i)}$$

and plugging in for $\rho_i = \frac{n-\kappa_i}{(n-1)\kappa_i}$ from Lemma 3 we have:

$$\begin{aligned} V_i(\kappa_i|n) &= \frac{\kappa_i(\kappa_i - 1)}{(\kappa_i + \tau)(n^2(\kappa_i + \tau) - \kappa_i(1 + \tau))} \\ \frac{dV_i(\kappa_i|n)}{d\kappa_i} &= \frac{n^2\tau^2(\kappa_i - 1) + \kappa_i^2(n^2 - 1)(1 + 2\tau) + \kappa_i\tau^2(n^2 - \kappa_i)}{(\kappa_i + \tau)^2(n^2(\kappa_i + \tau) - \kappa_i(1 + \tau))^2} > 0 \end{aligned}$$

where the sign of the last line follows from $1 < \kappa_i \leq n$, as $n \geq 2$ and $\rho_i \in [0, 1)$.

We have that n_i^* is determined by

$$\arg \max_{n_i} \quad - (a - b_i - \mathbb{E}_i[x])^2 - \text{Var}_i[x|n_i] - cn_i \quad \Rightarrow \quad \frac{d\text{Var}_i[x|n_i^*]}{dn} = -c$$

and thus n_i^* is defined implicitly by

$$\frac{1 - \rho_i}{(n_i^* + \tau(1 + (n_i^* - 1)\rho_i))^2} - c = 0$$

and using the implicit function theorem we have

$$\frac{dn_i^*}{d\rho_i} = -\frac{\tau(n_i^* - 1)(1 - \rho_i) + n_i^*(1 + \tau)}{2(1 - \rho_i)(1 + \rho_i\tau)} < 0$$

so n_i^* is increasing in correlational neglect $\rho - \rho_i$. ■

Proof of Proposition 11: The posterior of citizen i about the bias of citizen j after observing \mathcal{I}_j is:

$$\begin{aligned} \mathcal{L}(b_j|\mathcal{I}_j) &\propto \mathcal{L}(\mathcal{I}_j|b_j)\mathcal{L}(b_j) \\ &\propto \left[\int_{-\infty}^{+\infty} \exp \left\{ -\frac{\kappa}{2} \left(x - \frac{\kappa + \tau}{\kappa} (\mathcal{I}_j - b) \right)^2 \right\} * \exp \left\{ -\frac{\kappa_i + \tau}{2} (x - \mathcal{I}_j)^2 \right\} dx \right] * \exp \left\{ -\frac{\tau_b b_j^2}{2} \right\} \end{aligned}$$

This is a normal distribution with mean

$$\left(\mathcal{I}_j \frac{\kappa + \tau}{\kappa} - (\mathcal{I}_j - b_i) \right) * \frac{\kappa(\kappa_i + \tau)(\kappa + \tau)}{(\kappa + \tau)(\tau^2 + \tau\kappa + \tau_b\kappa) + \kappa_i(\tau^2 + 2\tau\kappa + \kappa\tau_b + \kappa^2)}$$

where the second term is positive and increasing in κ_i . Thus, if $\mathcal{I}_j > (\mathcal{I}_j - b_i) \left(\frac{\kappa}{\kappa + \tau} \right)$, then $\mathbb{E}_i[b_j] > 0$ and $\mathbb{E}_i[b_j]$ is increasing in κ_i . If $\mathcal{I}_j < (\mathcal{I}_j - b_i) \left(\frac{\kappa}{\kappa + \tau} \right)$, then $\mathbb{E}_i[b_j] < 0$ and $\mathbb{E}_i[b_j]$ is decreasing in κ_i . Thus, $|\mathbb{E}_i[b_j]|$ is increasing in κ_i .

The existence of $\alpha_i, \beta_i \in \mathbb{R}_{++}$ s.t. the ideology of citizen i after communication is $\alpha_i \mathcal{I}_i + \beta_i \mathcal{I}_j$ is a standard result of Bayesian updating. $\alpha_i + \beta_i \neq 1$ because ideology is a signal of both bias and beliefs. The fact that α_i increases and β_i decreases in κ_i is a direct consequence of the standard result that citizens with a high prior precision update less, and also because here they will tend to assign a higher probability to the fact that the differences in ideologies

are due to differences in preference biases. In particular, solving for $\alpha_i, \beta_i \in \mathbb{R}_+$:

$$\alpha_i = \frac{(\kappa_i + \tau)(\tau^2 + 2\tau\kappa + \kappa\tau_b + \kappa^2)}{(\kappa + \tau)(\tau^2 + \tau\kappa + \tau_b\kappa) + \kappa_i(\tau^2 + 2\tau\kappa + \kappa\tau_b + \kappa^2)}, \quad \beta_i = \frac{\kappa(1 - \alpha_i)}{\kappa + \tau}$$

and thus

$$\frac{d\alpha_i}{d\kappa_i} = \frac{\tau_b\kappa^2(\tau_b\kappa + (\kappa + \tau)^2)}{(\kappa_i + \tau)(\kappa + \tau)^2 + \tau_b\kappa(\kappa_i + \kappa + \tau)^2} > 0.$$

Thus, α_i is increasing in κ_i , so β_i is decreasing in κ_i . ■

Proof of Proposition 12: We begin with the second and third parts of the proposition.

By Bayes' rule: $\mathcal{L}(\kappa_j|\mathcal{I}_j) \propto \mathcal{L}(\mathcal{I}_j|\kappa_j)\mathcal{L}(\kappa_j)$. Note that $\mathcal{L}(\mathcal{I}_j|\kappa_j) = \phi_{\mathcal{I}_i, \kappa_i + \tau}(\mathcal{I}_j(\frac{\tau + \kappa_j}{\kappa_j}))$, where $\phi_{\mu, \tau}(\cdot)$ denotes the p.d.f. of a normal distribution with mean μ and precision τ . To prove that $\mathbb{E}_i[\kappa_j]$ is increasing in κ_i , it is sufficient to prove that, for any $\kappa_j, \kappa'_j \in \text{supp}(F)$, $\kappa_j < \kappa'_j$, the ratio

$$\begin{aligned} \frac{\mathcal{L}(\mathcal{I}_j|\kappa'_j)}{\mathcal{L}(\mathcal{I}_j|\kappa_j)} &= \frac{\sqrt{\frac{\kappa_i + \tau}{2\pi}} \exp\left\{-\frac{(\kappa_i + \tau)}{2} \left(\mathcal{I}_j\left(\frac{\tau + \kappa'_j}{\kappa'_j}\right) - \mathcal{I}_i\right)^2\right\}}{\sqrt{\frac{\kappa_i + \tau}{2\pi}} \exp\left\{-\frac{(\kappa_i + \tau)}{2} \left(\mathcal{I}_j\left(\frac{\tau + \kappa_j}{\kappa_j}\right) - \mathcal{I}_i\right)^2\right\}} \\ &= \exp\left\{-\frac{\kappa_i + \tau}{2} \left(\left(\mathcal{I}_j\left(\frac{\tau + \kappa'_j}{\kappa'_j}\right) - \mathcal{I}_i\right)^2 - \left(\mathcal{I}_j\left(\frac{\tau + \kappa_j}{\kappa_j}\right) - \mathcal{I}_i\right)^2\right)\right\} \end{aligned}$$

is increasing in κ_i . This holds if and only if

$$\left(\mathcal{I}_j\left(\frac{\tau + \kappa'_j}{\kappa'_j}\right) - \mathcal{I}_i\right)^2 < \left(\mathcal{I}_j\left(\frac{\tau + \kappa_j}{\kappa_j}\right) - \mathcal{I}_i\right)^2 \quad (18)$$

for all $\kappa_j, \kappa'_j \in \text{supp}(F)$, $\kappa_j < \kappa'_j$. If the converse of (18) holds for all $\kappa_j, \kappa'_j \in \text{supp}(F)$, $\kappa_j < \kappa'_j$, this is sufficient for $\mathbb{E}_i[\kappa_j]$ to be decreasing in κ_i .

As $\frac{\tau + \kappa_j}{\kappa_j}$ is decreasing in κ_j , $\mathcal{E}_j(\frac{\tau + \kappa'_j}{\kappa'_j}) < \mathcal{E}_j(\frac{\tau + \kappa_j}{\kappa_j})$ since $\kappa_j < \kappa'_j$. This implies (18) holds if $\mathcal{I}_i * \mathcal{I}_j < 0$ or $\mathcal{E}_j > \mathcal{E}_i$. By contrast, the converse holds if $\kappa_j, \kappa'_j \in \text{supp}(F)$, $\kappa_j < \kappa'_j$ if $\mathcal{I}_i * \mathcal{I}_j > 0$, and $\mathcal{E}_i > \frac{\tau + \kappa_j}{\kappa_j} \mathcal{E}_j$.

Finally, as in the Proof of Proposition 11, the first part follows from standard properties

Appendix B Survey Details—For Online Publication

The typical way psychologists measure overconfidence is not well suited to surveys. They often use a very large number of questions—up to 150 (see, for example, Alpert and Raiffa, 1969/1982; Soll and Klayman, 2004)—and elicit confidence using confidence intervals, which may be difficult for the average survey respondent to understand (see, for example, Juslin et al., 1999; Rothschild, 2011).

Our methodology for measuring overconfidence on surveys uses three innovations. The first two are due to Ansolabehere et al. (2011). First, the questions we use are about either quantities that everyone knows the scale of, such as dates, or the scale is provided, as in the case of unemployment or inflation. That is, when asking about unemployment rates, the question gives respondents the historical minimum, maximum, and median of that rate. This has been shown to reduce the number of incorrect answers simply due to a respondent not knowing the appropriate scale (Ansolabehere et al., 2013). Second, confidence is elicited on a qualitative scale, which is easily understandable by survey respondents and allows for more conservative controls for actual knowledge.

The third innovation is a modification of the second, and was only utilized on the 2011 CCES. For our general knowledge questions—the year the telephone was invented, the population of Spain, the year Shakespeare was born, and the percent of the U.S. population that lives in California—we elicited confidence using an inverted confidence interval.¹ That is, rather than asking for a confidence interval directly, which we felt may have been too challenging for survey respondents, we asked them to give their estimates of the probability that the true answer was in some interval around their answer. So, for example, after giving their best guess as to the date of Shakespeare’s birth, respondents were asked:

¹Note that these general knowledge questions were all from previous research on overconfidence.

What do you think the percent chance is that your best guess, entered above, is within 50 years of the actual answer?

Given a two-parameter distribution, such as a normal, this is enough to pin down the variance of a respondent's belief.

The sum total of these innovations is that overconfidence can be elicited using a small number of questions that are understandable to most survey respondents, rather than just to university undergraduates.

Appendix B.1 Survey Questions

We next present the text of the questions used to construct our overconfidence measure on the 2010 and 2011 CCES, as described in Section 2.2.1. Instructions in brackets indicate limitations on possible answers implemented by the survey company—these were not displayed to respondents. If a survey respondent tried to enter, say, text where only a positive number was allowed, they would be told to edit their entry to conform with the limitations placed on the response field. If a respondent tried to skip a question, the survey would request that the respondent give an answer. If the respondent tried to skip the same question a second time, they were allowed to do so.

1. The unemployment rate is the percent of people actively searching work but not presently employed. Since World War II it has ranged from a low of 2 percent to a high of 11 percent.

What is your best guess about the unemployment rate in the United States today? Even if you are uncertain, please provide us with your best estimate of the percent of people seeking work but currently without a job in the United States.

----% [only allow a positive number]

2. How confident are you of your answer to this question?
 - No confidence at all
 - Not very confident
 - Somewhat unconfident

- Somewhat confident
- Very confident
- Certain

3. The inflation rate is the annual percentage change in prices for basic goods like food, clothing, housing, and energy. Since World War II it has ranged from a high of 14 percent (a 14% increase in prices over the previous year) to a low of -2 percent (a 2% decline in prices over the previous year).

What is your best guess about the inflation rate in the United States today? Even if you are uncertain, please provide us with your best estimate of about what percent do you think prices went up or down in the last 12 months.

Do you think prices went up or down?

- Up
- Down

4. By what percent do you think prices went up or down?
----% [only allow a positive number]

5. How confident are you of your answer to this question?

- No confidence at all
- Not very confident
- Somewhat unconfident
- Somewhat confident
- Very confident
- Certain

6. The unemployment rate is the percent of people actively searching work but not presently employed. Since World War II it has ranged from a low of 2 percent to a high of 11 percent.

What do you expect the unemployment rate to be a year from now? Even if you are uncertain, please provide us with your best estimate of the percent of people who will be seeking but without a job in the United States in November, 2011.

----% [only allow a positive number]

7. How confident are you of your answer to this question?

- No confidence at all
- Not very confident
- Somewhat unconfident

- Somewhat confident
- Very confident
- Certain

8. The inflation rate is the annual percentage change in prices for basic goods like food, clothing, housing, and energy. Since World War II it has ranged from a high of 14 percent (a 14% increase in prices over the previous year) to a low of -2 percent (a 2% decline in prices over the previous year).

What do you expect the inflation rate to be a year from now? Even if you are uncertain, please provide us with your best estimate of about what percent do you expect prices to go up or down in the next 12 months.

Do you expect prices to go up or down?

- Up
- Down

9. By what percent do you expect prices to go up or down?
 ----% [only allow a positive number]

10. How confident are you of your answer to this question?

- No confidence at all
- Not very confident
- Somewhat unconfident
- Somewhat confident
- Very confident
- Certain

Next, we list the questions from the 2011 CCES used to construct the overconfidence measures discussed in Section 6.1. Note that the unemployment questions were changed from 2010, in accordance with the evolving research agenda of Ansolabehere et al..

1. In what year was the telephone invented? Even if you are not sure, please give us your best guess.

2. How confident are you of your answer to this question?

- No confidence at all

- Not very confident
 - Somewhat unconfident
 - Somewhat confident
 - Very confident
 - Certain
3. As a different way of answering the previous question, what do you think the percent chance is that your best guess, entered above, is within 25 years of the actual answer?
----%
4. What is the population of Spain, in millions? Even if you are not sure, please give us your best guess.

5. How confident are you of your answer to this question?
- No confidence at all
 - Not very confident
 - Somewhat unconfident
 - Somewhat confident
 - Very confident
 - Certain
6. As a different way of answering the previous question, what do you think the percent chance is that your best guess, entered above, is within 15 million of the actual answer?
----%
7. In what year was the playwright William Shakespeare born? Even if you are not sure, please give us your best guess.
----%
8. How confident are you of your answer to this question?
- No confidence at all
 - Not very confident
 - Somewhat unconfident
 - Somewhat confident
 - Very confident
 - Certain

9. As a different way of answering the previous question, what do you think the percent chance is that your best guess, entered above, is within 50 years of the actual answer?
----%
10. What percent of the US population lives in California? Even if you are not sure, please give us your best guess.

11. How confident are you of your answer to this question?
- No confidence at all
 - Not very confident
 - Somewhat unconfident
 - Somewhat confident
 - Very confident
 - Certain
12. As a different way of answering the previous question, what do you think the percent chance is that your best guess, entered above, is within 5 percentage points of the actual answer?
----%
13. According to the Bureau of Labor Statistics, since World War II the most non-agricultural jobs the US economy has lost in a year is 5.4 million. The most jobs gained in a year has been 4.2 million. Over the same period, the US economy has gained an average of 1.4 million jobs a year.
- What is your best guess about the number of jobs gained or lost in the last year?
Over the past year, I think the US economy has overall
- Lost jobs
 - Gained jobs
14. How many jobs do you think have been lost or gained over the past year?
---- million jobs [only allow a positive number]
15. How confident are you of your answer to this question?
- No confidence at all
 - Not very confident
 - Somewhat unconfident
 - Somewhat confident
 - Very confident

- Certain

16. The inflation rate is the annual percentage change in prices for basic goods like food, clothing, housing, and energy. Since World War II it has ranged from a high of 14.4 percent (a 14.4% increase in prices over the previous year) to a low of -1.2 percent (a 1.2% decline in prices over the previous year).

What is your best guess about the inflation rate in the United States today?

Do you think prices went up or down?

- Up
- Down

17. By what percent do you think prices went up or down?

----% [only allow a positive number]

18. How confident are you of your answer to this question?

- No confidence at all
- Not very confident
- Somewhat unconfident
- Somewhat confident
- Very confident
- Certain

19. According to the Bureau of Labor Statistics, since World War II the most non-agricultural jobs the US economy has lost in a year is 5.4 million. The most jobs gained in a year has been 4.2 million. Over the same period, the US economy has gained an average of 1.4 million jobs a year.

What is your best guess about the number of jobs that will be gained or lost over the next year?

Over the next year, I think the US economy will overall

- Lose jobs
- Gain jobs

20. How many jobs do you think the US economy will lose or gain over the next year?

---- million jobs [only allow a positive number]

21. How confident are you of your answer to this question?

- No confidence at all
- Not very confident

- Somewhat unconfident
 - Somewhat confident
 - Very confident
 - Certain
22. The inflation rate is the annual percentage change in prices for basic goods like food, clothing, housing, and energy. Since World War II it has ranged from a high of 14.4 percent (a 14.4% increase in prices over the previous year) to a low of -1.2 percent (a 1.2% decline in prices over the previous year).
- What do you expect the inflation rate to be a year from now?
- Do you expect prices to go up or down?
- Up
 - Down
23. By what percent do you expect prices to go up or down?
 ----% [only allow a positive number]
24. How confident are you of your answer to this question?
- No confidence at all
 - Not very confident
 - Somewhat unconfident
 - Somewhat confident
 - Very confident
 - Certain

Appendix C Historical Data—For Online Publication

While the results in the text support our theory, they raise the concern, briefly discussed in Section 3.3, that overconfidence and conservatism are somehow linked in a way not accounted for in our theory. This section contains a limited analysis to address this concern, and concludes by gathering together a number of facts in order to construct a post-hoc rationalization of this fact that goes beyond the findings in Section 3.3.

As the data in the text are the only we are aware of that provide both good measures of political ideology and of overconfidence, we turn to a survey with greater coverage over time, but more limited measures of ideology, and only a proxy for overconfidence: the American National Election Study (ANES). In particular, we follow a strategy based on the fact that many studies over time, including ours, have found men to be more overconfident than women and use male as a proxy for “more overconfident”.¹

To begin the analysis we add a basic result.

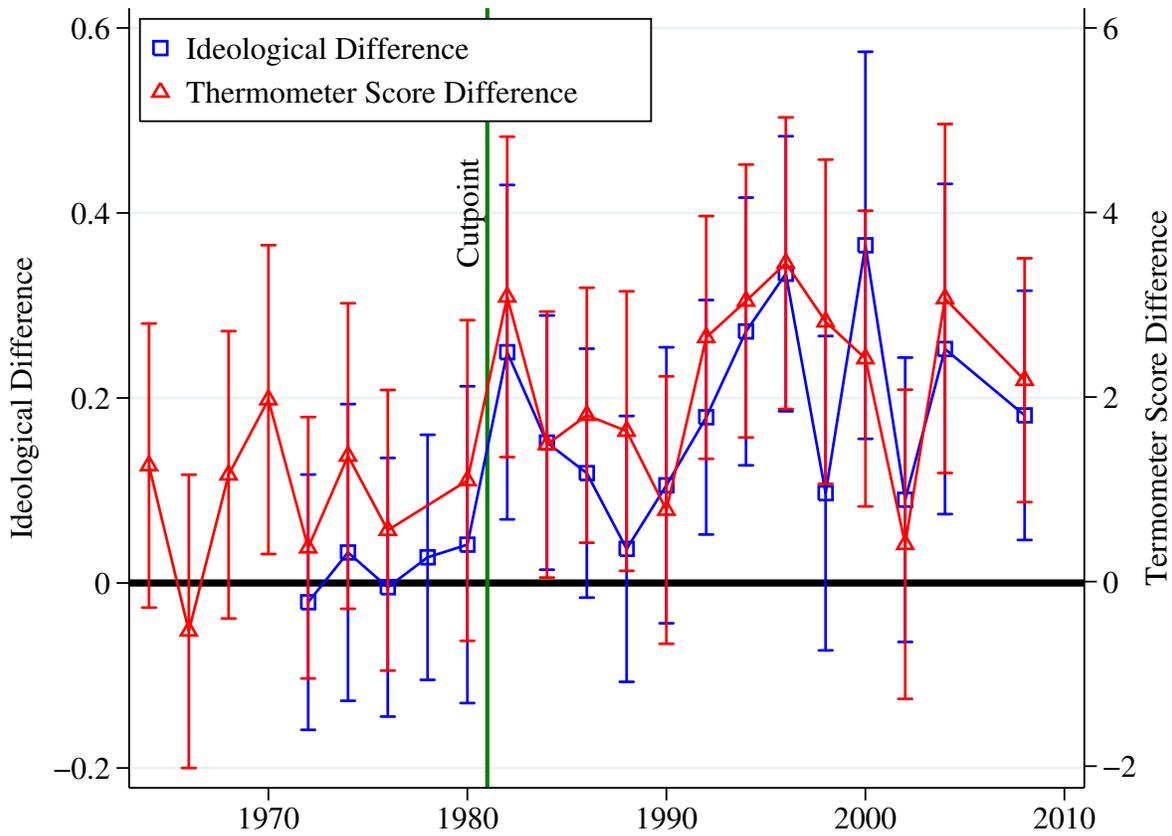
Proposition C.1. *If more overconfident citizens have the same average ideology as less overconfident citizens, then overconfidence is equally correlated with ideological extremeness for both those to the right and to the left of center.*

Proof of Proposition C.1: Consider two citizens with $\kappa_1 > \kappa_2$. As $\mathbb{E}[\mathbb{E}_i[\mathcal{I}|\kappa]] = \frac{\kappa x}{\tau + \kappa}$, we have that $\frac{\kappa_1 x}{\tau + \kappa_1} = \frac{\kappa_2 x}{\tau + \kappa_2} \iff x = 0$. Thus, $\mathcal{I}|\kappa \sim \mathcal{N}\left[0, \frac{\tau_b(\tau + \kappa)^2}{\tau_b \kappa^2 + (\tau + \kappa)^2}\right]$. As this is symmetric about zero for all κ , it implies $\text{Cov}[\mathbb{E}[\mathcal{E}|\kappa, \mathcal{I} \geq 0], \kappa] = \text{Cov}[\mathbb{E}[\mathcal{E}|\kappa, \mathcal{I} \leq 0], \kappa]$ and $\text{Var}[\mathcal{I}|\mathcal{I} \geq 0] = \text{Var}[\mathcal{I}|\mathcal{I} \leq 0]$. Finally, as this implies $f(\kappa|\mathcal{I} \geq 0) = f(\kappa|\mathcal{I} \leq 0) = f(\kappa)$, thus, $\text{Var}[\kappa|\mathcal{I} \geq 0] = \text{Var}[\kappa|\mathcal{I} \leq 0]$. Taken together this implies $\text{Corr}[\mathcal{E}, \kappa|\mathcal{I} \geq 0] = \text{Corr}[\mathcal{E}, \kappa|\mathcal{I} \leq 0]$. ■

Next, we investigate if there is variation over time in the difference between the average ideology of men and women. In particular, we have both self-reported ideology and the difference between respondent’s thermometer scores for “liberals” and “conservatives”, which is intended as a measure of ideology. Figure 5 plots the difference between men and women on both of these scales over time with 95% confidence intervals in each year we have data. There is a clear rightward shift for men between 1980 and 1982. We divide the sample into two parts around 1981, and conduct a similar analysis to Table 5. The results can be found

¹Barber and Odean (2001) use male as an instrument for overconfidence in a study of financial risk taking. We have not adopted this strategy as being male is likely correlated with numerous other factors which may also affect the dependent variables we are interested in (Grinblatt and Keloharju, 2009). The curious reader may be interested to know that doing so approximately triples the effect size of overconfidence in the regressions presented in the main text.

Figure 5: Men became significantly more conservative after 1980.



Note: Thermometer scores were not collected in 1978. There was no panel survey in 2006.

in Table C.1.

The results in Table C.1 are broadly consistent with the patterns predicted by Proposition 5 and Proposition C.1. For self-reported ideology, there is no statistical difference in average ideology between men and women before 1982, and, consistent with Proposition C.1, men are equally more ideologically extreme, regardless of their ideological direction. After 1982, men are significantly further to the right than women on average, and, consistent with Proposition 5, being male exhibits greater correlation with ideological extremeness for those to the right of the population median than for those to the left of the median.² For the thermometer

²The magnitudes of the coefficients are similar in magnitude to the coefficient on gender in the analysis of the 2010 CCES in Sections 3.2 and 3.3. After 1988, the self-reported ideological extremeness measure exhibits no statistically significant correlation with gender for those to the left of the median, which is consistent with the analysis in Table 5.

Table C.1: Data from the ANES is broadly consistent with Proposition 5 and Proposition C.1.

Time Frame	Up to 1980			1982 and After		
Dep. Variable	Ideology	Extremeness		Ideology	Extremeness	
Sample		Left of Median	Right of Median		Left of Median	Right of Median
Panel A: Self-Reported Ideology						
Male	0.013 (.032)	0.14*** (.027)	0.10*** (.025)	0.18*** (.022)	0.012** (.018)	0.13*** (.017)
Difference		0.035 (.037)			0.12*** (.025)	
Year Fixed Effects	Y	Y	Y	Y	Y	Y
<i>N</i>	6,880	4,241	5,132	16,862	9,821	12,593
Panel B: Thermometer Scores						
Male	0.88*** (.28)	0.72*** (.24)	1.61*** (.25)	2.17*** (.22)	-0.15 (.18)	1.92*** (.20)
Difference		0.89** (.35)			2.07*** (.27)	
Year Fixed Effects	Y	Y	Y	Y	Y	Y
<i>N</i>	11,439	6,551	8,709	20,192	11,795	14,428

Notes: ***, **, * denote statistical significance at the 1%, 5% and 10% level with standard errors in parentheses. The *N* of the split-sample regressions do not sum to the *N* of the ideology regression due to the fact that those respondents with the median ideology are included in both regressions.

scores, the difference in correlation between right and left expands as the ideological difference between men and women increases.

While the results presented here are broadly consistent with theory, and suggest that overconfidence and ideological extremeness are correlated for both left and right, depending on the time-frame under study, further research is needed. In particular, gender is correlated with a multitude of political differences, and the shift in ideology that occurred in the 1980s has many potential explanations that have nothing to do with overconfidence. We believe it is best to note that the available data is consistent with theory, but that better data is

clearly needed.³

Is There a Connection between Overconfidence and Conservatism? Table 3 shows a clear correlation between overconfidence and conservatism. But is this a more general phenomenon? While our data is limited, and our thinking about this issue is decidedly post-hoc, we believe the answer is no.

There are three pieces of weak evidence against a more general relationship between overconfidence and conservatism. The first piece is noted in Section 3.3: if overconfidence and conservatism were both caused by some underlying factor, then there should be a negative correlation between extremeness and overconfidence for those left-of-center in Table 5, yet there is not. Second, as noted in Section 3.3 older people on both the left and the right are more ideologically extreme. Third is the analysis in this section, which suggests that in the past overconfidence was equally linked to liberalism and conservatism.

So if there is no general relationship between overconfidence and conservatism, what can explain this relationship in 2010 (and 2011)? This relationship, and the facts above, are consistent with our theory if we add that a citizen's ideological leaning, left or the right, is the product of the political environment when he or she became politically active. In technical terms, this would specify that the aggregate bias in x discussed in Footnote 11 would be the political zeitgeist as a citizen comes of age. More descriptively, correlational neglect gives people the tendency to become both more ideologically extreme and more overconfident as they age. However, the theory makes no prediction about which ideological direction they will tend towards, and it is known that this responds to environmental factors when a person first becomes politically active (Meredith, 2009; Mullainathan and Washington, 2009). As the most ideologically extreme and overconfident people in 2010 began participating in politics in the late 1970s and 1980s, when conservatism was in the ascendency, this would rationalize

³Another proxy for overconfidence, especially given the results in Section 3.1, is age. However, across the entire timespan of the ANES cumulative dataset, age has a roughly constant, statistically significant, positive correlation with ideology. That is, the hypothesis of Proposition C.1 is never met, and thus, there is no way to contrast that proposition with the results in Section 3.3.

the patterns we see in the data. This further implies that in other periods in time there may be a relationship between overconfidence and liberalism.

Appendix D Additional Specifications—For Online Publication

Appendix D.1 Theoretical

This section addresses, in a casual way, a number of theoretical questions that have been posed to us. While the result of our inquiry into these questions did not produce results that merit a discussion in the main text, we thought it would be useful to record the results.

Distributional Assumptions Throughout the paper we make heavy use of normal distributions. This has advantages for both tractability and interpretation. In particular, tractability is helped by the fact that a normal is a self-conjugate prior, and that properties of the normal are well studied in statistics. The advantage in interpretation comes from the fact that the normal is a two-parameter distribution (the mean and precision), so it is straightforward to implement and interpret overconfidence as a function of precision without worrying about the effects of higher (or lower) order moments.

However, this leads to questions about how much our results are driven by the use of normal distributions. Or, conversely, many seminar attendees have conjectured that it would be straight-forward to extend our results to well-behaved distributions. Here we give some guidance on these questions.

We start by discussing how our results might generalize to other distributions. Without the normal distribution, the correlational neglect model becomes intractable. The value of this model is that it allows us to make predictions about the role of age that could not be obtained under any fully Bayesian model, as discussed in Section 3.1.

However if one is willing to put aside these predictions, it is possible to discuss the role of

the normal when citizens receive uncorrelated signals they over-interpret (as in the “model” of Lemma 3). The proof of Proposition 2 relies on the fact that both overconfidence and extremeness are increasing in correlational neglect. Intuitively this seems as though it would hold for a wide range of distributions (at least when $x = 0$), but it is quite difficult to show this analytically. Using the normal distribution, then, gives two advantages: tractability, as just discussed, and clarity, as unique among commonly used multi-variate conjugate distributions the variance-covariance matrix is a parameter. This makes the definition of correlational neglect very clear as it does not require tweaking other parameters of the distribution. Indeed, while we have verified that our primary result holds when priors are distributed according to a Beta (or uniform) distribution, and signals are Bernoulli, the interpretation of even this simple model is much more difficult.

If one uses a support with only two possible states, then our results may not always hold. However, it is known that such a setup (without overconfidence) produces perverse results: see McMurray (2013a). In particular, with only two states, the precision of beliefs may decrease, rather than increase with more signals. However, this would be inconsistent empirical results in Section 3.1.

Multi-Dimensional Issue Spaces: Our theory has implications for how ideology on different dimensions would be related to overconfidence. For example, if the information on a given dimension were all public, with agreed upon correlational structure, then there should be no relationship between ideology and overconfidence on that dimension. While this implication is straight-forward to work out, we did not feel that it was testable with current data.

In particular, in order to test this, one would need to know quite a bit about where citizens get their data from, and how citizens infer about how this data affects them. For example, even if most economic information is public, how that information relates to a citizen’s permanent income is more opaque. Learning about that relationship would entail

seeing how nationwide economic performance seemed to affect a citizen's own employment situation. As these very personal signals would have an unknown correlational structure, there is plenty of room for correlational neglect.

Likewise, positions on a social issue like gay marriage may appear to have no informational content at all, and hence, there should be no relationship between overconfidence and ideology on this dimension. However, it is perfectly reasonable that one's position on gay marriage may depend on beliefs about the likelihood that a loved one, say a child, is gay. This likelihood may be drawn, in part, from the number of openly gay people in a citizen's social environment. If a citizen neglects the fact that they live in a religious community where others are not open about their sexuality, then they will tend to underestimate the probability that a loved one may turn out to be gay. This will lead to both overconfidence and more extreme positions, as before.

We believe that applying our theory to multi-dimensional spaces would be interesting, and possibly fruitful. We refrain from doing so in this paper because it does not add to the predictions we can test in our data.

Other Dimensions of Personality: We treat correlational neglect as akin to a personality trait, which has raised questions of how this might be related to other personality traits. In particular, previous research has found that overconfidence is related to the extraversion of the "Big Five" personality inventory (Schaefer et al., 2004), although Moore and Healy (2007) has found that it is orthogonal to all traits in the Big Five. Regardless, extraversion does not have any significant explanatory power for the political behaviors we consider here (Gerber et al., 2010, 2012). Other studies have noted a link between overconfidence and narcissism. Little is known about the relationship between narcissism and political behavior, nor are there formal theories (that we are aware of) that relate narcissism to decision making more generally.

Appendix D.1.1 Voting

Our model of voter turnout, and partisan identification, is based on a specific form of expressive voting (Fiorina, 1976; Brennan and Hamlin, 1998). In particular a citizen i votes if and only if

$$\left| \text{Prob}_i[U_R(b_i|x) > U_L(b_i|x)] - \frac{1}{2} \right| - c_i > 0, \quad (19)$$

where c_i is an i.i.d. draw from some distribution F_c , which is strictly increasing on $(0, \frac{1}{2})$. In addition $c_i \perp (\rho_i, b_i, e_{it})$.

While any political economy model where turnout is exogenous implicitly uses an expressive voting model (and others use it more explicitly, see Knight, 2013), there are a number of other approaches in the political economy literature. As each approach has its partisans, we thought it worthwhile to discuss those models, and show, where possible, how our model relates to them.

Before discussing alternative models, we should note that we focused on the expressive approach because we believe it is correct, and because it is compatible (as shown below) with a promising approach in the literature, that voters are choice- or regret-avoidant (Matsusaka, 1995; Degan and Merlo, 2011; Degan, 2013).

In addition, this modeling approach allows for both non-trivial turnout and strong partisan identification even if the policies proposed by political parties are similar to each other, as seems to be the case in reality (Snowberg et al., 2007a,b). This is generally not possible in more traditional models. To make this specific, suppose that both parties propose very similar platforms, and consider a citizen who is very confident that the best policy for her is proposed by party R . According to our model, this citizen would strongly identify with, and turn out to vote, for party R . However, if these behaviors were rooted in expected utility, and the parties espoused similar platforms, this would not hold. For any reasonably smooth utility function there is a small difference in utility between the two parties—and hence no reason to strongly identify with one party or the other, or turnout.

Pivotal Voting: In these models, the turnout decision is driven largely by whether or not a voter is likely to be pivotal—that is, change the outcome of the election (Riker and Ordeshook, 1968). In this model a citizen turns out to vote if and only if

$$pB_i - C_i + D_i > 0 \tag{20}$$

where p is the probability an individual citizen’s vote is pivotal—that is, changes the winner of the election—and B_i is the benefit to the citizen of the citizen’s favored candidate winning over the other candidate. The remaining terms, C_i and D_i , are the instrumental costs and benefits of voting, which are unrelated to the outcome of the election.

It seems reasonable to assume that more-overconfident citizens would over-estimate their probability of being pivotal. This would lead to the prediction that more overconfident citizens would be more likely to turnout.

However, whether or not more ideologically extreme people are more likely to turn out will depend on their utility function. It is well known in the literature on pivotal voting that in order for more ideologically extreme people to be more likely to turn out, utilities need to be very concave: that is, they care much more about small differences in policy when those policies are very far away from their ideal, than when those policies are close to them. Adding overconfidence adds some additional issues: in particular, in order to have more extreme citizens be more likely to turn out the utility function has to be more concave than a quadratic loss function. We have examined a quartic loss function, and even this degree of concavity will not guarantee the result: it holds only for specific parameters and values of the fourth moment of the distribution of beliefs.

Finally, we do not know if it is possible to replicate our conditional predictions about the role of overconfidence and extremeness using a pivotal voter model. As such, it seems that turning in our model for a pivotal voter model would be a poor choice.

Group Utilitarian: In the group-utilitarian framework a citizen votes not just because voting may improve her utility, but because it will improve the utility of others like her as well (Coate and Conlin, 2004; Feddersen and Sandroni, 2006). In these models there is heterogeneity in the costs of voting, and this selects who, from a group, actually turns out. In order to use our model of overconfidence, there needs to be a mapping from beliefs to the cost of voting. An expression for the cost of voting like the left-hand-side of (19) works, and once this is nested in the group-utilitarian framework will produce the same comparative statics as in Proposition 8. This occurs because in the group utilitarian framework those with the lowest costs of voting vote (up to some threshold), and the overconfident, and ideologically extreme, have the lowest costs according to (19). While it would have been possible to use the full group-utilitarian framework in Section 4.2, we felt that, for concision, it was best to avoid that machinery and show directly the important assumption that gives the predictions in that section.

The remaining two models we discuss—like the expressive voting model—focus on the idiosyncratic costs and benefits of turning out to vote. In particular, they focus on large elections where the number of voters grows large, and hence, $p_i \rightarrow 0$.

Regret-Avoidance: Matsusaka (1995) argues that voter turnout is driven in part by whether citizens anticipate they will regret their vote. We view this theory as descriptively accurate: indeed, we ran a survey on a convenience sample using Mechanical Turk, and found that over 60% of respondents reported that they took into account whether they might regret their vote when deciding whether or not to vote. Almost 40% could name someone they regretted voting for.¹

It is straightforward to show that our model is consistent with a model of regret-avoidant

¹For more on regret-avoidance, see Connolly and Zeelenberg (2002), Zeelenberg (1999), Zeelenberg et al. (2001). Models of regret have then been frequently used to explain behavioral patterns which are not compatible with standard, expected-utility, models (Bell, 1982; Loomes and Sugden, 1982; Loomes and Sugden, 1987; Sugden, 1993; and Sarver, 2008). Indeed, Matsusaka’s approach is a direct instantiation of Sugden (1993), applied to politics.

voting. In particular, as $p_i \rightarrow 0$, a citizen's turnout decision depends only on the idiosyncratic, instrumental costs and benefits of voting in (19), C_i and D_i . We decompose the instrumental cost into two parts: direct costs C'_i , such as the opportunity cost of going to vote, and a regret penalty \mathcal{R}_i that accrues if the citizen votes for a candidate whose platform turns out to be worse for the citizen, given the state. That is

$$D_i - C_i \equiv D_i - \mathcal{R}_i \mathbb{I}_{\text{vote=wrong}} - C'_i$$

with D_i , \mathcal{R}_i and C'_i i.i.d. draws from some (possibly different) distributions.² We then have:

Proposition D.1. *In large elections when $D_i - C_i \equiv D_i - \mathcal{R}_i \mathbb{I}_{\text{vote=wrong}} - C'_i$, comparative statics on voter turnout and partisan identification are the same as comparative statics on*

$$\left| \text{Prob}_i[U_R(b_i|x) > U_L(b_i|x)] - \frac{1}{2} \right| - c_i > 0.$$

Proof of Proposition D.1: When elections are large $p \rightarrow 0$ in (20). Supposing citizen i favors candidate R if he or she were to vote, citizen i will vote if and only if

$$\begin{aligned} D_i - \mathcal{R}_i \mathbb{E}[\mathbb{I}_{\text{vote=wrong}}] - C'_i &> 0 \\ \text{Prob}[\text{vote} = \text{wrong}] &< \frac{D_i - C'_i}{\mathcal{R}_i} \\ 1 - \text{Prob}[U_R(b_i|x) > U_L(b_i|x)] &< \frac{D_i - C'_i}{\mathcal{R}_i} \\ \text{Prob}[U_R(b_i|x) > U_L(b_i|x)] - \frac{1}{2} &> \frac{1}{2} - \frac{D_i - C'_i}{\mathcal{R}_i} \equiv c_i. \end{aligned}$$

The absolute value follows from considering the case where i favors candidate L . ■

We chose to display this chain of logic here to simplify and shorten exposition in the text.

Choice-Avoidance: Degan and Merlo (2011) use the same idea as Matsusaka (1995).

²We emphasize that, although we pick a particular formalization, (expected) regret can be seen as either a reduction in the benefit of voting, or an increase in the cost of voting.

However, they note that as it is unlikely that a citizen will discover the actual state, they will not anticipate regretting their decision; instead, they discuss their model in terms of choice-avoidance. It should be clear from the form of (19) that citizens who make their voting decision in this way are choice-avoidant. In particular, a citizen avoids choice unless the choice is clear.³

Appendix D.1.2 Strength of Partisan Identification

Our initial model of strength of partisan identification assumed that citizens would invest in a partisan identity only if they believed there was a sufficiently high probability that they would stay on the same side of the ideological spectrum as they received more signals.

This yields the same predictions as Corollary 9. More overconfident citizens would believe that, with high-probability, future signals would just confirm what they already knew. As such, there is little chance that they would end up on the opposite side of the ideological spectrum. Thus, more-overconfident citizens would be more likely to strongly identify with a party.

More ideologically extreme citizens would know that they would need a more extreme signal that the state is on the other side of the ideological spectrum in order to cross-over to that side. As such, there is little chance they would end up on the opposite side, and they would thus be more likely to strongly identify with a party.

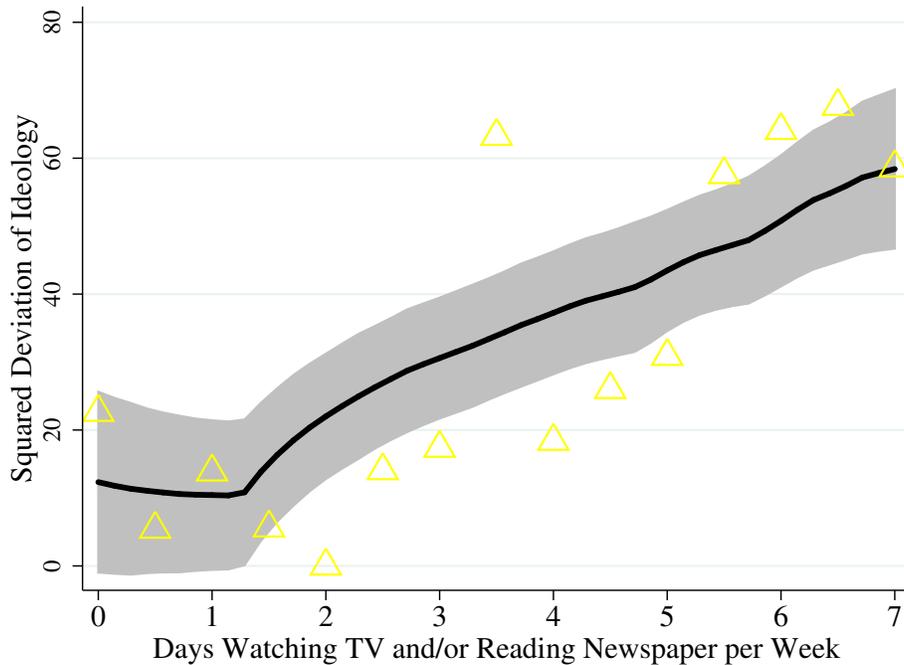
We removed this additional model from the text of the paper in order to simplify and shorten the exposition.

Appendix D.2 Empirical

In the text we present our preferred specifications. Here we provide additional specifications that we excluded from the text for brevity.

³For examples of choice avoidance in other contexts see Iyengar et al. (2004), Iyengar and Lepper (2000), Boatwright and Nunes (2001), Shah and Wolford (2007), Schwartz (2004), Choi et al. (2009), DellaVigna (2009), Reutskaja and Hogarth (2009), and Bertrand et al. (2010).

Figure D.1: Media Exposure and Ideological Extremeness



Notes: Data from the ANES cumulative data file for 1984–2008, all years for which media measures are available. Each point is the average for a particular level of media exposure. Trendiness, in black, and 95% confidence intervals, in gray, use an Epanechnikov kernel with a bandwidth of 1. Graph shows the relationship between residuals, controlling for year-by-income, year-by-education, and year-by-state fixed effects.

Given the centrality of the relationship between media exposure and extremeness in falsifying fully Bayesian alternatives, it is important to show that this relationship is not just a flash in the pan. We note in the text that the relationship between ideological extremeness and media consumption found in Figure 1 can also be found in other datasets. Here we present results from the ANES cumulative data file, discussed in Appendix C. Figure D.1 shows how the squared deviation of ideology evolves with media exposure— $\text{Var}[Z|n]$, in the language of this paper. As can be seen, the general pattern agrees with Figure 1. Note that the scale here is much larger because the baseline scale for ideology here is a 0–100 scale (see Appendix C). This data, taken from 1984–2008, and controlling for year-by-income, year-by-education, and year-by-state fixed effects shows that this is truly a stylized fact that any theory must take seriously.

Table D.1: Ideology, and ideological extremeness is robustly related to overconfidence.

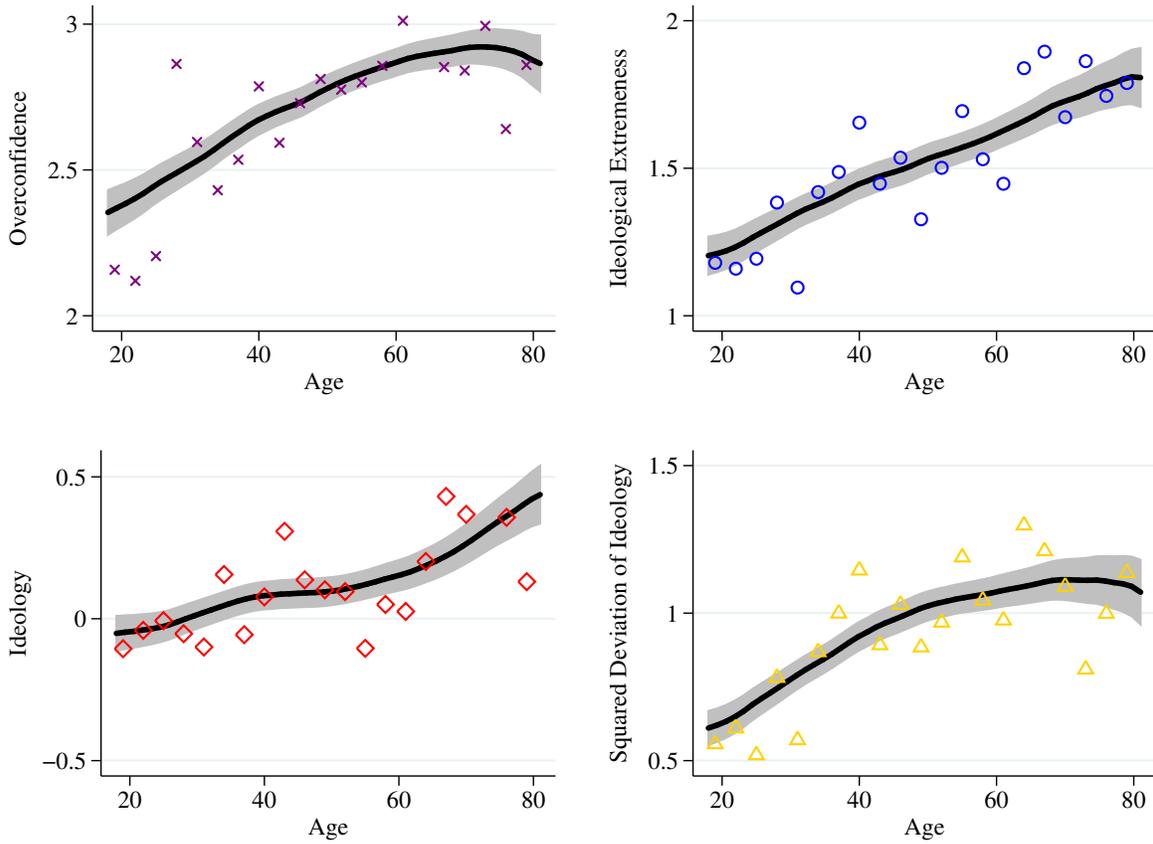
Ideology Measure:	Scaled	Self-Reported Treatment of “Don’t Know”							
		Centrist (0)	Missing (.)						
Panel A: Ideology (Right-Left)									
Overconfidence	0.22*** (.028)	0.22*** (.024)	0.22*** (.019)	0.17*** (.021)	0.25*** (.020)	0.25*** (.020)	0.21*** (.022)		
Economic Controls	Y	Y	Y	Y	Y	Y	Y		
Number of Signals	Y	Y	Y	Y	Y	Y	Y		
R^2	0.047	0.16	0.23	0.048	0.15	0.23	0.057		
N	2,868		2,910				2,754		
Panel B: Extremism (Generated from Right-Left Ideology Purged of Economic Controls)									
Overconfidence	0.23*** (.028)	0.17*** (.027)	0.12*** (.026)	0.26*** (.024)	0.24*** (.024)	0.20*** (.025)	0.28*** (.026)	0.26*** (.026)	0.23*** (.026)
Economic Controls	Y	Y	Y	Y	Y	Y	Y	Y	
Number of Signals	Y	Y	Y	Y	Y	Y	Y	Y	
R^2	0.05	0.19	0.29	0.067	0.084	0.16	0.069	0.083	
N	2,868		2,910					2,754	

Notes: ***, **, * denote statistical significance at the 1%, 5% and 10% level with standard errors, clustered by age (73 clusters), in parentheses. All specifications estimated using WLS with CCES sampling weights.

Next we show the regression results for all three—including the two alternate—measures of ideology that Table 4 is based on. As can be seen, the patterns in Table 3 are also found in the alternative measures. Indeed, in most cases, the results using the alternative measures are more robust.

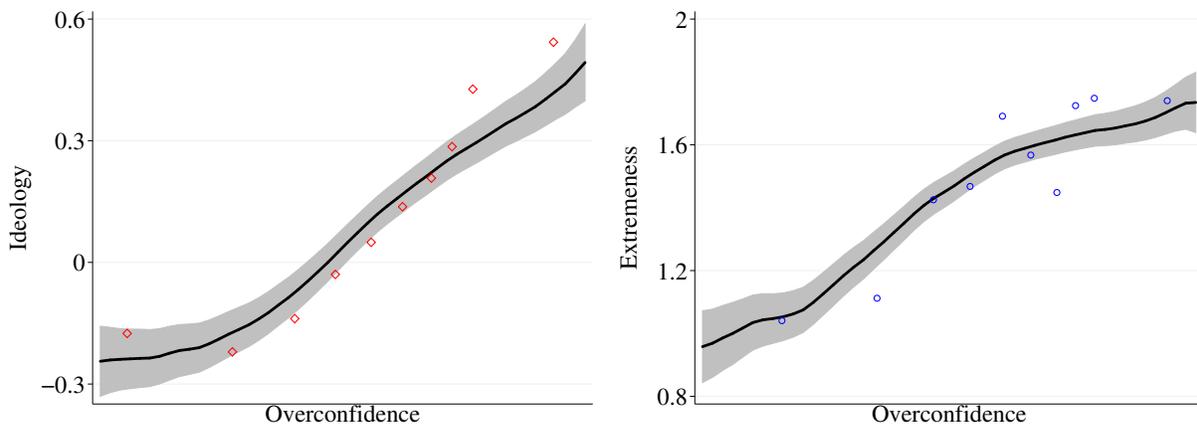
A feature of the theory that we do not emphasize in the text is that our model predicts that many of the relationships in the data should be monotonic. For example, the increase in overconfidence with media exposure and age should be monotonic. Moreover, the relationship between overconfidence and (average) extremeness, and overconfidence and (average) ideology (when $x > 0$) should both be monotonically increasing relationships. We do not discuss those results in the text because the econometrics of testing for monotonicity is still in its infancy (see Patton and Timmermann, 2010, 2012). As such, we prefer to show graphical representations of some of these monotonic relationships. The relationships in Figure 1 appear monotonic; equivalent patterns for age are found in Figure D.2. Figure D.3 shows the relationship between overconfidence, ideology, and extremeness. Figure D.4 shows the results in Table 5 graphically.

Figure D.2: Age, Overconfidence, and Ideology



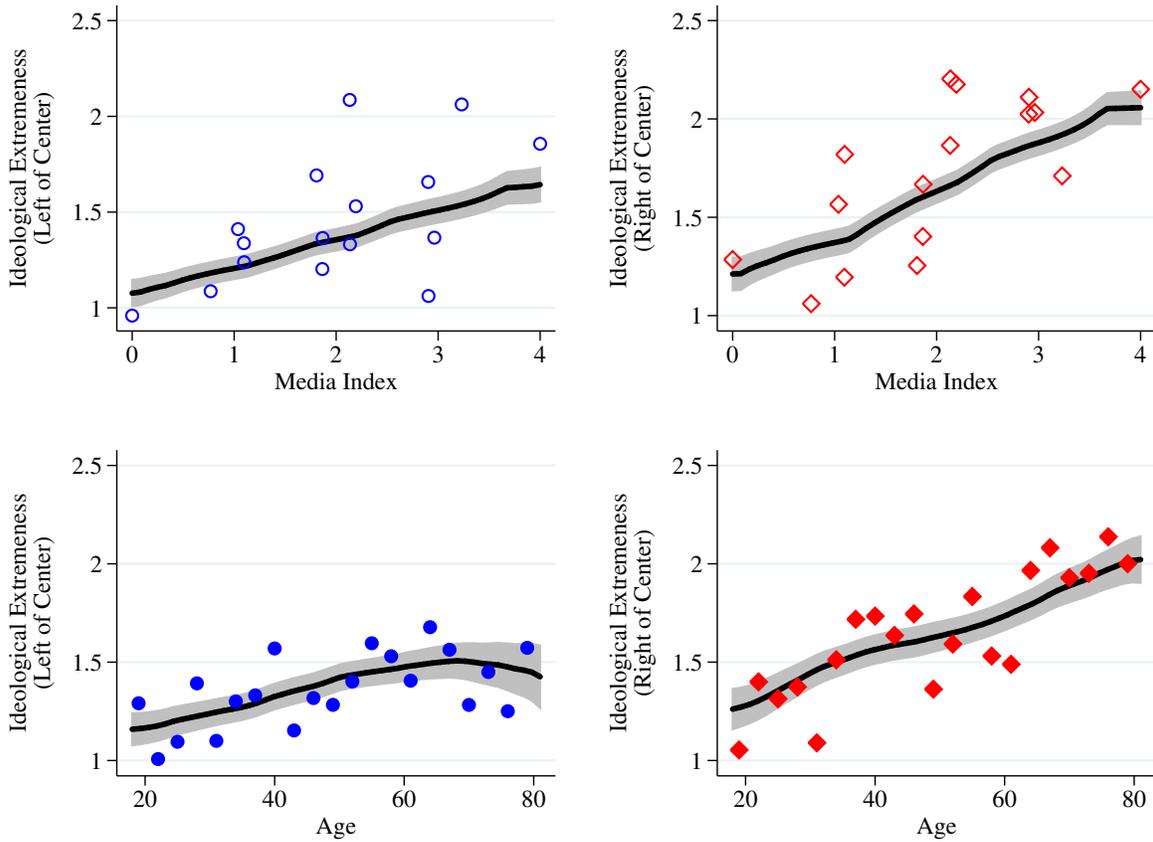
Notes: Each point is the average for three years of age. Trendiness, in black, and 95% confidence intervals, in gray, use an Epanechnikov kernel with a bandwidth of 8.

Figure D.3: The relationship between Overconfidence and Extremeness, and Overconfidence and Ideology Appears to be Monotonic.



Notes: Each point is the average for a decile of overconfidence, or three years of age. Trendiness, in black, and 95% confidence intervals, in gray, use an Epanechnikov kernel with a bandwidth of 0.5.

Figure D.4: There is a greater covariance between extremeness and overconfidence for right-of-center citizens than left of center citizens.



Notes: Each point is the average for a value of the media index, or three years of age. Trendiness, in black, and 95% confidence intervals, in gray, use an Epanechnikov kernel with a bandwidth of 0.8 for media figures, and 8 for age figures.

References

- Alpert, M. and H. Raiffa**, “A Progress Report on the Training of Probability Assessors,” in Daniel Kahneman, Paul Slovic, and Amos Tversky, eds., *Judgement Under Uncertainty: Heuristics and Biases*, Cambridge and New York: Cambridge University Press, 1969/1982, pp. 294–305.
- , **Marc Meredith**, and **Erik Snowberg**, “Mecro-Economic Voting: Local Information and Micro-Perceptions of the Macro-Economy,” 2011. California Institute of Technology, *mimeo*.
- , – , and – , “Asking About Numbers: Why and How,” *Political Analysis*, 2013, *21* (1), 48–69.
- Barber, Brad M. and Terrance Odean**, “Boys will be Boys: Gender, Overconfidence, and Common Stock Investment,” *Quarterly Journal of Economics*, 2001, *116* (1), 261–292.
- Bell, David E.**, “Regret in Decision Making Under Uncertainty,” *Operations Research*, 1982, *30* (5), 961–981.
- , **Dean S. Karlan**, **Sendhil Mullainathan**, **Eldar Shafir**, and **Jonathan Zinman**, “What’s Advertising Content Worth? Evidence from a Consumer Credit Marketing Field Experiment,” *The Quarterly Journal of Economics*, 2010, *125* (1), 263–306.
- Boatwright, Peter and Joseph C. Nunes**, “Reducing Assortment: An Attribute-Based Approach,” *The Journal of Marketing*, 2001, *65* (1), 50–63.
- Brennan, Geoffrey and Alan Hamlin**, “Expressive Voting and Electoral Equilibrium,” *Public Choice*, 1998, *95* (1-2), 149–175.
- Choi, James, David Laibson, and Brigitte Madrian**, “Reducing the Complexity Costs of 401 (k) Participation Through Quick Enrollment,” in David A. Wise, ed., *Developments in the Economics of Aging*, Chicago, IL: University of Chicago Press, 2009, pp. 57–82.
- Coate, Stephen and Michael Conlin**, “A Group Rule-Utilitarian Approach to Voter Turnout: Theory and Evidence,” *The American Economic Review*, 2004, *94* (5), 1476–1504.
- Connolly, Terry and Marcel Zeelenberg**, “Regret in Decision Making,” *Current Directions in Psychological Science*, 2002, *11* (6), 212–216.
- Degan, Arianna**, “Civic Duty and Political Advertising,” *Economic Theory*, 2013, *52* (2), 531–564.
- and **Antonio Merlo**, “A Structural Model of Turnout and Voting in Multiple Elections,” *Journal of the European Economic Association*, 2011, *9* (2), 209–245.
- DellaVigna, Stefano**, “Psychology and Economics: Evidence from the Field,” *Journal of Economic Literature*, 2009, *47* (2), 315–372.

- Feddersen, Timothy and Alvaro Sandroni**, “A Theory of Participation in Elections,” *The American Economic Review*, 2006, 96 (4), 1271–1282.
- Fiorina, Morris P.**, “The Voting Decision: Instrumental and Expressive Aspects,” *The Journal of Politics*, 1976, 38 (2), 390–413.
- Gerber, Alan S., Gregory A. Huber, David Doherty, and Conor M. Dowling**, “Personality and the Strength and Direction of Partisan Identification,” *Political Behavior*, 2012, 34 (4), 653–688.
- , – , – , – , and **Shang E. Ha**, “Personality and Political Attitudes: Relationships Across Issue Domains and Political Contexts,” *American Political Science Review*, 2010, 104 (1), 111–133.
- Grinblatt, Mark and Matti Keloharju**, “Sensation Seeking, Overconfidence, and Trading Activity,” *The Journal of Finance*, 2009, 64 (2), 549–578.
- Iyengar, Sheena S. and Mark L. Lepper**, “When Choice Is Demotivating: Can One Desire Too Much of a Good Thing?,” *Journal of Personality and Social Psychology*, 2000, 79 (6), 995–1006.
- , **Grub Huberman, and Wei Jiang**, “How Much Choice is Too Much? Contributions to 401 (k) Retirement Plans,” in Olivia S. Mitchell and Stephen P. Utkus, eds., *Pension Design and Structure: New Lessons from Behavioral Finance*, Oxford University Press, 2004, chapter 5, pp. 83–95.
- Juslin, Peter, Pia Wennerholm, and Henrik Olsson**, “Format Dependence in Subjective Probability Calibration,” *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 1999, 25 (4), 1038–1052.
- Knight, Brian**, “Evaluating Competing Explanations for the Midterm Gap: A Unified Econometric Approach with Microfoundations,” 2013. Brown University, *mimeo*.
- Loomes, Graham and Robert Sugden**, “Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty,” *The Economic Journal*, 1982, 92 (368), 805–824.
- and – , “Some Implications of a More General Form of Regret Theory,” *Journal of Economic Theory*, 1987, 41 (2), 270–287.
- Matsusaka, John G.**, “Explaining Voter Turnout Patterns: An Information Theory,” *Public Choice*, 1995, 84 (1), 91–117.
- McMurray, Joseph C.**, “Ideology as Opinion: A Spatial Model of Common-value Elections,” 2013. Brigham Young University, *mimeo*.
- Meredith, Marc**, “Persistence in Political Participation,” *Quarterly Journal of Political Science*, 2009, 4 (3), 186–208.
- Moore, Don A. and Paul J. Healy**, “The Trouble with Overconfidence,” 2007. Carnegie Mellon University, *mimeo*.

- Mullainathan, Sendhil and Ebonya Washington**, “Sticking with Your Vote: Cognitive Dissonance and Political Attitudes,” *American Economic Journal: Applied Economics*, 2009, 1 (1), 86–111.
- Patton, Andrew J and Allan Timmermann**, “Why do Forecasters Disagree? Lessons from the Term Structure of Cross-sectional Dispersion,” *Journal of Monetary Economics*, 2010, 57 (7), 803–820.
- and –, “Forecast Rationality Tests Based on Multi-Horizon Bounds,” *Journal of Business & Economic Statistics*, January 2012, 30 (1), 1–17.
- Reutskaja, Elena and Robin M. Hogarth**, “Satisfaction in Choice as a Function of the Number of Alternatives: When “Goods Satiates,”” *Psychology and Marketing*, 2009, 26 (3), 197–203.
- Riker, William H. and Peter C. Ordeshook**, “A Theory of the Calculus of Voting,” *The American Political Science Review*, 1968, 62 (1), 25–42.
- Rothschild, David**, “Expectations: Point Estimates, Probability Distributions, Confidence, and Forecasts,” 2011. Yahoo! Research, *mimeo*.
- Sarver, Todd**, “Anticipating Regret: Why Fewer Options May Be Better,” *Econometrica*, 2008, 76 (6), 263–305.
- Schaefer, Peter S., Cristina C. Williams, Adam S. Goodie, and W. Keith Campbell**, “Overconfidence and the Big Five,” *Journal of Research in Personality*, 2004, 38 (5), 473–480.
- Schmidt, Klaus D.**, “On the Covariance of Monotone Functions of a Random Variable,” 2003. Lehrstuhl für Versicherungsmathematik Technische Universität Dresden, *mimeo*.
- Schwartz, Barry**, *The Paradox of Choice: Why More is Less*, New York, NY: Harper Collins, 2004.
- Shah, Avni M. and George Wolford**, “Buying Behavior as a Function of Parametric Variation of Number of Choices,” *Psychological Science*, 2007, 18 (5), 369–370.
- Snowberg, Erik, Justin Wolfers, and Eric Zitzewitz**, “Partisan Impacts on the Economy: Evidence from Prediction Markets and Close Elections,” *The Quarterly Journal of Economics*, 2007, 122 (2), 807–829.
- , –, and –, “Party Influence in Congress and the Economy,” *Quarterly Journal of Political Science*, 2007, 2 (3), 277–286.
- Soll, Jack B. and Joshua Klayman**, “Overconfidence in Interval Estimates,” *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 2004, 30 (2), 299–314.
- Sugden, Robert**, “An Axiomatic Foundation for Regret Theory,” *Journal of Economic Theory*, 1993, 60 (1), 159–180.

Zeelenberg, Marcel, “Anticipated Regret, Expected Feedback and Behavioral Decision Making,” *Journal of Behavioral Decision Making*, 1999, 12 (2), 93–106.

– , **J. Jeffrey Inman, and Rik G.M. Pieters**, “What we do when Decisions go Awry: Behavioral Consequences of Experienced Regret,” in Elke U. Weber, Jonathan Baron, and Graham Loomes, eds., *Conflict and Tradeoffs in Decision Making*, Cambridge University Press, 2001, pp. 136–155.