Positive, Small, Homogeneous, and Durable: Political Persuasion in Response to Information

Alexander Edwards Coppock

Submitted in partial fulfillment of the
requirements for the degree
of Doctor of Philosophy
in the Graduate School of Arts and Sciences

Columbia University

2016

©2016
Alexander Edwards Coppock
All Rights Reserved

ABSTRACT

Positive, Small, Homogeneous, and Durable: Political Persuasion in Response to Information

Alexander Edwards Coppock

This dissertation offers a theory of political persuasion rooted in a Bayesian model of information processing. I support this theory with the results of 20 survey experiments, conducted variously on convenience samples and nationally-representative surveys. From these data, I draw four main conclusions. First, when confronted with persuasive messages, individuals update their views in the direction of information. Second, people change their minds about political issues in small increments. Third, persuasion in the direction of information occurs regardless of background characteristics, initial beliefs, or ideological position. Finally, these changes in political attitudes are not ephemeral, in most cases lasting at least 10 days.

These findings stand in contrast to the predictions of the attitude polarization literature, which posits that the effects of persuasion attempts will be positive for some, but negative for others. Across these 20 experiments, I was unable to find any evidence of negative effects (defined as attitude changes away from treatment information) for any subgroup defined by standard demographics, prior attitudes, or their intersections. Instead, people appear to update their views in a manner consistent with Bayes' rule, i.e., as a weighted average of prior beliefs and new information.

Table of Contents

List of Figures

List of Tables

Acknowledgments

I am indebted to a great many individuals for their support throughout the process of producing this dissertation. In particular, I would like to thank my advisor and mentor, Donald Green, whose infectious enthusiasm for experimental research is the principal cause of my academic ambitions (though I can't be sure); the members of my dissertation committee, James Druckman, Robert Erikson, Shigeo Hirano, and Robert Shapiro for their guidance and encouragement; my longtime collaborator Andrew Guess, who helped shape my thinking on many of the central ideas advanced here.

During all stages of this project, but especially at the end, I have been filled with gratitude for my wife and partner, Penelope Van Grinsven, whose idea of fun includes spending many hours happily working in parallel in her studio.

Throughout my (now very long) education, I have been fortunate to have the support of three sets of loving parents: Martha Edwards and Robert Green, Bruce Coppock and Lucia May, and Diane and Charles Van Grinsven. My sister Elizabeth Coppock and my aunt Jane Coppock, in addition to being loving and supportive family members, were also early readers of this project; their critical feedback greatly improved the final product.

I wish also to acknowledge the assistance of Peter Aronow, a great source of statistical advice and more, as well as David Kirby and Emily Ekins, my collaborators on the Newspapers study discussed in Chapter 5.

I owe a debt to the many hundreds of developers, programmers, and statisticians who donate their time and expertise to the `R` ecosystem of open source statistical software (R Core Team 2015). Hadley Wickham's `dplyr` package (Wickham and Francois 2015) was indispensable for manipulating complex data structures. I used Wickham's `ggplot2` package (Wickham 2009) for all figures and Marek Hlavac's `stargazer` package (Hlavac 2014) for all regression tables. Thomas Leeper's `MTurkR` (Leeper 2015) made the panel design of the studies conducted on Mechanical Turk possible.

This dissertation was supported by a Columbia University Dissertation Development Grant and by the Time-sharing Experiments for the Social Sciences organization.

Preface

The overarching claim of this dissertation is that information designed to persuade can and does change minds. The persuasive effects of information have four salient features:

1. Effects are nearly uniformly *positive*: individuals are persuaded in the direction of evidence.
2. Effects are *small*: changes in opinion are incremental.
3. Effects are relatively *homogeneous*: regardless of background, individuals respond to information by similar degrees.
4. Effects are *durable*: at a minimum, effects endure for weeks, albeit somewhat diminished.

These four conclusions are drawn from 20 survey experiments conducted on both convenience and nationally-representative samples. They stand in contrast to theories of "minimal effects" that claim attitudes are largely unmoved by the messages transmitted via the media. These conclusions are also at odds with theories of motivated reasoning that assert that individuals only hear what they want to hear.

The evidence assembled in this dissertation leads me to conclude that political persuasion is a commonplace occurrence. The political attitudes and beliefs of Americans, as expressed in public opinion surveys, respond to the introduction of new evidence, arguments, and information. Persuasive effects of a single-shot dose of information, delivered in a controlled survey environment to willing participants, are small but real. Subjects update their attitudes and beliefs *at the margin*. In no cases did I measure an effect that would turn an ardent Democrat into a staunch Republican. Persuasion, as it unfolded in the studies detailed here, did not make people re-evaluate their opinions top-to-bottom; instead, persuasion consisted in the small updates that take a person, say, from a 2 to a

2.5 on a 7-point scale from Strongly Disagree to Strongly Agree. Put another way, the average effect size in these experiments is 0.21 standard units. A widely-adopted rule of thumb suggests that small, medium, and large social science effects are approximately 0.2, 0.5, and 0.8 standard units, respectively (Cohen 1992). By this rubric, persuasive effects are small, or at least they are small relative to the variance in attitudes held by people of diverse backgrounds and experiences.

I explain these results with a theory of Bayesian reasoning, which posits that people take a weighted average of their prior opinions and new information. This theory makes two basic predictions. The first is that people update their opinions in the direction of evidence. The second is that, under some basic assumptions about how strongly held prior opinions are, the effects of information are approximately the same for most people. This basic sameness leads to an optimistic view of the democratic process. Even in the most contentious of political disputes, we see that persuasion can does occur, on all sides of the debate.

Research Approach

The classic definition of an attitude given by Allport (1935) holds that "An attitude is a mental and neural state of readiness, organized through experience, exerting a directive or dynamic influence upon the individual's response to all objects and situations with which it is related." In this dissertation, no attempt will be made to model, account for, or otherwise explain this baseline state of readiness. Attitudes are the result of extraordinarily complex social processes involving family context, religious upbringing, political socialization, educational attainment, friendship networks, media exposure, occupational experience, and the interactions among these factors, not to mention individuals' idiosyncratic application of personal judgment and reason; untangling this web while successfully distinguishing spurious correlations from causal relationships is, simply put, too difficult.

Instead, I will focus on persuasion, which refers to the *change* in attitudes in response to some treatment. The study of persuasion is particularly well-suited to the experimental research paradigm. There is a strong match between the research question, "How does opin-

ion change in response to information?" and the research approach of randomly assigning information to some subjects but not others and then comparing the expressed attitudes of the treated and the untreated.

Randomized experiments are often described as the "gold standard" of research designs because of their uniquely strong licensing of causal inferences. Setting aside whether the analogy to monetary policy is appropriate, the basic point stands: whereas causal inferences drawn from non-experimental research rest on assumptions that are often difficult to verify, randomized experiments justify their assumptions by design (where possible). The crucial assumption of the independence of potential outcomes from exposure to information is established by a description of the physical random assignment procedure rather than through a statistical control strategy such as regression or matching. In my experiments, the assumption of non-interference is justified by the design feature that subjects all participate in private settings where communication with other subjects is impossible; in non-experimental studies of the persuasive effects of information delivered in the media, the "subjects" can and do talk to one another about the news, possibly (probably!) affecting outcomes.

Randomized experiments can also be thought of as a tool for measuring an otherwise difficult-to-measure quantity: the difference between two mutually-exclusive states of the world. In one state, everybody is exposed to the persuasive message; in the other state, nobody is. If we could measure the world in both states, there would be no need for randomization – we could simply inquire about the differences between the two states. Thinking of randomization as a measurement tool places special emphasis on the goal of an experiment: to measure an unobservable but nevertheless real, fixed characteristic of subjects. We use surveys or administrative data to measure most fixed characteristics about people, such as their age, political ideology, or policy attitudes. Unfortunately, we cannot use surveys to directly measure individuals' response to treatment. For that, we have to turn to experiments and content ourselves with estimating group-level average treatment effects, not individual treatment effects.

Outline of Chapters to Follow

Chapter 1 explores the micro foundations of the parallel publics hypothesis (Page and Shapiro 1992). Republicans and Democrats hold divergent views in many domains, but evidence from repeated cross-sectional surveys suggests that these groups respond to changing political conditions in a parallel fashion. Evidence from a large set of survey experiments confirms that the probable mechanism for this parallel movement is that partisans of different stripes update their attitudes by approximately the same amount.

Chapter 2 is the theoretical heart of this project. The Bayesian reasoning hypothesis suggests that individuals change their attitudes about political beliefs by updating in the direction of evidence. The principal competitor to this hypothesis is some form of motivated reasoning. These two theories – motivated reasoning and Bayesian reasoning – make divergent predictions about how individuals should react to counter-attitudinal evidence and arguments. Under motivated reasoning, counter-attitudinal evidence causes individuals to hold their prior opinions even more strongly; under Bayesian reasoning, everyone updates in the direction of evidence, counter-attitudinal or not. Evidence from three survey experiments designed to elicit "attitude polarization" suggests that at minimum, no one updates away from evidence.

Chapter 3 addresses a potential critique of the evidence presented in Chapter 2 in favor of the Bayesian reasoning hypotheses: that because subjects are answering survey questions in a neutral environment, the conditions for backlash and attitude polarization are not met. In an attempt to assess the causal impact of the "conditions for backlash," I randomize whether subjects are in a contentious processing condition before being randomly assigned to information. The evidence from this experiment is in line with the results of Chapter 2: regardless of the processing condition, subjects update in the direction of evidence.

The finding that most individuals, regardless of predisposition, update their attitudes and beliefs in the same direction and by approximately the same amounts has a fascinating implication for survey experimental methodology: persuasion experiments conducted on nationally representative samples and convenience samples should arrive at very similar answers. Chapter 4 demonstrates that this implication is borne out: survey experiments

originally conducted on probability samples replicate very nicely on non-probability samples. The explanation offered in this chapter is that treatment effect heterogeneity is low in persuasion experiments.

Chapter 5 presents the results of an effort to measure the effects of persuasive treatments over time. The persistence of treatment effects appears to vary from 20% to 60% after 10 days, depending on features of the persuasive messages. Messages that prime existing considerations were found to have relatively fleeting effects, whereas treatments that provided new information were found to exhibit stronger persistence.

The majority of the empirical claims made in this dissertation rest on a set of 20 survey experiments. Because the focus of each chapter is usually on summarizing the pattern of evidence found across experiments, some details of the experiments are omitted in the main text. Appendix A gathers together the complete descriptions of each study, including subject pools, randomization procedure, outcome question wordings, and basic results.

The main theoretical antagonist of this dissertation is attitude polarization, the idea that the effects of an information treatment could be positive for some but negative for others. This idea was given its first empirical test in the classic Lord, Ross and Lepper (1979) demonstration. In Appendix B, I present a reanalysis and replication of that original study to show how errors in design and analysis led those authors to draw incorrect inferences about the effects of information on attitudes.

1 The Micro Foundations of Parallel Publics

The parallel publics thesis holds that the opinions of subgroups in society tend to change in tandem. Although this pattern has been demonstrated by many scholars using time-series cross-sectional survey evidence, the causal engine driving the parallel publics finding is ambiguous. Using knowledge of political history, analysts can sometimes offer plausible reasons for why public opinion changed over time as it did; sometimes no convincing explanation is forthcoming. In this chapter, I offer evidence in support of the parallel publics thesis that derives from a different research design: survey experiments conducted on both convenience and probability samples. Drawing on results from 20 survey experiments, I find highly correlated treatment effects across politically salient subgroups defined by partisanship, education, and gender. These experiments shed light on a longstanding hypothesized explanation for the parallel publics finding; namely, that attitudes move in parallel because all segments of the population are exposed to, receive, and accept the same messages.

The parallel publics thesis, as originally formulated by Page and Shapiro (1992) in *The Rational Public*, makes a strong prediction about how public opinion evolves over time: "Among most groupings of Americans, opinions tend to change (or not change) in about the same manner: in the same direction and by about the same amount at about the same time (318)." Page and Shapiro support this conclusion with evidence from an over-time analysis of thousands of opinion-subgroup dyads. Almost invariably, opinion shifts occur in tandem.

Table 1.1 collects together 31 studies that use either repeated cross-section or panel survey designs to test the parallel publics claim.[1] Fully 30 of these 31 confirm the basic claim that the average opinions of subgroups move in parallel. The range of demographic splits explored by these studies is large, including education (Nincic 1997; Erikson, MacKuen and Stimson 2002; Enns and Kellstedt 2008; Soroka and Wlezien 2008; Gillion, Ladd and Meredith 2015), gender (Nincic and Nincic 2002; Kellstedt, Peterson and Ramirez 2010; Eichenberg and Stoll 2012; Gillion et al. 2015), partisanship (Green, Palmquist and Schickler 2002; Soroka and Wlezien 2008; Snyder, Shapiro and Bloch-Elkon 2009; Enns and McAvoy 2012; Brooks and Manza 2013; Huxster, Carmichael and Brulle 2015; Morgan and Kang 2015), sophistication (Erikson et al. 2002; Enns and Kellstedt 2008), ideology (Erikson et al. 2002; Degeorges and Gonthier 2012; Huxster et al. 2015), and race (Nincic and Nincic 2002; Kellstedt 2003). Three of these studies cover a large number of demographic splits including those just mentioned as well as others like age and religion (Page and Shapiro 1992; DiMaggio, Evans and Bryson 1996; Gonthier 2015). Some studies document parallel trends according to geography rather than demographics: (Enns and Koch 2013; Pacheco 2014). While most studies were conducted with American subjects, two studies (Degeorges and Gonthier 2012; Gonthier 2015) used European samples. In only one of these studies was

[1]These studies were gathered by searching Google Scholar for the search string "parallel publics" and were limited to studies that present estimates of opinion over time by subgroup. Some studies (Mayer 2004; Taydas, Kentmen and Olson 2012, e.g.,) use cross-sectional design with no over time component, typically exploring the "impact" of some covariate on opinion in different subgroups. These were excluded on the grounds that the research designs were ill-suited to uncovering the causal effect of the covariate on opinion, rendering the question of parallel effects difficult to interpret.

any evidence of contrary motion uncovered: Brooks and Manza (2013) found that support for government regulation increased over time for Democrats but decreased over time for Republicans, though the robustness of this finding depends on the time frame examined. The remaining studies all found clear evidence in favor of parallel publics. The range of topics covered is too wide to list here, but they are indicated in the last column of Table 1.1. These opinion items cover nearly every major issue area that has been of interest to political and other social scientists over the past 30 years. The parallel publics pattern uncovered by Page and Shapiro in 1992 has been replicated across enough contexts, places, and times to earn the description "empirical regularity."

Why do opinions move in parallel? Page and Shapiro explain their findings with a conjecture that the root cause is "centralization of the mass media – all segments of the population receive essentially the same information at the same time." Whether or not this hypothesis is true is difficult to demonstrate with survey data. First, it is unclear whether all segments of the population do in fact receive similar messages. Some groups report relatively low media exposure while others follow the news minute-by-minute. Second, the media have become dramatically decentralized in recent years, but there has been no corresponding decrease in the validity of the parallel publics description of American politics (as evidenced by the studies in Table 1.1), so it may not be that media centralization *per se* is the cause (or was for the period studied by Page and Shapiro). Finally, the time-series cross-sectional design used by many scholars to demonstrate parallel trends in public opinion is limited in its ability to draw inferences about the causes of opinion shifts. We would like attribute changes in public opinion to changes in the information environment generated by the media. Unfortunately, the independent variable in these studies is time. While it is true that the information environment changes as a function of time, so too does every other possible explanation: the economy, long-term demographic change, international tensions, among many others.

The approach I will take in this chapter is to control the information environment to which individuals are exposed in a series of survey experiments. By randomly assigning information to some subjects but not others, I am able to measure the impact of information

Table 1.1: 31 Studies of Parallel Publics

Study	Demographic Split	Opinion Item
Page and Shapiro (1992)	Gender, Race, Education, Occupation, Income, Religion, Age, Partisanship, Region, and Community	169 repeated opinion questions
Bartels (1994)	Political Information	Defense spending
DiMaggio et al. (1996)	Age, Education, Gender, Race, Religion, Ideology, and Partisanship	Sexuality, Racial attitudes, Aid to minorities, Feelings toward poor, Gender roles, Crime and Justice, Abortion, Divorce law, Sex education
Nincic (1997)	Education	Internationalism
Erikson et al. (2002)	Education, Sophistication, Ideology	Policy mood
Green et al. (2002)	Partisanship	Presidential approval
Nincic and Nincic (2002)	Gender, Race	Alienation
Kellstedt (2003)	Race	Busing
Enns (2007)	Education	Defense and welfare spending
Enns and Kellstedt (2008)	Education, Sophistication	Policy mood
Soroka and Wlezien (2008)	Income, Education, Partisanship	Defense spending, Foreign aid, Education spending, Health spending, Welfare spending, Cities, Crime, Environment, Taxes
Ura and Ellis (2008)	Income	Policy liberalism
Snyder et al. (2009)	Partisanship	Military power, the UN, International terrorism
Kelly and Enns (2010)	Income	Policy Mood
Kellstedt et al. (2010)	Gender	Policy liberalism
Enns and Wlezien (2011)	Income, Education, Partisanship, Race	Welfare, Taxes, Mood, Policy liberalism
Wlezien and Soroka (2011)	Income	Welfare, heath, education, environment and crime spending
Ellis and Ura (2011)	Education, Income	Scope of government, Cultural preferences
Degeorges and Gonthier (2012)	Ideology	Economic liberalism
Eichenberg and Stoll (2012)	Gender	Defense spending
Enns and McAvoy (2012)	Partisanship	Economy ratings
Hopkins (2012)	Income	Economy ratings
Enns and Koch (2013)	U.S. State	Policy mood
Brooks and Manza (2013)[1]	Partisanship	Government regulation
Johnson and Kellstedt (2014)	Newspaper reading, TV watching	Policy mood
Pacheco (2014)	U.S. State	Partisanship, Ideology, Education spending, Welfare spending, Death penalty, Abortion, Presidential approval, Consumer sentiment
Gillion et al. (2015)	Education, Gender	Partisanship
Gonthier (2015)	Gender, Age, Employment status, Income, Education, Ideology, Political interest	Economic liberalism
Huxster et al. (2015)	Partisanship, Ideology	Climate change
Morgan and Kang (2015)	Partisanship	Healthcare spending
The present study	Partisanship, Gender, Education	Death penalty, Gun control, Welfare, Racial attitudes, Presidential approval

[1] Brooks and Manza (2013) is unique among these studies in finding evidence of contrary motion.

on attitudes, subgroup by subgroup. If the Page and Shapiro explanation of parallel publics is correct, then the experimental estimates of the effects of information treatments should be approximately the same among different groups of people.

Two alternative theories of information processing stand in contrast to parallel publics. These theories make predictions about treatment effect heterogeneity. They predict that information, or certain types of information, will be especially effective (or ineffective) among particular subgroups. The first of these is, broadly, the theory of motivated reasoning. The term motivated reasoning encompasses a large body of work, not all of which is mutually consistent (Kunda 1990). The portions of motivated reasoning that are relevant here are the interrelated concepts of perceptual bias, biased assimilation, disconfirmation bias, and attitude polarization. Within political science, the notion of perceptual bias can be traced back at least as far as the American Voter (Campbell, Converse, Miller and Stokes 1960), which claimed that "Identification with a party raises a perceptual screen through which the individual tends to see what is favorable to [his or her] partisan orientation (p. 133)." Presumably, by "tends to see," these authors meant that individuals receive, accept, and hence are influenced by only messages that are good for their partisan group. Stated in terms of causal effects, the perceptual screen argument predicts no effect of negative information for Democrats among Democrats, but positive effects of positive information.

Within psychology, the first demonstration of this sort of biased assimilation of information is due to Lord et al. (1979). That study claimed that proponents of capital punishment became more pro-death penalty in response to mixed information, while the opposite occurred among opponents of capital punishment. In my estimation, the empirical foundations of Lord et al. (1979) are weak; the authors mistake the correlation between subjects' background characteristics and their post-information responses for the causal effect of information. For much more detail on this study and how exactly the authors draw incorrect inferences from their data, see Chapter 2 and Appendix B, for a reanalysis of the original study and a replication and extension conducted in 2014. Regardless of whether the conclusions of Lord et al. (1979) are correct, that study has inspired a long tradition seeking to explain the "attitude polarization" effect. Common explanations include biased

assimilation, in which counter-attitudinal information is simply discounted, similar to the "perceptual screen" argument of Campbell et al. (1960). Another possible mechanism is disconfirmation bias (Edwards and Smith 1996; Taber, Cann and Kucsova 2009) by which individuals actively seek out counterarguments to counter-attitudinal information, thereby reinforcing their original positions. Another way of describing attitude polarization is that information may have "backlash" effects among some segments of the population. If this theory were true, we should see negative effects (relative to the direction of the information) among those for whom the information is counter-attitudinal and positive effects among those for whom the information is pro-attitudinal.

The second body of theory that differs from the parallel publics view of information processing focuses on political knowledge as a major moderator of framing and information effects. Althaus (2003) predicts that those with high political information will be "less prone (p. 167)" to question wording effects.[2] This view holds that doses of *new* political information should have smaller effects among the knowledgeable, both because they already have access to the relevant information and the have more stable (read: less influencable) opinions generally.

Both strands of theory are well-represented in Zaller's 1992 Receive-Accept-Sample (RAS) model of political information processing and attitude change. In this model, survey responses are the product of sampling (the "S" in RAS) from a set of considerations held in the mind of the survey respondent. If the preponderance of considerations are liberal, the survey response will tend to be liberal. Importantly, in this model, the survey response

[2]In some respects, Althaus (2003) is an extended argument against the notion that collectively (but possibly rationally) ignorant public opinions represent true policy preferences well. Althaus defines an "information effect" differently from how I am using it here. I am using the term to refer to the causal effect of information on opinions, whereas Althaus is using information effect to describe the simulated difference in average opinion that would obtain if uninformed people held the same views as informed people who are otherwise similar to them on other observable background characteristics. In essence, Althaus is confirming that the informed and uninformed hold different policy views. I do not contest this point, but do note that this some (possibly large) portion of this difference is likely to reflect unmeasured or unobservable differences between the informed and uninformed rather than the causal effect of information.

is the result of a stochastic process, which can explain instability in survey responses. The process of attitude change in response to a persuasive message depends on the extent to which subjects receive and subsequently accept the message. Messages that are not received cannot be persuasive, nor can messages that are not accepted.

From this model, Zaller derives two axioms that are of special importance here. The first is the reception axiom, which states that "The greater a person's level of cognitive engagement with an issue, the more likely he or she is to be exposed to and comprehend – in a word, to receive – political messages concerning that issue." This level of cognitive engagement (referred to as "political awareness" throughout Zaller 1992) is proxied for with tests of political knowledge (p. 43) and sometimes education (Price and Zaller 1993). Zaller predicts that as political awareness increases, so too should reception. Therefore, holding the acceptance of the message constant (which may not be possible), people who are more politically aware should be more persuadable than the politically unaware. If this axiom holds, we should not see parallel updating. We should see no updating among the politically unaware, but evidence of opinion change only among the politically aware. Althaus (2003) also suggests that political knowledge should moderate persuasion, but offers the opposite prediction.

The second axiom holds that, "People tend to resist arguments that are inconsistent with their political predispositions. . . ." (p.44) This axiom concerns the "acceptance" process in the RAS model. I am interpreting "resist" to mean that people do not update their attitudes in reponse to counter-attitudinal information, i.e., that under this theory, we should see small to zero effects of counter-attitudinal information but positive effects of pro-attitudinal information. With these two axioms, Zaller's theory combines both the "political knowledge matters" and the "perceptual bias" critiques of parallel publics (Shapiro and Bloch-Elkon 2008).

Figure 1.1 displays the joint distribution of treatment effects under four theories of information effects. Hypothetical (average) repsonses to treatments are plotted for two subgroups of the population, one on the vertical axis and the other on the horizontal. Suppose that these subgroups divide the population according to some demographic characteristic.

This characteristic could be partisanship, gender, race, education, or any other division. In order to keep things general, imagine that the group on the horizontal axis are "pessimists" and those on the vertical axis are "optimists".

Each point represents a pair of treatment effects that describe how optimists and pessimists respond to a particular persuasion treatment. The treatments themselves are *signed* in the sense that negatively signed treatments "should" engender negative treatment effects and positively signed treatments "should" cause positive treatment effects. I place "should" in quotation marks because the sign of the effects cause by these treatments is precisely where these theories differ.

The top left panel describes the responses under the theory of parallel updating. Both pessimists and optimists are moved in positive directions by positive treatments and in negative directions by negative treatments. The treatment effects are highly correlated across groups. The negative treatments that cause large (in magnitude) negative shifts for pessimists also cause equivalently large changes for optimists. In short, under parallel updating, everyone moves in response to treatment by approximately the same amount and in the same direction.

Contrast this pattern with the effects in the lower left hand panel, labeled "Resistance" Under this theory, individuals are responsive to messages they agree with, but resist messages they disagree with (predicted by Campbell et al. 1960; Zaller 1992) Optimists respond to positive messages but not to negative ones; the reverse holds for pessimists. This theory produces a backwards "L" shape: Optimists move in a positive direction or not at all; pessimists move in a negative direction, or not at all. If this theory were true at the micro level, we would be unlikely to see parallel updating at the macro level. Instead, we would see different groups moving in response to different stimuli at different times.

Continuing in a counterclockwise direction, consider the lower right hand panel labeled "Backlash." This panel represents the theory of information processing that posits that counterattitudinal information is not only resisted, it actually causes a polarization of opinions (predicted by Lord et al. 1979; Taber and Lodge 2006). When confronted with negative information, an optimist has a positive treatment effect; when confronted with postive in-

formation, a pessimist has a negative treatment effect. In this world, people *never* update their views in a direction that is contrary to their prior beliefs. Pessimists always update in a negative direction and optimists always update in a positive direction. A distinctive feature of the backlash prediction is the negative correlation of treatment effects across subgroups.

Thus far, I have been using positive and negative to stand in for the usual dimensions of political conflict. Optimists and Pessimists are analogous to groups that may find themselves (on average) on different sides of some political conflicts: conservatives and liberals, Republicans and Democrats, the rich and the poor, blacks and whites, the educated and the uneducated, and so on.

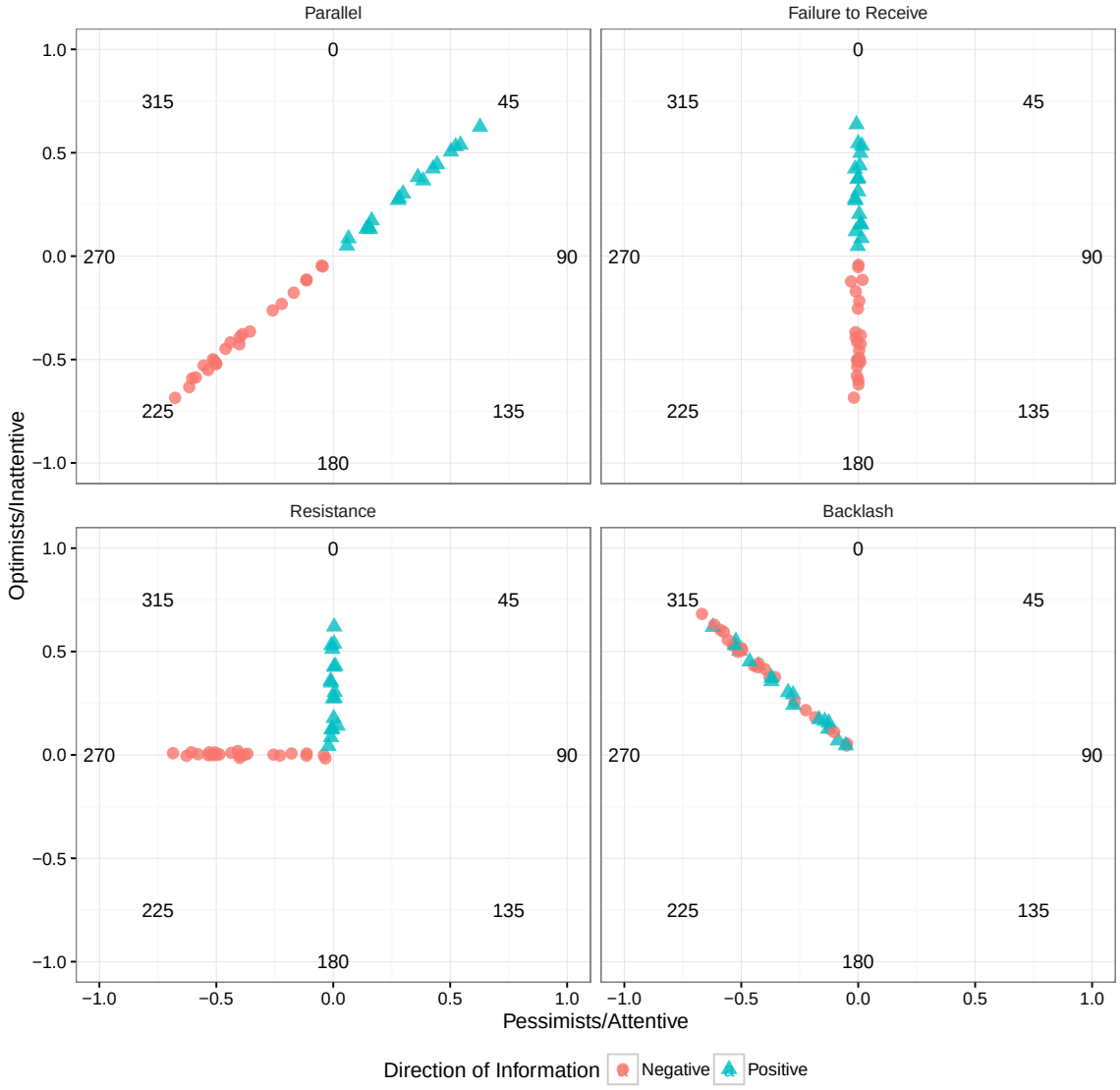
In the panel labeled "Failure to Receive," imagine that the analogy is not to groups on either side of a political conflict, but instead to groups that differ in their attentiveness in Zaller's sense of propensity to receive political messages. The attentive group, which is plotted on the vertical axis, receives all messages and responds in the correct direction for each treatment. The inattentive group, by contrast, never changes its views, because it never receives the message. This panel also corresponds to the prediction of Althaus (2003), but with the axes switched: larger effects are predicted for the inattentive because their political opinions are more malleable.

These four examples are highly stylized versions of each theory, and one might imagine that the truth is some mixture. Fix the point (0, 1) at zero degrees and the point (0, -1) at 180 degrees. Under parallel updating, the effects of the positive treatments (the triangular points) run from the origin to the 45 degree mark; the effects of the negative treatments (the circular points) fall along a line from the origin to the 225 degree mark. For any dose of motivating reasoning from none at all to resistance through backlash, imagine sweeping the segment of triangular points from 45 degrees over to 315 degrees while sweeping the circular points from 225 degrees to 315 degrees. Any mixture of parallel updating and motivating reasoning can be depicted as "folding" these two segments from fully open in parallel updating to fully closed under motivated backlash.

Similarly, any mixture of parallel updating and differential receptivity can be described

as a progressively steepening slope. The slope is 1 in the upper left hand panel and infinite in the upper right hand panel; slopes in between these values could represent different levels of receptivity in the inattentive group.

Figure 1.1: Joint Distributions of Treatment Effects Under Four Theories



If the micro foundations of the parallel publics finding lie in the parallel responses to the same information, then when we conduct randomized experiments to measure the responses

of different subgroups to persuasion, the pattern of results should look like the upper-left hand panel of Figure 1.1. If, however, the results look like any of the other three patterns, then we will have to reconsider the roots of the parallel publics trends.

Before proceeding to new evidence on parallel publics, a brief digression on the importance of measurement for this discussion. Disagreements about just what the "parallel" in parallel publics means have cropped up from time to time since the publication of *The Rational Public*. For example, in summarizing their finding that the policy liberalism of men and women appears to move together, Kellstedt et al. (2010) write: "Scholars who argue for parallel movement assume that sub-aggregate publics move in response to the same stimuli, in the same direction, and by the same magnitude. Although we find the first two assertions to be true – men and women both move in the same direction to changes in public policy – the third requirement for parallel opinion movement is challenged by these results."

Unfortunately, arguments on either side of this third requirement depend crucially on the underlying scale of the outcome. Effects that are homogeneous in one scale (e.g., dollars) would be heterogeneous in another scale (e.g., log-dollars). Ding, Feller and Miratrix (2015) provide a cogent discussion of this problem (p. 4) and a proof of a theorem by G.E.H. Reuter that shows if the CDFs for the treated and untreated units do not cross, there exists some monotone transformation of the outcome that renders the treatment effects constant. By the same token, most demonstrations of "parallel" movement can be shown to not be parallel by some monotone transformation of the outcome (an exception would be if the untreated and treatment outcomes were identical for all units, in which case the movements would be parallel regardelss of transformation).

As an illustration of this problem, consider two groups A and B. A moves from 5% to 10% support while B moves from 50% to 55%. In percentage *points*, these shifts are homogeneous. However, in percentage terms, the former shift is much larger: $100\% = \frac{10\%-5\%}{5\%}$ versus $10\% = \frac{55\%-50\%}{50\%}$. Other transformations are possible: If 5% support in a group is taken to mean that all members of that group have a 0.05 probability of support, then we can express the same average support with a probit transformation, $-1.64 = \Phi^{-1}(0.05)$, where

Φ^{-1} indicates the inverse of the standard normal cumulative distribution function. In probits, the change from 5% to 10% would be $0.36 = \Phi^{-1}(0.10) - \Phi^{-1}(0.05)$, whereas the change from 50% to 55% would be much smaller, at $0.13 = \Phi^{-1}(0.55) - \Phi^{-1}(0.050)$. The appropriate outcome scale will differ from domain to domain. One situation in which scaling is not an issue arises when groups A and B have very similar baseline outcomes. In such a case, changes relative to baseline can be directly compared, regardless of scale.

An agnostic approach to this problem is to fall back on the "weak form" of the parallel publics, which holds that among most groupings of Americans, we do not observe *contrary* motion. People update either not at all or in the same direction as others. The weak form of parallel publics is satisfied if the conditional average treatment effects (CATEs) do not have opposite signs: $(CATE|A \geq 0 \wedge CATE|B \geq 0) \vee (CATE|A \leq 0 \wedge CATE|B \leq 0)$. The only pattern that is not consistent with the weak form is contrary motion: $(CATE|A > 0 \wedge CATE|B < 0) \vee (CATE|A < 0 \wedge CATE|B > 0)$.

In the remainder of this chapter, I will first present a brief update to the macro level trends to confirm that the parallel publics finding holds through 2016. I will then proceed to the micro-level evidence of parallel updating I have found in a set of 20 survey experiments in which I randomly exposed some subjects to information but not others. I show that when all segments of the population receive the same information, they update their views in parallel. As a preview of the results to come, I find that the pattern of results strongly resembles the pattern predicted by the parallel publics theory in the upper-left hand panel of Figure 1.1.

1.1 Time-Series Cross-Sectional Evidence

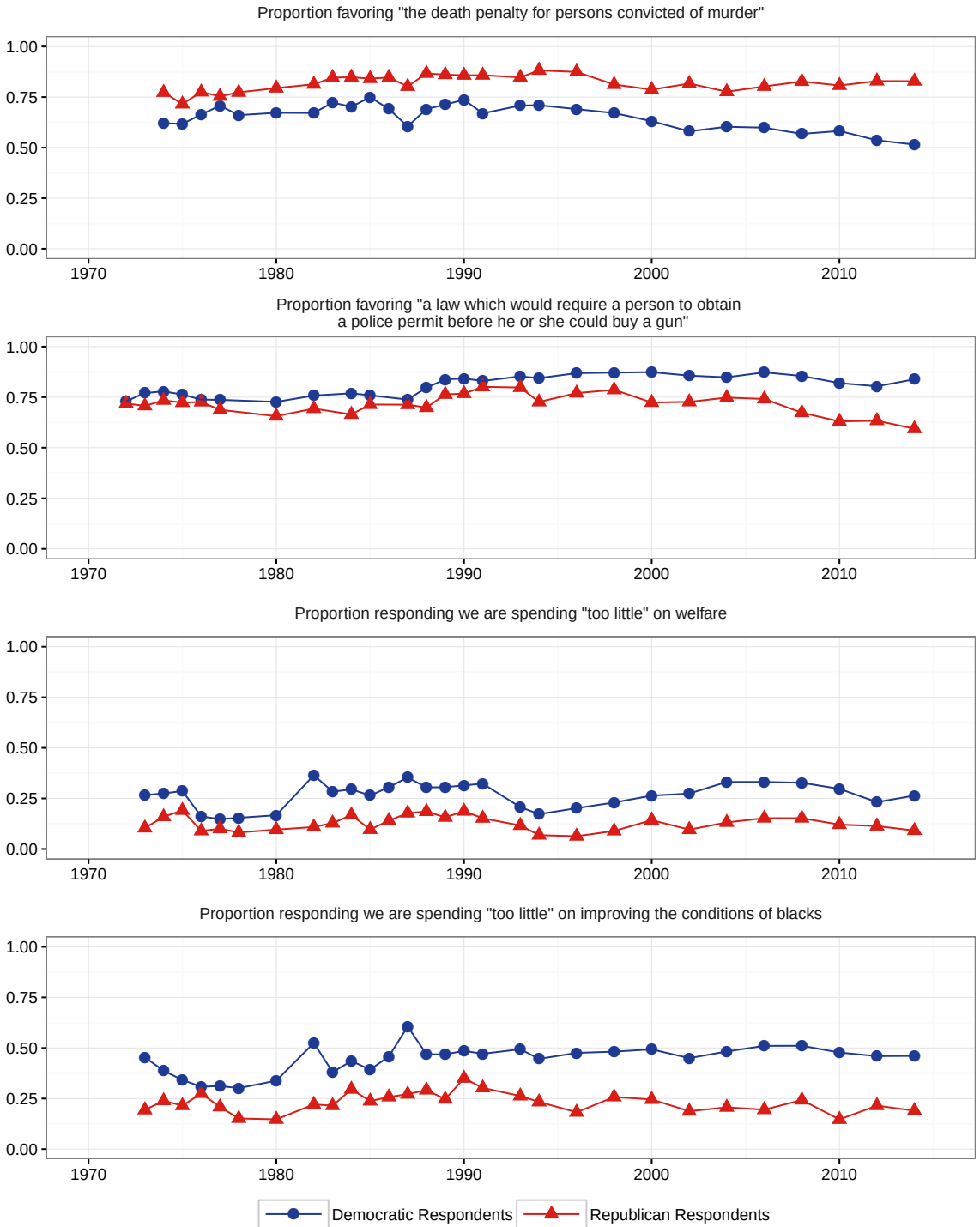
Among other sources, Page and Shapiro drew on the General Social Survey (GSS) from 1972 though 1983. Figure 1.2 updates some their findings using the GSS from 1972 to 2014. It plots the parallel movements of Republican and Democratic average opinion on four items: the death penalty, gun control, welfare spending, and spending to improve the condition of blacks. All four items have been largely stable for the past 40 years and the little movement we do see has largely been in parallel.

The top panel displays the proportion favoring the death penalty for those convicted of committing murder. Republican respondents have held steady over the past four decades at approximately 80% support. Democrats are consistently lower, hovering around the 60% mark, with a decline to approximately 50% over the past 15 years. While it is clear that the gap between partisans has widened on this issue, we do not see evidence of contrary motion. The second panel shows a similar story for the proportion favoring requiring police permits before gun purchases: from a baseline of approximately 75%, both Republicans and Democrats increased their support slightly in the 1990s. In recent years, Republicans' support has dropped off slightly, while Democratic support for the policy has remained constant. The third panel shows attitudes toward welfare spending. We see very little over-time movement, but small shifts of 2 to 5 percentage points appear to move in tandem. The final panel shows how, for many years now, 50% of Democrats respond that we are spending too little to improve the conditions of blacks, while the comparable figure among Republicans is 25%.

The parallel publics thesis extends to other, more obviously partisan opinions. Green et al. (2002) showed that while Democrats, Independents, and Republicans have vastly different baseline levels of presidential approval, corresponding quite plainly with the party of the current presidential office-holder, *changes* in presidential approval take place in parallel. Green et al. (2002) conducted their analysis for Harry Truman through Ronald Reagan using the presidential approval series from the Gallup organization. Figure 1.3 updates their analysis from Bill Clinton through Barack Obama.

These graphs are remarkable for their within-president consistency. Support for Bill Clinton rose for all three groups over the course of his presidency; support declined over the course of George W. Bush's term, and has stayed essentially stable for Barack Obama's time in office. Some of the changes have obvious causes: the spikes in approval after 9/11 and the invasion of Iraq; the gentle rise in Obama's approval among Democrats and Independents during the 2012 campaign. Other changes, such as the long decline in Bush's approval likely have many causes. These graphs also conform to the "weak form" of the parallel publics thesis. Clearly, these partisan groupings change their approval of the president in the same

Figure 1.2: Issue Positions by Partisanship



direction as one another, but not obviously by the same amount. For example, one might conclude that, following 9/11, Republicans' approval of the President did not move much, at least not as much as that of Independents or Democrats. A critic of this reading might point out that Republicans were constrained by the scale, and that a change from 88% support to near universal support (98%) is at least as impressive as Democrats' jump from 30 to 81.

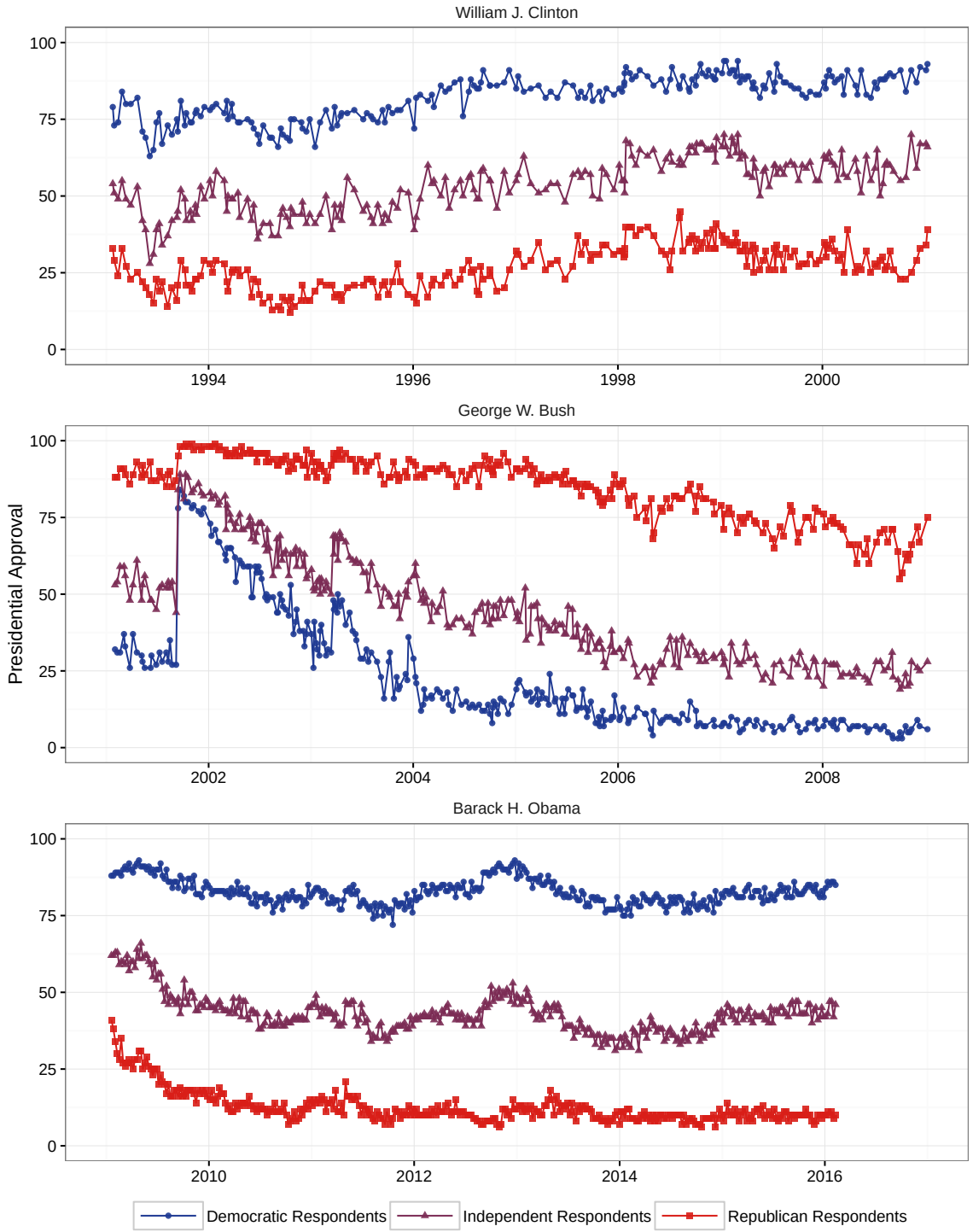
1.2 Experimental Evidence

The remainder of this chapter will explore the micro foundations of parallel publics, specifically the hypothesis that people of varied backgrounds and political views all update their views in the same direction when exposed to new information. As discussed above, observational survey evidence cannot establish this relationship conclusively, for two reasons. First, researchers are not in control of who sees what kind of information, so there is no guarantee that different subgroups see the same information. Second, because the repeated cross-sectional design relies on the passage of time to reveal changes in opinions, we cannot be sure if it is information itself, rather than some other factor that information happens to be correlated with, that is changing minds.

The survey experimental research design will help on both counts. The researcher is in control of the information seen by subjects; there is no worry that some subgroups will be exposed to different information. Further, because information treatments are allocated at random before outcomes are collected, direct causal inferences can be drawn. If the distribution of outcomes differs across the randomly-formed groups, we can attribute those differences to the information treatment itself.

To test the claim that subgroups tend to respond to information in parallel, I will draw on a set of 20 survey experiments I conducted between May 2014 and August 2015. Fifteen of the experiments are replications of other scholars' designs, 12 of which were conducted on Amazon's Mechanical Turk and three of which were conducted on nationally-representative samples obtained by the survey firm GfK. With collaborators, I designed the remaining five experiments and conducted them with subjects on Mechanical Turk.

Figure 1.3: Presidential Approval by Partisanship



These experiments cover a wide range of topics: capital punishment, the minimum wage, global warming, gun control, tax preferences, the Patriot Act, home foreclosures, immigration, free trade, crime, health care, economic policy, terrorism, public financing of elections, mental illness, infrastructure, and veterans' assistance. Taken together, these topics cover the majority of the contentious political issues discussed at a national level in the last decade. Notable exceptions include same-sex marriage and abortion, though I have no theoretical reason to expect different results for those issues.

The treatments in these experiments are, in one way or another, persuasion treatments, covering a wide array of persuasion strategies: descriptions of scientific evidence in graphs and text, video clips, newspaper stories, small pieces of factual information, op-eds, excerpts from speeches, or small changes in question wording.

Full descriptions of all 20 experiments are given in Appendix A, including the number of subjects, the precise treatments, outcome items (along with any transformations), and summaries of basic results. Except where noted, I use Ordinary Least Squares without any controls for pre-treatment covariates to obtain average treatment effects relative to a control, baseline, or placebo condition, depending on the experiment. In three experiments (Studies 1, 2, and 3), a baseline survey was conducted in advance of the experiment, so in those cases I include a pre-treatment measure of the outcome as a covariate.

1.2.1 Results 1: Parallel Treatment Effects by Subgroup

I will summarize the relationship between treatment effects across subgroups with Pearson correlation coefficients. Because treatment effects are measured with error, the raw correlations will be attenuated by measurement error. I will disattenuate the correlations using the standard formula: $\rho_{x'y'} = \frac{\rho_{xy}}{\sqrt{\rho_{xx}\rho_{yy}}}$, where ρ_{xx} and ρ_{yy} refer to the reliabilities of x and y. I will estimate these reliabilities using the same simulation method employed in Coppock and Green (2015): obtain the correlation between two draws from normal distributions implied by the coefficient vector and corresponding vector of standard errors and square it, repeat this operation 10,000 times, then take the average reliability. Sometimes this method returns a disattenuated correlation coefficient above 1, which is obviously incorrect. In such

cases, it is sufficient to infer that the true correlation is very high.

Figure 1.4 presents striking evidence of parallel updating in persuasion survey experiments according to three different demographic splits: partisanship (Republican versus Democrat), gender (men versus women), and education (college or higher versus some college or lower). For each subgroup pair in the figure, the left panel plots the baseline levels of the outcome variables and the right panel plots the treatment effects. All outcomes are standardized, so a baseline level of 0 indicates the sample mean in the control group.

The top row of Figure 1.4 presents estimates by partisanship, with Republicans on the vertical axis and Democrats on the horizontal axis. The baseline levels of opinion are very negatively correlated ($p < 0.001$). The raw correlation is -0.72; when correcting for measurement error the correlation is -0.77. This is unsurprising: adherents of the major parties, on average, hold opposing views on most political issues.[3] Turning now to the right panel of the top row, we see that Democratic and Republican responses to treatment are strongly and *positively* correlated (raw: 0.86, corrected: > 1.00). This positive correlation is strong evidence that partisans update their views in parallel. The finding that Republicans and Democrats update their views in the same direction by very similar amounts directly contradicts the "perceptual screen" theory as well as the theory that these groups would be differentially receptive to information, depending on whether subjects agree with its direction.

The second row of Figure 1.4 shows the estimates by gender. The attitudes of men and women on these issues are weakly correlated (raw: 0.31, corrected: 0.42) at baseline. Nevertheless, the effects of persuasive treatments are very similar for both groups. The correlation is estimated to be very high (raw: 0.90, corrected: > 1.00), stronger than the already very strong correlation observed for Republicans and Democrats. This analysis shows that the pattern of parallel updating does not depend on any particular correlation structure in baseline attitudes: whether baseline attitudes are positively, negatively, or not

[3]Their views are opposing in the sense that when Democrats hold above average views, Republicans tend to hold below average views; it does often happen that Republicans and Democrats "agree" in the sense that a majority of both groups supports an issue, but one group supports it more strongly.

at all correlated, treatment effects are very similar.

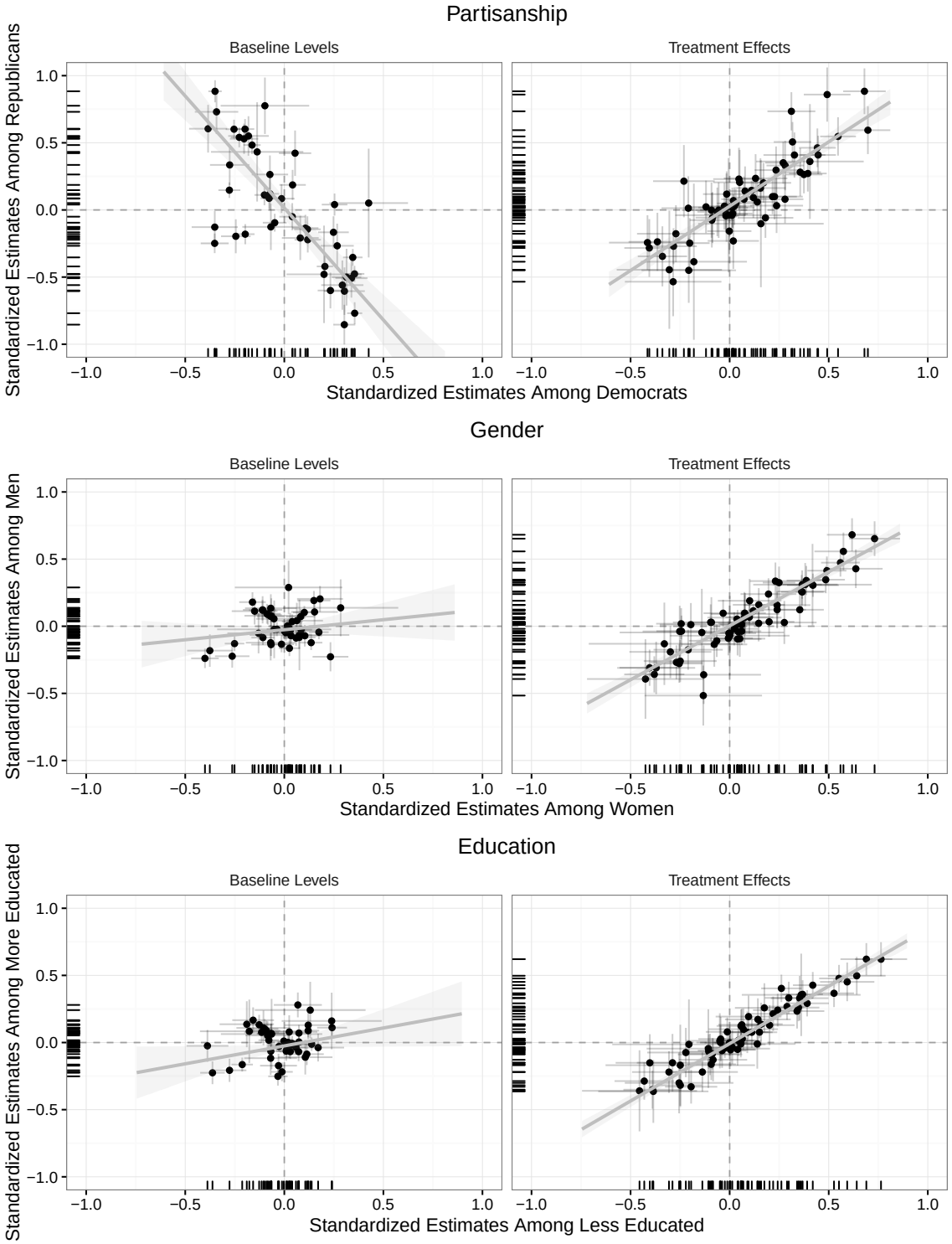
Finally, the third row of Figure 1.4 shows the effects for education. Those with a college education or higher are plotted on the vertical axis and those with some college or less are on the horizontal axis. If we assume that education is a reasonable proxy for political attentiveness (Price and Zaller 1993), then this analysis tests Zaller's notion that the more politically attentive will experience higher treatment effects. While baseline attitudes by eduction are only weakly correlated (raw: 0.29, corrected: 0.41), we do not observe any differences in the ways individuals respond to information by educational attainment. In fact, the correlation across this demographic split is higher than it was for either party or gender. Raw, the correlation is 0.95 and exceeds 1.00 when disattenuated.

In figure 1.1, I presented the implications of four different theories of how members of two groups would respond to information. The second column of panels in Figure 1.4 are very similar to the "Parallel Publics" panel. I find no support for the three alternative theories in this set of experiments.

1.2.2 Results 2: Parallel Treatment Effects with BART

The previous section presented the results of an analysis of many experiments. In this section, I will turn to an analysis of a single experiment where treatment effect heterogeneity should, in principle, be possible. Chong and Druckman (2010) conducted an experiment that tested the effects of pro and con information on attitudes toward the Patriot Act; I replicated their experiment on Mechanical Turk. Subjects assigned to the "Strong Pro" condition saw six pieces of information that emphasized the national security benefits of the law. Those in the "Strong Con" condition saw six pieces of information that emphasized the threats to civil liberties. Considering the highly polarized nature of the debate over the Patriot Act, we might expect that, in contrast to the homogeneous effects predicted by the parallel publics theory, the effects of treatment would be heterogeneous. We saw in the previous analysis that effects did not differ by party, gender or education. The concern is that the subgroup averages masked the true differences in responses that may have occurred along some other dimension.

Figure 1.4: Baseline Attitudes and Treatment Effects in 20 Survey Experiments



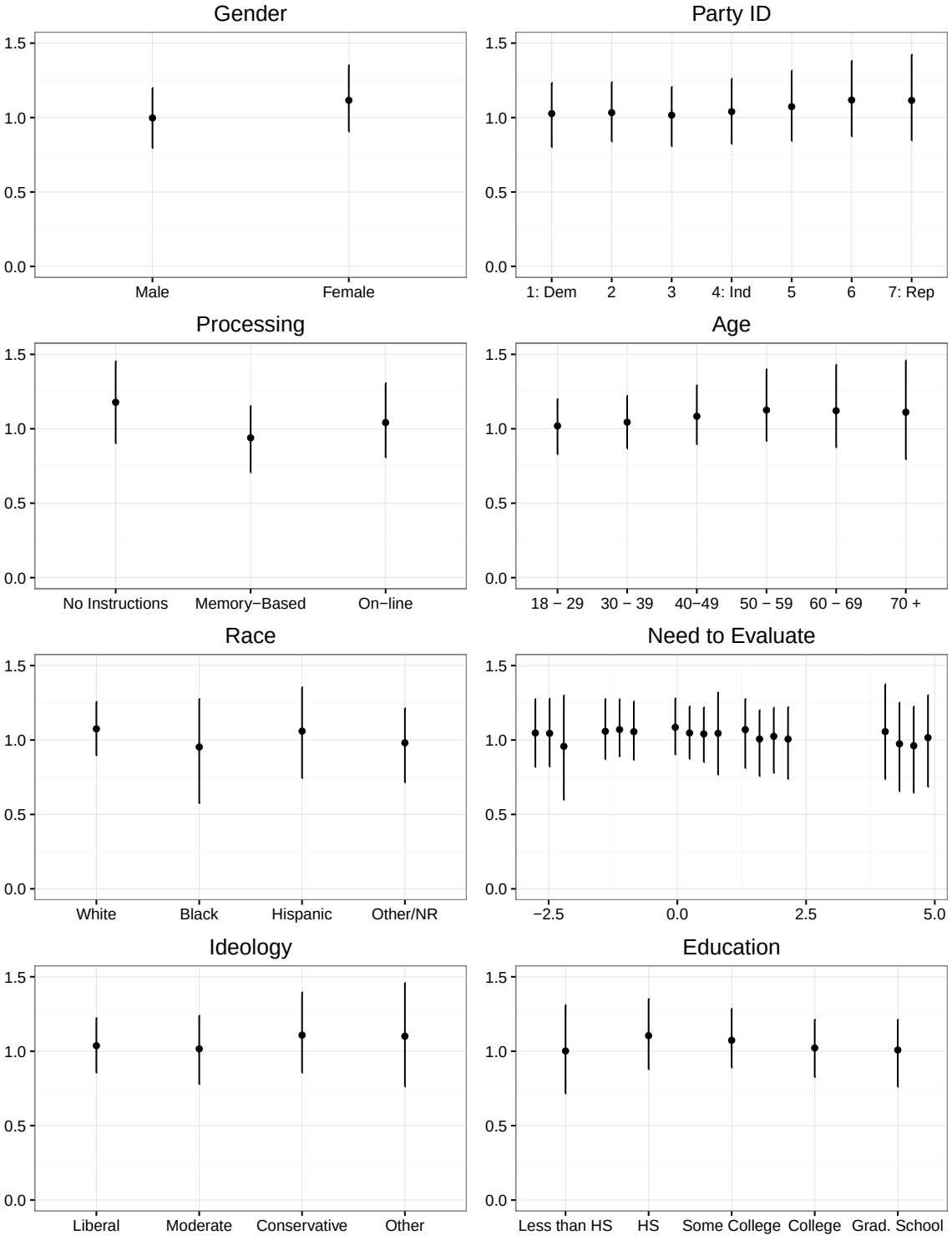
In my replication, the average difference in outcomes according to the type of information seen was large: 1.07 scale points on a seven point scale of support for the Patriot Act. (Robust SE: 0.10). The goal of the remainder of this analysis is to test whether any subgroups defined by age, gender, race, education, partisanship, ideology, "need-to-evaluate," the randomly-assigned "processing condition" (See Appendix A, Study 7 for a description of this feature of the original experiment), or their interactions, experienced much larger or smaller treatment effects. These covariates were chosen because they cover the vast majority of the demographics that are of concern to social scientists, as evidenced by the large overlap with the demographics listed in Table 1.1.

The multidimensional space described by these covariates and their intersections is vast. One unreliable approach to searching for heterogeneous effects is to engage in "fishing" (Humphreys, Sanchez de la Sierra and van der Windt 2013) by iteratively including interaction terms until a "statistically significant" interaction effect appears. This data-dredging procedure can lead to incorrect inferences because the nominal p-values do not reflect the multiplicity of tests conducted. One approach to this problem is to specify ex ante which treatment-by-covariate interactions are of theoretical interest in a pre-analysis plan. A second approach is to model all treatment effect interactions at once, but defend against false discovery by slightly shrinking all interaction effects toward zero. This second approach is handled nicely by a relatively new statistical method, Bayesian Additive Regression Trees (BART, Chipman, George and McCulloch (2007, 2010)). For further explanation of the BART algorithm and its uses for causal inference in the social sciences see Hill (2011) and Green and Kern (2012), as well as Chapter 2 of this dissertation.

For now, it is sufficient to understand that BART, just like any other statistical model, makes a prediction for both the potential outcomes for each unit. The difference between these predictions is the model's estimate of the causal effect of being in the "Pro" condition versus the "Con" for that particular unit. Figure 1.5 plots the averages of the BART-estimated individual-level treatment effects according to the demographic divisions described by each covariate. The estimates are remarkably stable across all demographic groupings: the conditional average treatment effects are all very close to the overall aver-

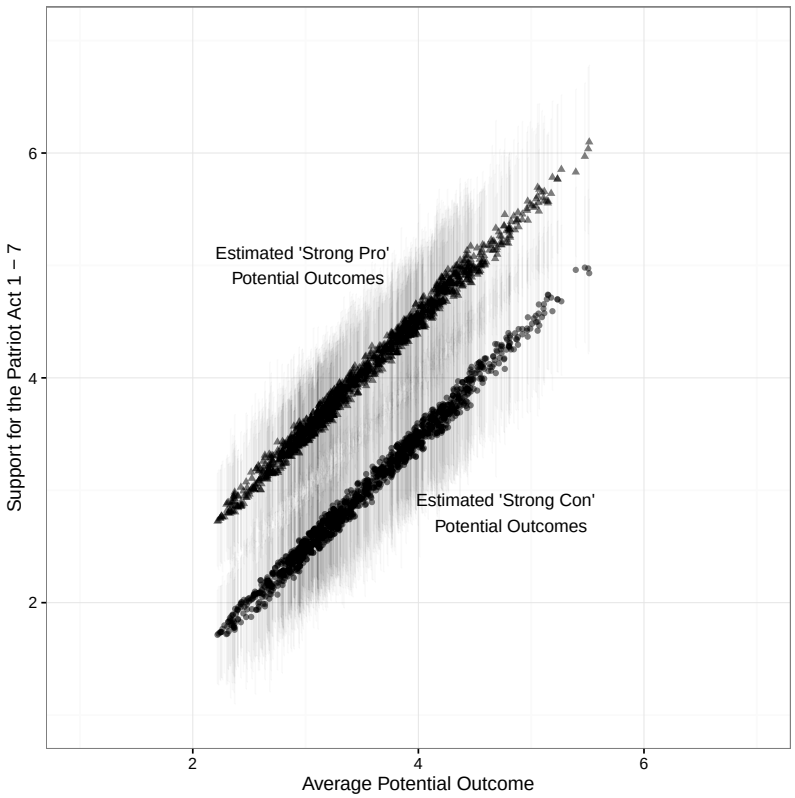
age of 1.07 scale points. This graph shows that there do not appear to be any significant interactions according to these demographic variables.

Figure 1.5: BART Estimated Treatment Effects in Study 7: Patriot Act



Another advantage of BART is that we are not limited to exploring heterogeneity according to the covariates we provide directly: BART automatically searches for treatment effect interactions that occur according to the intersection of covariates. If, for example, treatment effects were especially low for young Democratic women, the BART-estimated response surface would flexibly accommodate this four-way interaction. Figure 1.6 provides evidence that no such interactions are likely. The BART-estimated potential outcomes for each unit are plotted on the vertical axis. The triangular points refer to the estimated potential outcome under the "Strong Pro" treatment while the circular points refer to each unit's estimated potential outcome under the "Strong Con" treatment; the difference between them is the estimated treatment effect. The points are arrayed on the horizontal axis according to the average of the two estimated potential outcomes in order to better emphasize the decidedly parallel estimated effects.

Figure 1.6: BART Estimated Potential Outcomes in Study 7: Patriot Act



The pattern in Figure 1.6 appears highly linear, but nothing in the BART model requires linearity; both the pro and con response surfaces were flexibly modeled. These results indicate that any differences in treatment response according to the measured covariates or their interactions are small to negligible. One caveat to this analysis is that there may still remain unmeasured dimensions along which subjects and their treatment responses differ. If this unmeasured dimension is uncorrelated with the measured covariates, then BART would not recover evidence of treatment effect heterogeneity. While I am certain that many unmeasured covariates affect baseline opinions regarding the Patriot Act, I find it implausible that these dimensions could be so wholly uncorrelated with the measured covariates explored here that substantial interactions exist, but are not even slightly evident in the figures above.

The BART analysis is strongly in line with both the cross-experiment evidence and the evidence from time series cross-sectional surveys: opinion shifts occur in parallel, even for a highly contentious issue such as the Patriot Act.

1.3 Discussion

The purpose of this chapter was to demonstrate that homogeneous change in opinions in response to information remains a viable explanation for the parallel publics finding. Despite low or even strongly negative correlations in baseline outcomes, treatment effects in 20 survey experiments were shown to be very similar across politically-important demographic splits: Republicans and Democrats, men and women, and the better and less-well educated. In one experiment, I used a recently developed statistical method (BART) to show evidence of treatment effect homogeneity across the whole of the observed covariate space. Exposure to information does change minds, and it changes minds by similar amounts for the subdivisions of the population that social scientists tend to measure.

These results provide strong evidence in favor of Page and Shapiro's conjecture that the explanation for parallel updating is due to mass exposure to political information via the media. When people are exposed to information, they change their views in parallel, regardless of background characteristics. In the next two chapters, I develop a theory of

Bayesian information processing that undergirds this pattern of uniform updating.

2 Back to Bayes

A large body of theory suggests that when individuals are exposed to counter-attitudinal evidence or arguments, their preexisting opinions and beliefs are reinforced, resulting in a phenomenon known as attitude polarization. Drawing on evidence from three well-powered randomized experiments designed to induce polarization, I find no such effect. To explain these results, I develop and extend a theory of Bayesian learning that, under reasonable assumptions, rules out attitude polarization. People, instead, update their prior beliefs in accordance with the evidence presented in a more or less uniform fashion. I further show, through an exact replication of the original Lord et al. (1979) study, that the seeming discrepancy between these results and those in the literature are driven by differences in experimental design. Using this insight, I reinterpret previous findings and argue that Bayesian reasoning remains the best available model of opinion change in light of evidence.

This chapter draws on joint work with Andrew Guess.

For several decades, the attitude polarization literature has presented a challenge for believers in evidence-based decision-making. Instead of prompting a reconsideration of long-held views, counter-attitudinal evidence may actually strengthen preexisting beliefs, resulting in polarization. How can agreement, even when the evidence is relatively clear, ever be achieved?

A tendency to follow "directional" goals in evaluating information (Kruglanski and Webster 1996) would seem to fly in the face of normative standards of cognition which posit a rational weighing of evidence against previous knowledge. The most straightforward model of this type invokes Bayes' rule, which provides a simple, mathematically coherent mechanism for updating one's prior beliefs in light of new evidence. The predictions of this model are subtle, leading to occasional disagreements about the expected pattern of evidence under various conditions (Gerber and Green 1999; Bartels 2002; Bullock 2009). Most clearly, however, it rules out polarization, defined here as a circumstance in which information has a positive effect for some but a negative effect for others. Bayesian reasoning posits that if two people receive the same information and have compatible beliefs about the likelihood of observing it, then they must update their views in the same direction, although not necessarily by the same amount. This is the case even if their prior beliefs are completely at odds.

Contrary to this model, some authors theorize that people work – consciously or not – to protect their worldviews via a series of complementary belief-preserving mechanisms (Kunda 1990). Examples include the *prior attitude effect*, the tendency to perceive evidence and arguments that support one's views as stronger and more persuasive than those that challenge them; *disconfirmation bias*, in which people exert effort to counter-argue vigorously against evidence that is not congruent with their beliefs; and various forms of selective exposure and selective attention to congenial information, sometimes referred to as *confirmation bias* (Taber and Lodge 2006). The cumulative effect of these "biased assimilation" mechanisms, as predicted in the literature, is attitude polarization: people exposed to the same information may respond by strengthening their preexisting views.

The canonical explication and demonstration of these mechanisms is Lord et al. (1979),

in which both pro- and anti-death penalty college students were exposed to mixed scientific evidence on the effectiveness of capital punishment on crime deterrence. To their surprise, the authors found that the subjects did not moderate their views; rather, those who initially supported the punishment became more pro-capital punishment on average by the end of the study, and those who opposed it became less so. Moreover, biased assimilation appeared to be the culprit: the participants were more skeptical of the parts of the evidence that challenged their beliefs than those that supported them.

Since then, a substantial literature in psychology (Miller, McHoskey, Bane and Dowd 1993; Kuhn and Lao 1996), political science (Redlawsk 2002; Taber and Lodge 2006; Lau and Redlawsk 2006; Taber et al. 2009; Nyhan and Reifler 2010), and other fields such as public health (Strickland, Taber and Lodge 2011; Nyhan, Reifler, Richey and Freed 2014) have sought to replicate, extend, and further delineate the mechanisms for this provocative hypothesis. A well-known early criticism of the Lord et al. results focused on the authors' use of a self-reported measure of attitude change rather than a more direct measurement of the difference between pre- and post-treatment attitudes (Miller et al. 1993). In my view, the more fundamental methodological issue is that the treatments in Lord et al. (1979) were not allocated on a random basis. Without randomizing the content of the information itself, it is difficult to adjudicate between the competing predictions of attitude polarization and Bayesian updating. The concern is that the "effect" of mixed information is actually a reflection of pre-existing differences among subjects.

In this chapter, I will provide evidence that calls into question the consensus on attitude polarization. Across three studies designed to induce differential treatment responses and covering different issue areas and types of stimuli, subjects update their beliefs roughly in parallel, regardless of predisposition. Given the large, diverse subject pools and broad set of issues, it is unlikely that these findings are due to the particularities of the experimental context. More fundamentally, these results indicate that several well-known results taken as evidence of attitude polarization are likely to be artifacts of inadequate research design.

These conclusions do not call into question the ample evidence for "biased assimilation" processes. For example, the prior attitude effect is perfectly consistent with Bayesian

reasoning. The relevant insight is that prior beliefs can change in response to evidence even as they color interpretations of that evidence. As Gerber and Green (1999, p. 199) argue in their review of studies on perceptual bias, "although it is not inconsistent with Bayesian learning for observers to comment that contrary information is unconvincing, in the simple version of Bayesian learning, we do not expect critics of new information to be altogether unmoved by it." This paper proceeds as follows. First, I present an overview of the Bayesian learning hypothesis and its predictions under different assumptions about the likelihood of observing evidence. Next, I outline my overall research approach for three separate studies and clarify how it can adjudicate between competing models of attitude change. Employing standard linear models, I present the experimental results from each study in turn, including evidence (where available) that the effects persist over one week later. I further explore the results in Section 7 with a relatively new method in political science, Bayesian Additive Regression Trees (BART), that enables omnibus modeling of treatment effect heterogeneity. The following section examines several well-known results in the attitude polarization literature and re-interprets them in light of these experiments. I conclude with a discussion of the implications of these findings.

2.1 Bayesian Learning

The Bayesian learning hypothesis may be traced at least as far back as de Finetti (1937), who appears to have been the first to explicitly invoke Bayes' rule as a model of human cognition.[1] The main thrust of that work is that although individuals make subjective probability judgments, they nevertheless respond to new information in the manner suggested by "objective" laws of probability:

> Observation cannot confirm or refute an opinion, which is and cannot be other than an opinion and thus neither true nor false; observation can only give us information which is capable of *influencing* our opinion. The meaning of this statement is very precise: it means that to the probability of a fact conditioned

[1] See, however, Ramsey (1931) for a formulation that comes close.

> on this information – a probability very distinct from that of the same fact not conditioned on anything else – we can indeed attribute a different value. (Reproduced in Kyburg and Smokler 1964, p. 154, emphasis in original.)

The hypothesis developed alongside the introduction of Bayesian statistical methods to the social sciences (Edwards, Lindman and Savage 1963). Bayesian statisticians were arguing that Bayes' rule offered a sensible method for the integration of research findings, leading to the normative claim that scientists and humans generally *should* incorporate new evidence according to the cool calculus of conditional probability. It was then a short step from advocating this normative position to testing the positive claim that Bayes' rule provides a fair description of human cognition.

A long series of demonstrations (Edwards and Phillips 1964; Peterson and Miller 1965; Phillips and Edwards 1966) showed that humans perform poorly compared to (a specific version of) the Bayesian ideal. In the prototypical experiment, subjects are shown two urns, one with majority blue balls and the other with majority red balls. They are told that they will be given a series of draws from a randomly selected urn and will have to estimate the probability that the draws come from the majority blue urn. Bayes' rule (with a binomial likelihood function) dictates how much subjects "should" update their priors; subjects are consistently too conservative and fail to update their beliefs far enough.

Within political science, the Bayesian models of cognition have been most frequently applied to models of partisan evaluation and choice. Zechman (1979) and Achen (1992) propose statistical models of party identification based on a Bayesian model in which individuals update their impressions of the political parties based on new evidence. Gerber and Green (1999) argue that so-called "perceptual bias" can be explained in Bayesian terms: partisans accord negative evidence for their candidates less weight not because they irrationally disregard discordant information, but because it conflicts with their priors. Partisans are nevertheless moved by "bad news," as evidenced by the parallel public opinion movements by partisans on both sides in response to changing political and economic conditions (see Chapter 1 of this dissertation). Bartels (2002) disagrees, arguing that a lack of convergence (i.e., parallel motion) indicates partisan bias. Bullock (2009) shows that the prediction of

convergence requires an assumption that the underlying parameter (in this case, the "true" value of the parties) is not changing.

To this critique of Bartels (2002), I would add that the convergence prediction requires that information and opinions be in the same scale. Suppose that the opinion item in question is support for a Democratic candidate; Democratic partisans support the candidate at 85%, Republicans at 25%. Gerber and Green note than in response to information, support for the candidate fluctuates in parallel. Bartels views this parallel fluctuation as evidence of partisan bias, because in his view, Bayesian Democrats and Republicans should converge over time. This is only true if "evidence" could indicate that the proper level of support is somewhere between 85% and 25%, in which case Democrats should flag in their support, whereas Republicans should become more supportive. Information, however, is usually not measured in the same scale as opinions: it comes as good or bad news for a candidate, not a target level of support.

Any Bayesian framework involves five quantities: a prior, new evidence, a likelihood function, a posterior, and an objective function. Bayes' rule itself is an accounting identity that relates the first four of these. In the discrete case, with two possible states of the world A and $\neg A$, the formulation is: $P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$. The prior is $P(A)$, the new evidence is B, the likelihood function is hidden in $P(B|A)$ and $P(B)$, and the posterior is $P(A|B)$. The objective function is not part of Bayes' rule – it describes what an individual decides or does as a result of the update. In the experiments presented below, the objective function describes how a subject responds to a survey item.

As noted by many scholars (e.g., Taber et al. 2009, p. 138), Bayes' rule on its own does not provide a prescription for how a "rational" person should update. As a result, "violations" of Bayesian rationality are difficult to demonstrate because we often lack the requisite information. For example, a typical persuasion experiment will measure prior attitudes $P(A)$, intervene with some evidence B, and then measure posterior attitudes $P(A|B)$. In order to determine whether the update $P(A|B) - P(A)$ moved in the "correct" direction – toward the evidence – we need either to measure or impose assumptions on the remaining terms in Bayes' rule: $P(B|A)$ and $P(B)$. In other words, we need a likelihood

function describing how individuals read or interpret the new evidence B. This function would describe whether, for example, we expect to observe domestic discontent more often when a leader is weak or when a leader is strong.

The Bayesian framework can flexibly accommodate individual differences on two dimensions: the prior and the likelihood function. Within political science, we often categorize individuals according to their imputed priors: Democrats presumably have a stronger prior belief that liberal policies will be effective than do Republicans. How might likelihood functions differ across individuals? For the simplest case in which A and B are both binary, a likelihood function is fully described by two numbers: $P(B|A)$ and $P(B|\neg A)$. Suppose that we believe that B is evidence in favor of the proposition that A is true: any likelihood function in which $P(B|A) \geq P(B|\neg A)$ would generate updates in the "correct" direction, i.e. $P(A|B) - P(A) > 0$. This restriction on the likelihood function follows the monotone likelihood ratio property (MLRP, Dixit and Weibull 2007, Karlin and Rubin 1956). If this restriction does not hold, individuals are not necessarily less Bayesian; they may just interpret evidence differently from others.[2]

The treatments used in the experiments described below operate by adding new considerations (and possibly changing the weights attached to pre-existing considerations). Some of the outcome questions measure "beliefs," by which I mean the content of some consideration. Others measure "attitudes," which are the weighted average of considerations. The distinction between attitudes and beliefs is important to maintain. I do, however, recognize that the "belief" survey questions may also (inadvertently) measure what should rightly be considered an attitude. In these experiments, subjects are given evidence in favor of some proposition (e.g., capital punishment decreases crime) and are then asked to report relevant political attitudes (e.g., support for capital punishment), not the probability that the proposition is true. I assume a monotonic relationship between the probability that the

[2]Benoît and Dubra (2014) advance a theory of *rational attitude polarization* in which fully Bayesian individuals can encounter the same evidence yet update in opposite directions. Their model accomplishes this by allowing individuals to have access to different "auxiliary information" which essentially colors how the new evidence is interpreted. Put simply, individuals in their models have different likelihood functions, some of which follow the MLRP and some of which do not.

proposition is true and survey measures of political attitudes (e.g., the higher the posterior probability that capital punishment decreases crime, the higher the support for the policy). This is reasonable if the treatments add or adjust the weights attached to individual considerations comprising an attitude.

A Bayesian theory of information processing makes specific predictions about how much different individuals should update their views in light of evidence. Those with more strongly held prior beliefs should update less than those with weak priors. Those with more discriminating likelihood functions should update more than whose likelihood functions are less finely tuned. Unfortunately, we lack the tools to measure the strength of individuals' priors and the quality of their likelihood functions (Kahan 2015). In view of these difficulties, I will make directional predictions, using the portions of the Bayesian framework that do not rely on such unmeasured subject attributes.

What, then, constitutes evidence against this "directional Bayes" theory in favor of the theory of attitude polarization? If we accept that subjects' likelihood functions all follow the MLRP (or, in non-Bayesian terms, subjects all agree that the evidence is "good for" the same proposition, even if they counter-argue, thereby making other evidence against the proposition more salient) and we grant the mild restriction that subjects' objective functions are monotonic, then we would have strong evidence against Bayesian reasoning if some subjects updated in one direction but others updated in a different direction. If all subjects update in the same direction, we have strong evidence against a key prediction of the attitude polarization hypothesis: namely, positive effects for some, but negative effects for others.

2.2 Research Approach

The three experiments presented in this chapter share some salient design features, described together here for convenience.

2.2.1 Design and Measurement

All three studies employ a between-subjects experimental design. Studies 1 and 2 were conducted in three waves. First, respondents were invited to complete a survey that collects pre-treatment covariates. A subset of those who completed the responses was then selected for the second wave, which contained the experimental treatments. The main outcomes of interest were recorded in this second wave (T2). Finally, to measure the persistence of effects, subjects were recontacted approximately 10 days later to collect a new set of outcome measures (T3). Study 3 did not feature a third wave of measurement. To address the above-mentioned measurement critiques of Lord et al. (1979), I use direct measures of post-treatment attitudes and beliefs rather than subjective self-reports (although both were collected in Study 1 in order to replicate earlier findings, as discussed below).

2.2.2 Subject Recruitment

Instead of engaging laboratory subjects, I conducted these studies on Amazon's Mechanical Turk (MTurk), an online marketplace for modular "human intelligence tasks." In recent years, social scientists have recognized its utility as a tool for recruiting subjects (e.g., Buhrmester, Kwang and Gosling 2011). While opt-in samples collected via MTurk should not be considered representative of the U.S. population, they have been shown to outperform laboratory-based convenience samples in ensuring adequate variation across sociodemographic and political characteristics of interest (Berinsky, Huber and Lenz 2012*a*). Typically, MTurk samples tend to skew younger, less wealthy, better educated, more male, whiter, and more liberal than the population as whole. I opted for this approach in order to boost the representativeness relative to student populations (Sears 1986; Clifford and Jerit 2014), and to better ensure generalizability of results via similarity of subjects, treatments, contexts, and outcome measures across domains (Coppock and Green 2015). The utility of this sample for drawing inferences about the causal effects of information treatments depends on treatment effect heterogeneity. If the treatment effects for these subjects are substantially different from the effects for others, then MTurk is a poor guide to effects more generally. The evidence to date on this question suggests that estimates obtained

from MTurk samples match national samples well (Mullinix, Leeper, Druckman and Freese (2015), Chapter 4 of this dissertation).

In Appendix B, I present an exact replication, using the original study materials, of the Lord et al. (1979) study. The results closely match the direction and significance of the original, and in the case of biased assimilation the point estimates are virtually identical. It is therefore unlikely that the overall findings are driven by any qualitative differences between the Mechanical Turk subject pool and those drawn from undergraduate populations.

2.3 Study 1: Capital Punishment

Study 1 embeds a replication of the original (Lord et al. 1979) demonstration within a larger multi-arm experiment. In that study, subjects were presented sequentially with apparent scientific evidence both challenging and affirming the notion that the death penalty deters crime. Subjects' attitudes about capital punishment and beliefs about its deterrent efficacy were subsequently measured. Subjects in the original study were exposed to a "mixed evidence" condition.[3] To this single condition, various combinations of pro-capital punishment, anti-capital punishment, and inconclusive evidence were added. I made minor updates to the text (changing the publication date of the fictitious research articles from 1977 to 2012, for example) and to the graphical and tabular display of the fabricated data using modern statistical software.

The results of the replication study are presented in full in Appendix B. To summarize: following the analytic strategy of the original authors, I recover nearly identical point estimates of the "effect" of mixed evidence on self-reported attitude changes. This similarity is remarkable considering the probable differences between the Mechanical Turk sample and late-1970s Stanford undergraduates. This estimate of the effect of mixed evidence is (almost surely) biased because it compares subjects according to their pre-existing differences, not

[3]Authors who performed an earlier replication of the original design (Kuhn and Lao 1996) generously shared the original experimental materials, so I was able to use the identical wordings of the study summaries and descriptions.

according to a random assignment.

2.3.1 Subjects

I initially recruited 1,659 subjects via Mechanical Turk. Subjects were paid $0.25 for their participation in this portion of the experiment. The pre-survey designed for screening subjects gathered a series of pre-treatment covariates (age, race, gender, and political ideology) and two items concerning capital punishment: attitude toward the death penalty and belief in its deterrent effect.

From the pool of 1,659, 933 subjects' pre-survey responses indicated clear and consistent support or opposition to capital punishment. These subjects were invited to participate in the main survey. Among these, proponents were defined as subjects whose answers to the pre-treatment attitude and belief questions were between 5 and 7 on a 7-point scale. Opponents were defined as subjects whose answers to these questions were between 1 and 3. In total, 683 subjects participated in the main survey (287 proponents and 396 opponents). Of those, 628 completed the 10-day follow-up, for a recontact rate of 92%. Subjects were paid $1.00 to complete the main survey and an additional $1.00 to complete the follow-up.

The two main dependent variables measured subjects' attitudes and beliefs about capital punishment. The *Attitude* question asked, "Which view of capital punishment best summarizes your own?" The response options ranged from 1 (I am very much against capital punishment) to 7 (I am very much in favor of capital punishment). The *Belief* question asked, "Does capital punishment reduce crime? Please select the view that best summarizes your own." Responses ranged from 1 (I am very certain that capital punishment does not reduce crime) to 7 (I am very certain that capital punishment reduces crime).

2.3.2 Procedure

Subjects were randomly assigned to one of six conditions,[4] as shown in Table 2.1 below. All subjects were presented with two research reports on the relationship between crime rates and capital punishment that varied in their findings: *Pro* reports presented findings that capital punishment appears to decrease crime rates, *Con* reports showed that it appears to increase crime rates, and *Null* reports showed that no conclusive pattern could be discerned from the data. The reports used one of two methodologies:[5] either cross-sectional (a comparison of 10 pairs of neighboring states, with and without capital punishment) or time-series (a comparison of the crime rates before and after the adoption of capital punishment in 14 states).

The statistical models will operationalize the "information content" of the pair of reports seen by subjects in a linear fashion. The information content of two *Pro* reports is coded as 2, one *Pro* and one *Null* as 1, and so on. In order to allow the coefficient on information content to vary depending on whether the information is pro- or counter-attitudinal, I will split the information content variable into positive information and negative information, as shown in Table 2.1.

Table 2.1 also presents the mean and standard deviation of both outcome variables by treatment group. These means provide a first indication that the treatments had average effects in the correct direction.

We can also use the standard deviations reported in the table to perform a test of whether the treatments polarized opinion. If the treatments were polarizing, the standard deviation in the successively more pro or more con treatment groups should be larger, but we do not observe this pattern. Formal statistical tests also reveal that the treatment

[4]Following the original Lord et al. (1979) procedure, in treatment conditions 1, 3, and 6, the order of the reports' methodology (time series or cross-sectional) was randomized, resulting in two orderings per condition. In treatment conditions 2, 4, and 5, both the order of the methodology and the order of the content were randomized, resulting in four orderings per condition. In total, subjects could be randomized into 18 possible presentations. This design was maintained in order to preserve comparability with the original study, but I average over the order and methodology margins to focus on the effects of information.

[5]The experimental materials are available in Appendix D.

groups do not differ with respect to the standard deviations of the outcomes. Relative to the Null Null condition, the differences-in-standard-deviations across groups are generally not statistically significant at the 5% level, according to a randomization inference test conducted under the sharp null hypothesis of no effect for any unit. The only exception is the difference-in-standard-deviations between the Null Null and Pro Null conditions for the support dependent variable, which has a p-value of 0.04. This inference does not survive common multiple comparisons corrections, including the Bonferonni, Holm, and Benjamini-Hochberg corrections. I conclude from these tests that the treatments do not polarize opinion in the sense of increasing its variance.

Table 2.1: Study 1 (Capital Punishment) : Treatment Conditions

				T2 Attitude		T2 Belief	
Condition	N	Positive Information	Negative Information	Mean	SD	Mean	SD
Con Con	117	0	2	3.521	2.074	3.274	1.649
				(0.191)	(0.075)	(0.157)	(0.089)
Con Null	116	0	1	3.147	2.115	3.112	1.576
				(0.200)	(0.089)	(0.148)	(0.085)
Null Null	112	0	0	3.179	2.063	3.134	1.562
				(0.195)	(0.089)	(0.147)	(0.084)
Pro Con	118	0	0	3.593	2.265	3.686	1.748
				(0.210)	(0.072)	(0.163)	(0.089)
Pro Null	121	1	0	3.620	2.285	3.810	1.665
				(0.207)	(0.071)	(0.151)	(0.086)
Pro Pro	102	2	0	3.686	2.225	4.127	1.609
				(0.226)	(0.072)	(0.165)	(0.097)

Bootstrapped standard errors in parentheses.

Subjects were exposed to both of their randomly assigned research reports – one time series and one cross-sectional within each treatment condition – according to the following procedure:

1. Subjects were first presented with a "Study Summary" page in which the report's findings and methodology were briefly presented. Subjects then answered two questions about how their attitudes toward the death penalty and beliefs about its deterrent efficacy had changed as a result of reading the summary.
2. Subjects were then shown a series of three pages that provided further details on

the methodology, results, and criticisms of the report. The research findings were presented in both tabular and graphical form.

3. After reading the report details and criticism, subjects answered a series of five questions (including a short essay) that probed their evaluations of the study's quality and persuasiveness.

4. Subjects then answered the attitude and belief change questions a second time.

Subjects completed steps one through four for both the first and second research reports. After reading and responding to the first and second reports, subjects were asked two endline *Attitude* and *Belief* questions, identical to the pre-treatment questions.

2.3.3 Analytic Strategy

The relatively complicated design described above can support many alternative analytic strategies. Each report is associated with seven intermediate dependent variables in addition to the two endline dependent variables. Subjects could have been assigned to eighteen different combinations of research reports. Reducing this complexity requires averaging over some conditions and making choices about which dependent variables to present. I present here my preferred analysis, though alternatives are offered in Appendix A.

I will focus on the separate effects of positive and negative information on subjects' T2 responses to the attitude and belief questions. Because the *Null Null* and *Pro Con* conditions are both scored 0 on both the positive information and negative information scales, I include an intercept shift for the *Null Null* condition, as shown in Equation 2.1:

$$Y_i = \beta_0 + \beta_1(POS_i) + \beta_2(NEG_i) + \beta_3(Condition_i = \text{Null Null}) + \epsilon_i \tag{2.1}$$

Under Bayesian reasoning, β_1 is hypothesized to be positive for all subjects and β_2 to be negative for all subjects. Under attitude polarization, β_1 is hypothesized to be positive for proponents, but negative for opponents, whereas β_2 is hypothesized to be *positive* for proponents and *negative* for opponents. That is, under attitude polarization, proponents always update in the positive direction regardless of the sign of the information; the opposite is predicted for opponents. β_3 represents the difference in average outcomes for subjects in the *Null Null* condition as compared with the *Pro Con* condition. Neither theory has a

strong prediction as to the sign of this coefficient; presumably it should be close to zero for all subjects. I will estimate Equation 2.1 via Ordinary Least Squares (OLS) among proponents and opponents separately, allowing for an easy comparison of $\hat{\beta}_1$ and $\hat{\beta}_2$ across subject types. In addition, I will estimate Equation 2.1 with and without a vector of covariates $\mathbf{X}_i$. Included in this vector of covariates are subjects' T1 responses to the attitude and belief questions. Including initial attitudes and beliefs as right-hand-side regressors is preferred to the differencing procedure implied by the "direct" measurement technique for attitude change, as it usually produces tighter estimates.[6]

2.3.4 Results

Tables 2.2 and 2.3 present estimates of the effects of information on attitudes and beliefs about capital punishment. Focusing on the covariate-adjusted models, I estimate that a unit increase in positive information causes an average increase in support for capital punishment of 0.015 scale points among proponents and 0.068 scale points among opponents, neither of which is statistically significant. Negative information has a strong negative effect among proponents (-0.326, $p < 0.01$), and a weakly negative effect among opponents (-0.042, $p = 0.43$). These estimates imply that moving from the *Con Con* condition to the *Pro Pro* condition would cause proponents to move $2 \cdot 0.015 + 2 \cdot 0.326 = 0.682$ scale points and opponents to move $2 \cdot 0.068 + 2 \cdot 0.042 = 0.220$ scale points. While the treatment effects do appear to differ by subject type ($p < 0.05$), we do not observe the "incorrectly" signed treatment effects predicted by attitude polarization.

Turning to Table 2.3, we observe that the effects of the information treatments on belief in the deterrent efficacy of capital punishment are nearly identical for proponents and opponents. For both groups, moving from *Con Con* to *Pro Pro* results in an entire scale point's worth of movement. This analysis shows clear evidence that subjects, when presented with information, update in a manner that is consistent with Bayesian reasoning, by making small but perceptible changes in the direction of the information. The attitude polarization hypothesis is roundly rejected by these data.

[6] See Gerber and Green (2012, pp. 96-102) for a fuller discussion of these two estimators.

Table 2.2: Effects of Information on Support for Capital Punishment

	Dependent Variable: T2 Attitude Toward Capitial Punishment			
	Among Proponents		Among Opponents	
Positive Information (0 to 2)	0.048	0.015	0.108	0.068
	(0.097)	(0.085)	(0.093)	(0.064)
Negative Information (0 to 2)	−0.267***	−0.326***	0.008	−0.042
	(0.102)	(0.082)	(0.078)	(0.054)
Condition: Null Null	−0.229	−0.194	−0.004	−0.068
	(0.196)	(0.157)	(0.138)	(0.104)
Constant	5.863	0.435	1.765	0.657
	(0.125)	(0.563)	(0.100)	(0.233)
Covariates	No	Yes	No	Yes
N	287	287	396	396
R^2	0.044	0.410	0.006	0.473

*p < .1; **p < .05; ***p < .01
Robust standard errors are in parentheses.
The information content of the Null Null condition is coded 0.
Covariates include T1 Attitude, T1 Belief, age, gender, ideology, race, and education.

Table 2.3: Effects of Information on Belief in Deterrent Efficacy

	Dependent Variable: T2 Belief in Deterrent Effect			
	Among Proponents		Among Opponents	
Positive Information (0 to 2)	0.146	0.118	0.348***	0.305***
	(0.106)	(0.101)	(0.105)	(0.091)
Negative Information (0 to 2)	−0.347***	−0.361***	−0.164	−0.207**
	(0.110)	(0.095)	(0.100)	(0.088)
Condition: Null Null	−0.322	−0.284	−0.275*	−0.302**
	(0.205)	(0.186)	(0.162)	(0.145)
Constant	5.078	1.695	2.472	1.441
	(0.143)	(0.679)	(0.115)	(0.329)
Covariates	No	Yes	No	Yes
N	287	287	396	396
R^2	0.084	0.255	0.090	0.341

*p < .1; **p < .05; ***p < .01
Robust standard errors are in parentheses.
The information content of the Null Null condition is coded 0.
Covariates include T1 Attitude, T1 Belief, age, gender, ideology, race, and education.

2.3.4.1 Persistence of Effects

As detailed in Appendix A and Chapter 4, I also measured the persistence of these effects 10 days after treatment. Put briefly, the effects on both dependent variables barely declined at all.

2.4 Study 2: Minimum Wage

One possible objection to the results from Study 1 is that the conditions for attitude polarization were not met. Specifically, the issue of capital punishment cuts across party lines and socioeconomic groups, so perhaps subjects are less inclined to "argue back" in the face of counter-attitudinal evidence. Study 2 addresses this concern by focusing on the minimum wage, a highly contentious issue area that divides the political parties and other demographic groups. The design of Study 2 is similar in spirit to Study 1, but instead of presenting treatments as figures with accompanying text, the treatments in this study were short web videos in favor and against raising the minimum wage.

2.4.1 Subjects

A large number (2,979) survey respondents on MTurk were recruited to participate in a pre-treatment survey measuring demographic characteristics (age, gender, race/ethnicity, education, partisan affiliation, and ideological leaning) as well as baseline attitudes toward the minimum wage and global warming (the focus of Study 3). From this large pool of survey respondents, 1,500 were invited to take part in the main survey testing the effect of videos on attitudes toward the minimum wage. Invitations to take part in the main survey were offered on a random basis, though more slots were offered to younger respondents and those with stronger views (pro or con) about the minimum wage. Of the 1,500 recruited to the main survey, 1,170 participated. Rates of payment were identical to Study 1: \$0.25 for the pre-survey and \$1.00 each for the main survey and the follow-up.

2.4.2 Procedure

Subjects were exposed to two videos on the subject of the minimum wage. Two of the videos were in favor of minimum wage increases, one presented by John Green, a popular video blogger, and the other presented by Robert Reich, former U.S. Secretary of Labor and established left-leaning public intellectual. On the "con" side of the debate, one video was presented by an actor, and the other by economics professor Antony Davies. Finally, two videos were included as placebos, addressing mundane requirements of state minimum wage laws. Links to all six videos are available in Appendix A, as well as screenshots and transcripts that convey the production quality and mood of the videos.

Subjects were randomized into one of thirteen conditions: placebo, or one of the twelve possible orderings of the four persuasive videos. Subjects answered intermediate questions relating to how well-made and persuasive they found each video, then at the end of the survey, they answered two questions which serve as our main dependent variables. The *Favor* question asked, "The federal minimum wage is currently $7.25 per hour. Do you favor or oppose raising the federal minimum wage?" The response options ranged from 1 (Very much opposed to raising the federal minimum wage) to 7 (Very much in favor of raising the federal minimum wage). The *Amount* question asked, "What do you think the federal minimum wage should be? Please enter an amount between $0.00 and $25.00 in the text box below."

2.4.3 Analysis

As in Study 1, I order the treatment conditions according to the amount of pro-minimum wage video content. The information content of the *Con Con* conditions is scored -1, the *Pro Con* and *Placebo* conditions 0, and the *Pro Pro* conditions 1, as shown in Table 2.4. I will model responses as in Equation 2.1 and estimate separate regressions for opponents, moderates, and proponents. Opponents are defined as subjects whose pre-treatment *Favor* response was 4 or lower and whose *Amount* response was below the median response ($10.00). Those with *Favor* responses of 4 or higher and *Amount* responses above the median are defined as proponents. All others are defined as moderates.

Table 2.4 presents the means and standard deviations by experimental group. Here again, we see an indication that the treatments had average effects in their intended directions. The means of the Con Young / Con Old condition are lower than the means of the mixed conditions, which are themselves lower than the means of the Pro Young / Pro Old conditions. These differences are all statistically significant. Turning to the differences-in-standard-deviations, formal tests under the sharp null of no effect lend some support to the hypothesis that the treatments lead to increases in the polarization of opinion – the differences between the placebo condition and the Pro Old, Con Old condition are statistically significant for both dependent variables at the $p < 0.01$ level. However, while increases in the standard deviations of outcomes would be a consequence of attitude polarization, these increases could also result from one group having larger treatment effects than another, but with both effects still positive.

Table 2.4: Study 2 (Minimum Wage): Treatment Conditions

				Favor		Amount	
Condition	N	Positive Information	Negative Information	Mean	SD	Mean	SD
Placebo	93	0	0	5.344	1.638	9.234	2.771
				(0.170)	(0.125)	(0.288)	(0.332)
Con Young / Con Old	162	0	1	4.765	1.837	8.671	2.651
				(0.144)	(0.081)	(0.208)	(0.254)
Pro Old / Con Old	165	0	0	4.988	2.089	9.993	4.282
				(0.160)	(0.093)	(0.323)	(0.291)
Pro Old / Con Young	195	0	0	4.872	1.934	9.851	3.320
				(0.142)	(0.077)	(0.244)	(0.245)
Pro Young / Con Old	169	0	0	5.012	1.902	9.303	3.031
				(0.147)	(0.084)	(0.240)	(0.267)
Pro Young / Con Young	192	0	0	5.177	1.640	9.379	2.938
				(0.115)	(0.087)	(0.209)	(0.333)
Pro Young / Pro Old	194	1	0	5.593	1.671	10.926	3.897
				(0.116)	(0.105)	(0.280)	(0.297)

According to attitude polarization, the effects of both positive and negative information should be negative for opponents, positive for proponents, and either positive or negative for moderates. According to Bayesian reasoning, the effect of positive information should be positive and the effect of negative information should be negative for all three subject

types.

2.4.4 Results

The results of Study 2 are presented in Tables 2.5 and 2.6. Again focusing on the covariate-adjusted estimates, the video treatments had powerful effects for all three subject types. Positive information had positive and statistically significant effects on subjects' preferred minimum wage amount; negative information had strongly negative effects.

A similar pattern of response is evident in Table 2.6. All coefficients have the sign predicted by "directional Bayes." With the exception of negative information among opponents, all these coefficients are statistically significant.

Table 2.5: Effects of Information on Preferred Minimum Wage Amount

	Dependent Variable: T2 Amount					
	Among Opponents		Among Moderates		Among Proponents	
Pos. Info (0 to 1)	0.434	0.601*	1.878***	1.507***	1.274***	1.371***
	(0.511)	(0.333)	(0.366)	(0.295)	(0.364)	(0.294)
Neg. Info (0 to 1)	−0.611	−0.696**	−0.831***	−0.927***	−1.929***	−1.699***
	(0.504)	(0.308)	(0.185)	(0.195)	(0.314)	(0.300)
Condition: Placebo	−0.115	−0.605**	−0.353	−0.463*	−0.790*	−0.493**
	(0.525)	(0.252)	(0.271)	(0.248)	(0.416)	(0.236)
Constant	6.891	2.493	9.365	5.636	11.991	1.309
	(0.212)	(0.863)	(0.097)	(2.657)	(0.185)	(1.190)
Covariates	No	Yes	No	Yes	No	Yes
N	343	343	356	356	471	471
R^2	0.008	0.664	0.187	0.326	0.100	0.443

*p < .1; **p < .05; ***p < .01
Robust standard errors are in parentheses.
The information content of the Placebo condition is coded 0.
Covariates include T1 Amount, T1 Favor, age, gender, ideology, party ID, and education.

2.4.4.1 Persistence of Effects

In this study, too, the estimated effects persisted 10 days later. While the treatment effects diminish somewhat, they retain somewhere between 50% and 90% of their original strength.

Table 2.6: Effects of Information on Favoring Minimum Wage Raise

	Dependent Variable: T2 Favor					
	Among Opponents		Among Moderates		Among Proponents	
Pos. Info (0 to 1)	0.860***	0.769***	0.564***	0.553***	0.291***	0.351***
	(0.255)	(0.215)	(0.164)	(0.140)	(0.088)	(0.073)
Neg. Info (0 to 1)	−0.017	−0.130	−0.626***	−0.657***	−0.556***	−0.485***
	(0.248)	(0.186)	(0.212)	(0.215)	(0.183)	(0.163)
Condition: Placebo	0.699**	0.338	0.277	0.198	0.008	0.067
	(0.275)	(0.248)	(0.238)	(0.215)	(0.194)	(0.166)
Constant	2.968	1.722	5.344	3.107	6.343	2.662
	(0.114)	(0.531)	(0.081)	(1.221)	(0.053)	(0.527)
Covariates	No	Yes	No	Yes	No	Yes
N	343	343	356	356	471	471
R^2	0.045	0.467	0.068	0.181	0.060	0.329

*p < .1; **p < .05; ***p < .01
Robust standard errors are in parentheses.
The information content of the Placebo condition is coded 0.
Covariates include T1 Amount, T1 Favor, age, gender, ideology, party ID, and education.

2.5 Study 3: Global Warming

Study 3 examines Bayesian reasoning in the context of global warming, a hotly contested issue area characterized by invective from both sides. For a number of reasons, this subject ought to be a fertile ground for attitude polarization. First, it is a salient political issue with clear partisan implications. Second, it involves complex considerations about risk, the reliability of projections into the future, the motives of industry players, and the costs of collective action. And third, the "facts on the ground" are themselves changing, from periodically issued reports from the Intergovernmental Panel on Climate Change (IPCC) to directly observable weather patterns.

With these difficulties in mind, I sought treatments with the clearest possible presentation of the current scientific evidence: graphical summaries of various projection models' performance in predicting both past and future warming trends. These graphs are based on actual figures included in the technical summary of the IPCC's Fifth Assessment Report.[7] In order to capture the complexity of the issue, I included dependent variables covering

[7]See `http://www.climatechange2013.org/images/report/WG1AR5_TS_FINAL.pdf`, page 64.

beliefs about the causes of climate change, the perceived threat of global warming, and whether sacrifices will be necessary to reduce its effects (designed to tap into subjects' assessment of relative costs and benefits as well as risk perceptions).

2.5.1 Subjects

Subjects were drawn from the same pre-treatment survey used in Study 2, which measured baseline attitudes toward global warming. Of the 2,000 invited subjects, 1,690 responded. These 1,690 comprise the experimental sample.

2.5.2 Procedure

Subjects were randomly assigned to be presented with one of two versions of the IPCC graph.[8] First, all subjects were given an introductory note about predicting future warming trends by "training statistical models on historical data, then running those models forward." The graph, which plots simulations from 138 different models, shows both their performance with actual data from the past and their projections into the future. For the *Warming* treatment, respondents were then told, "There is a great deal of agreement between the models and the observed temperature. This agreement means that we can have greater confidence in the models' projections into the future. The Earth is warming, and the models predict that the warming trends will continue into the future."

For the *Warming Hiatus* treatment, the line plotting past temperature changes is brought partially forward into the "projected" side of the graph to show that the simulations generally seem to be overstating the warming trend. Subjects assigned to this treatment are then told, "To the left of the dotted line, the models and the observed temperature agree. But to the right of the dotted line, the actual levels of warming clearly disagree with the predictions of the models. This disagreement means we cannot be confident of the models' projections into the future. The shaded region describes the updated projections based on more recent data."

After some intermediate questions asking about generic political preferences, subjects

[8] See Appendix A for reproductions of the treatment materials.

responded to questions measuring the main outcome variables. The *Cause* question asked, "Some people believe that increases in the Earth's temperature over the last century are due to the effects of pollution from human activities. Others believe that temperature increases are due to natural changes in the environment. Which comes closer to your view?" The response options were "Temperature increases are mainly due to the effects of pollution from human activities," "... due to natural changes in the environment," "... due to both human activities and natural changes," and "I do not believe the Earth's temperature has risen." The *Threat* question asked, "Do you believe global warming will pose a serious threat to your way of life in your lifetime?" Respondents answered yes or no.

2.5.3 Analysis

All subjects in the experimental sample were exposed to one of two treatments, *Warming* (coded 0) or *Warming Hiatus* (coded 1). I recoded the *Cause* question to equal 1 if the respondent said that temperature increases were due to natural changes in the environment or both human activities and natural changes in the environment – this new variable is called *Natural Cause*. Similarly, the *Not a Threat* variable is the reverse-coded version of the *Threat* question. Because this experiment only involves two conditions, I estimate treatment effects with the difference-in-means (or regression-adjusted difference-in-means), as in Equation 2.2. Under attitude polarization, β_1 is hypothesized to be positive for global warming "skeptics" and negative among global warming "believers." Skeptic/Believer status was measured with pre-treatment versions of the dependent variable questions. Under Bayesian reasoning, β_1 should be nonnegative for all subjects.

$$Y_i = \beta_0 + \beta_1(Hiatus_i) + \epsilon_i \tag{2.2}$$

2.5.4 Results

Table 2.7 shows that the effects of the *Warming Hiatus* treatment were strongest among moderates, causing a 15.8-percentage-point increase in belief that the Earth is warming, at least in part due to natural causes. Among climate change believers, this effect was 13.3 percentage points. The group hypothesized by the attitude polarization theory to have

the largest average treatment effect – climate skeptics – in fact had the smallest, at 5.7 percentage points. This small coefficient may be in part due to a ceiling effect, as 85% of skeptics in the control group believe global warming is due to natural causes.

Table 2.7: Effects of Information on Belief that Global Warming is due to Natural Causes

	Dependent Variable: T2 Natural Cause					
	Among Skeptics		Among Moderates		Among Believers	
Treatment: Hiatus	0.053**	0.056**	0.166**	0.157**	0.113**	0.131**
	(0.035)	(0.034)	(0.042)	(0.042)	(0.036)	(0.034)
Constant	0.850**	0.901	0.574**	0.584	0.395**	0.692
	(0.028)	(0.113)	(0.030)	(0.137)	(0.025)	(0.112)
Covariates	No	Yes	No	Yes	No	Yes
N	354	354	495	495	773	773
R^2	0.007	0.120	0.030	0.131	0.013	0.120

*p < .1; **p < .05; ***p < .01
Robust standard errors are in parentheses.
Covariates: T1 Natural Cause, T1 Not a Threat, age, gender, ideology, party ID, education.

Table 2.8 shows that the treatments had minimal effects on the belief that global warming is not a threat. While all the average treatment effects are estimated to be positive, none is statistically significant. When pooled, the average effect for the entire sample is estimated to 0.016, with a standard error of 0.019. It appears that subjects changed their minds as to the probable cause of global warming, but not with respect to how dangerous it is to their way of life. This finding is puzzling, given that the evidence presented to subjects in the *Warming Hiatus* condition showed that temperatures had not risen as much as projected – evidence that speaks not to the causes of global warming, but rather its extent. This pattern cannot be explained by a ceiling effect, at least among moderates and believers: baseline belief that global warming is not a threat was 47.4 and 27.9 percent, respectively. One possible (admittedly post-hoc) explanation is that treated subjects might have interpreted the revised projections as evidence of the uncertainty attending to climate science – just because it is difficult to know how much warming will happen and why does not mean that global warming is not dangerous.

Table 2.8: Effects of Information on Belief that Global Warming is not a Threat

	Dependent Variable: T2 Not a Threat					
	Among Skeptics		Among Moderates		Among Believers	
Treatment: Hiatus	0.004	0.003	0.023	0.023	0.028	0.027
	(0.030)	(0.030)	(0.045)	(0.039)	(0.033)	(0.028)
Constant	0.910	0.815	0.474	0.750	0.278	0.859
	(0.022)	(0.094)	(0.030)	(0.126)	(0.023)	(0.092)
Covariates	No	Yes	No	Yes	No	Yes
N	354	354	495	495	773	773
R^2	0.0001	0.156	0.001	0.288	0.001	0.280

$^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$
Robust standard errors are in parentheses.
Covariates: T1 Natural Cause, T1 Not a Threat, age, gender, ideology, party ID, education.

2.6 Exploring Treatment Effect Heterogeneity with BART

The difference between attitude polarization and Bayesian reasoning comes down to divergent predictions about treatment effect heterogeneity. Under attitude polarization, treatment effects are hypothesized to be positive for some subgroups but negative for others. Under Bayesian reasoning, treatment effects are hypothesized to be strictly non-negative for all subgroups. In the analyses presented above, I explored treatment effect heterogeneity according to pre-treatment measures of support for the issue area. While this choice was theoretically motivated, it is only one of a multitude of pre-treatment covariates that could be used to check for differences in responses to treatment. The Bayesian reasoning hypothesis is a strong one: non-negative treatment effects for *all* subgroups.

In order to fully evaluate this hypothesis, a method that takes a comprehensive approach to detecting differential treatment effects is required. Bayesian Additive Regression Trees (BART) is a recent advance in statistical learning (Chipman et al. 2007, 2010) that has been recommended by social scientists (Hill 2011; Green and Kern 2012) as a method for flexibly and automatically detecting treatment effect heterogeneity. BART is a sum-of-trees model that makes robust predictions of the conditional mean of the outcome variable without overfitting. A principal benefit of using BART over other machine learning algorithms is that it is robust to the choice of tuning parameters; I use the default settings in the `dbarts`

package for `R`, version 0.8-5.

I implemented the procedure recommended by Hill (2011), first fitting the model to the observed data, then using the model to make predictions about the outcome that each unit would express under each treatment condition. In Studies 1 and 2, I focused on the difference between the *Pro Pro* and *Con Con* conditions. These differences can be thought of as an estimate of the conditional average treatment effect (CATE) for each unique covariate profile. There are 494 unique covariate profiles among the 683 subjects in Study 1, 909 among the 1,170 subjects in Study 2, and 815 among the 1,690 subjects in Study 3. A fully saturated OLS model with all possible treatment-by-covariate interactions would also produce estimates for each covariate profile, but these estimates would be extremely noisy, and moreover, there might not be variation on the treatment indicator within each cell. In essence, the BART procedure partially pools across cells to achieve greater precision. Estimates of uncertainty are obtained by repeatedly drawing from the posterior and generating 95% credible intervals.

Standard statistical methods such as OLS have the advantage of providing relatively simple data summaries. The average effect of treatment can be summarized as the coefficient on the treatment indicator and heterogeneity can be characterized by the coefficients on interaction terms. BART models, by contrast, cannot easily be summarized by a series of coefficients. In order to understand the heterogeneity discovered by BART, it is best to inspect the model graphically. In the graphs that follow, I plot both estimated potential outcomes for each unit in the sample along with a 95% credible interval on the vertical axes. On the horizontal axes, I sort the units according to the average of their estimated potential outcomes. That is, if a unit is estimated to express a relatively low outcome regardless of treatment condition, that unit is plotted further to the left. This choice allows for a quick grasp of the general pattern: do the response surfaces generally move in parallel, do they cross, or do they diverge?

Figure 2.1 presents the BART-estimated potential outcomes for both dependent variables in all three studies. The first row shows estimates of treatment on the *Support* and *Belief* outcomes. The blue, circular points represents the model's best guess for the outcome

that each unit would express in the "Pro Pro" condition, whereas the red, triangular points represent the estimate of the outcome under the "Con Con" condition. The difference between these points is the estimated treatment effect for that unit. For *Support*, the average difference between the circular and triangular points is approximately 0.35. The difference is slightly larger for those who have generally higher support for capital punishment and slightly smaller for those who generally oppose it. For *Belief*, the average difference is about 0.92, and is approximately equal to this value, regardless of background characteristic.

The second row presents the output of the BART model for the *Amount* and *Favor* dependent variables in the minimum wage experiment. Turning first to the left panel, the blue, circular points represent the estimated preferred minimum wage for each subject under the "Pro Pro" condition, while the red, triangular points represent the estimated outcomes if subjects had seen both "Con" videos. For those subjects generally opposed to the minimum wage, the average difference between the curves is about $1.00, but it increases to about $4.00 at the high end of the scale. Here we see clear evidence of treatment effect heterogeneity. By contrast, the curves for the *Favor* dependent variable could hardly be more parallel. Nearly all subjects appear to have a treatment effect very close to 1.0.

Both the *Natural Cause* and *Not a Threat* dependent variables in the global warming study are binary, so I express the outcomes that each unit would express under each treatment condition in terms of a probability. For binary outcomes, BART employs a version of a probit model: I have translated these probits back into probabilities for ease of interpretation. For *Natural Cause*, we see large differences in the estimated probabilities on the order of a 0.10 difference in the probability of believing that global warming is due to natural causes. This difference gets smaller as we approach zero or unit probability. Interestingly, this pattern does not arise when inspecting the treatment versus control differences in probits: the differences appear far more parallel in probits than in probabilities (analysis not reported here). As noted in Chapter 1, the scale of the dependent variable has strong implications for treatment effect heterogeneity. In probits, the treatment effects are homogeneous but in probabilities, the effects are heterogeneous. In terms of evaluating the Bayesian reasoning hypothesis, however, the question of the correct outcome scale is

less important: in either scale, treatment effects are all estimated to be positive. The *Not a Threat* dependent variable is plotted in the lower right hand panel and shows that across the entire response surface, there is very little difference between the treated and control potential outcomes. Not only does there appear to be no average effect, there does not not appear to be any effect among any covariate profile in the sample.

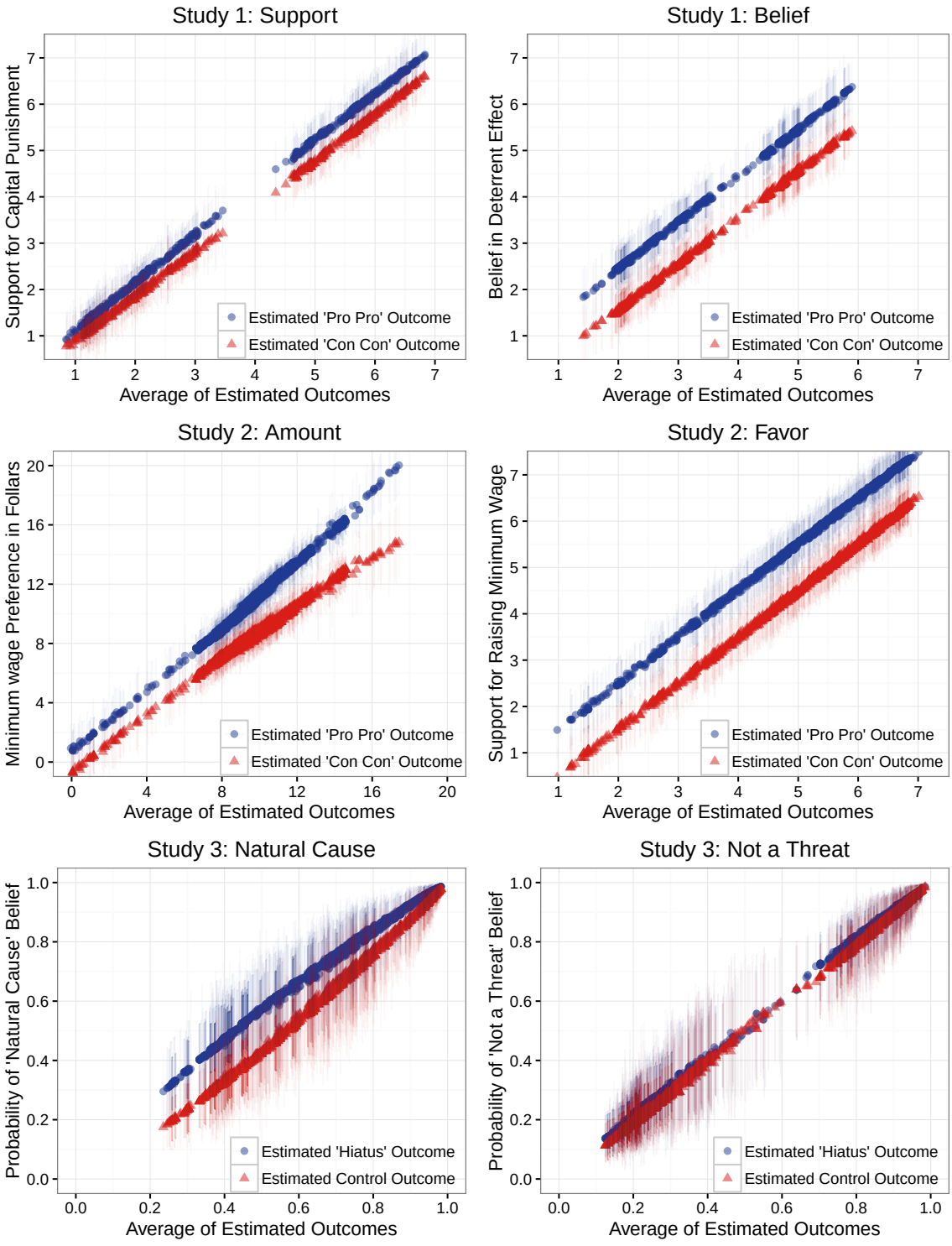
Figure 2.1 echoes many of the findings highlighted by the OLS models presented above. For example, in the Capital Punishment experiment, treatment had slightly bigger effects on "Support" among proponents than among opponents. Whereas the OLS model focuses attention on how effects were statistically significant for one group but not the other, the BART analysis emphasizes that the response surfaces, while not exactly parallel, do not diverge wildly for subjects with different backgrounds.

Table 2.9 summarizes the BART models for all six dependent variables presented in the previous sections. The estimated ATE is the average of the estimates of the individual-level treatment effects. The SE is the standard deviation of the distribution of estimated ATEs across 1,000 posterior draws. The 95% credible interval is the 2.5^{th} and 97.5^{th} percentiles of that distribution. These BART summaries are very similar to those returned by OLS. The last column shows the proportion of estimated treatment effects across all 1,000 posterior draws that were non-negative. With the exception of *Threat*, which is essentially unmoved by the treatment, we observe almost exclusively positive treatment effects. Regardless of the subgroup analyzed, there are almost no subjects who are estimated to experience negative treatment effects, providing strong evidence that attitude polarization was not at work in these experiments.

Table 2.9: Summaries of BART Estimation

	$\widehat{ATE}$	SE	95% CI	$Pr(\hat{\tau}_i \geq 0)$
Study 1: *Belief*	0.923	0.137	[0.654, 1.208]	1.000
Study 1: *Support*	0.355	0.111	[0.134, 0.572]	0.938
Study 2: *Amount*	2.458	0.204	[2.070, 2.834]	0.998
Study 2: *Favor*	1.014	0.110	[0.800, 1.240]	1.000
Study 3: *Natural Cause*	0.110	0.021	[0.067, 0.154]	0.971
Study 3: *Not a Threat*	0.017	0.019	[-0.019, 0.053]	0.685

Figure 2.1: BART Estimated Potential Outcomes



2.7 Relationship of these Results to Previous Literature

The foregoing experimental evidence is seemingly at odds with the previous literature describing how individuals evaluate information and update their beliefs. There are, in my view, at least two possible explanations for this discrepancy. First, it is possible that the conditions for attitude polarization that *were* present in previous experiments were not present in these experiments. This is unlikely to be the case, first and foremost because the results of study 1 are nearly identical to the original Lord et al. (1979) study when the data are analyzed as those authors did. Further, the mechanisms hypothesized to drive attitude polarization are often described in general terms: Taber and Lodge (2006, p. 756) take as their starting premise the claim that "all reasoning is motivated." Lord et al. (1979, p. 2108) worry that due to biased assimilation, evidence will "frequently fuel rather than calm the fires of debate." While many contexts may be more politically contentious than the survey experimental environment, it would seem to be an arbitrary domain restriction to rule these experiments as outside the theoretical scope.

A second possibility for reconciling these findings with the work of others is that previous evidence has been misinterpreted. The Lord et al. study is suspect because the analyses condition not on a randomly assigned treatment, but rather on subjects' background characteristics. The observed correlation between opponent/proponent status and the self-reported measure of change is not a good estimate of the effect of information; it is biased due to unobserved heterogeneity.

A related study by Taber and Lodge (2006) uses an information board experiment to demonstrate attitude polarization. In an information board experiment, the information to which subjects are exposed is not in control of the researcher. Instead, subjects are presented with a series articles they can choose to read or not. The main evidence for attitude polarization in Taber and Lodge (2006) is a series of regressions of T2 attitude extremity on T1 attitude extremity, subgroup by subgroup. Regression coefficients greater than 1 are taken to indicate a strengthening of attitudes at T2 over T1. To be clear,

this study did not condition its analyses according to a random assignment[9] The authors themselves seem to agree that such a design is prone to bias: "Since several independent variables are measured rather than manipulated (prior attitude and sophistication), this is more properly thought of as a quasi-experimental design" (p. 737).

Redlawsk, Civettini and Emmerson (2010) reports the results of a study that also makes use of an information board but randomly assigns subjects to varying doses of counter-attitudinal information. Subjects could still choose to read some articles or not, but the "media environment" in which subjects made their choices was controlled by the researchers. The premise of that article is finding the "affective tipping point," or the dose of counter-attitudinal information that will induce subjects to start moving in the direction of evidence. Those authors hypothesize (p. 564) that small doses of counter-attitudinal information will have negative treatment effects (movements away from evidence) but that at some point, the dosage will be large enough to induce positive treatment effects (movements toward evidence). The results of that study do not convincingly support the hypothesis that small doses produce negative treatment effects. While the authors do not emphasize this point (it is made in a footnote on page 579), the effects of small doses are too small to be distinguishable from zero. The effects of large doses are positive and significant. In sum, the evidence from these information board experiments corroborates the findings in this chapter: subjects, when exposed to counter-attitudinal information, update their priors in the direction of evidence.

I am aware of two studies that do demonstrate a negative treatment effect for a subgroup when conditioning on a randomly assigned treatment. Nyhan et al. (2014) show that providing a "correction" about vaccine misperceptions can decrease vaccine-averse subjects' reported intention to vaccinate; this finding was replicated in Nyhan and Reifler (2015). While I have no reason to doubt the validity of this effect, I do interpret it in light of the studies' other findings. The correction decreases vaccine-averse subjects' stated belief that vaccines cause the flu (Nyhan and Reifler 2015, p. 462) and decreases their belief that

[9]The authors did randomize subjects into one of two studies – affirmative action or gun control – but differences in outcomes due to this random assignment are not presented.

vaccines cause autism. The corrections had unintended consequences among a subgroup, but subjects in that subgroup *did* update their beliefs in the direction of evidence. For a more thorough description of these results, see Appendix C.

A related literature considers "partisan motivated reasoning" (Leeper and Slothuus 2014) in which party-branded messages are sometimes (but not universally) found to cause positive effects for co-partisans but negative effects for out-partisans (Arceneaux and Johnson 2013; Druckman, Peterson and Slothuus 2013; Druckman, Levendusky and McLain 2015). I view these effects as being consistent with Bayesian reasoning insofar as subjects hold differing likelihood functions. Democrats, for example, may view a Republican-branded argument for a policy as a bundle of two pieces of information: the argument itself (which may be presumed to have a positive effect) and the partisan cue, which, according to the Democrats' likelihood functions, decreases the posterior probability that the policy is a good one. The negative effect of the partisan cue may frequently overwhelm the positive effect of information, but this does not constitute evidence against Bayesian reasoning as I have conceived of it here.

2.8 Discussion

At the heart of the attitude polarization theory is the prediction that individuals evaluate evidence relative to their pre-existing beliefs. I argue that this connection is best understood as part of a Bayesian learning framework rather than a belief-preserving defensive mechanism.

The experimental evidence is consistent across all three of the studies reported here. Pro-capital-punishment evidence tends to make subjects more supportive of the death penalty and to strengthen their beliefs in its deterrent efficacy. Evidence that the death penalty increases crime does the opposite, while inconclusive evidence does not cause significant shifts in attitudes or beliefs. Arguments about the minimum wage likewise move respondents in the predicted direction – toward supporting a higher or a lower dollar amount according to the slant of the evidence presented in the video treatments. Finally, scientific projections about the likely trajectory of warming trends have a significant effect on subjects' views

about the causes of climate change. These findings are strongly in line with the "directional Bayes" prediction: individuals' beliefs are anchored by their priors and refreshed in the direction of the evidence.

The studies reported here were designed to encompass a variety of different types of information – scientific evidence described in tables and graphs in addition to more colloquial video appeals. The issues covered vary along numerous dimensions: both "hot" (death penalty) and "cold" (minimum wage), easily mapped to the partisan divide (global warming) or not, and of varying degrees of salience. The results do not depend on the particular issue or idiosyncratic features of the topics chosen. Importantly, the effects are not fleeting: In the two studies in which I collected follow-up responses, subjects' updated beliefs persisted for at least 10 days after the initial experiment.

The experimental results reveal a surprising *lack* of treatment effect heterogeneity. Many theories of persuasion predict that different groups will respond to information differently. The direction and magnitude of the effect of new evidence depends on people's prior beliefs, their level of knowledge, or other factors such as whether they have a personal stake in the issue. Contrary to such expectations, I find that subjects update in a parallel fashion regardless of predispositions, level of education, or ideology. This lack of heterogeneity raises important questions for the portion of the Bayesian theory that predicts differences in effects depending on the strength of priors or idiosyncrasies of individuals' likelihood functions. The overall similarity of effects across subjects suggests that people might not differ much with respect to the strength of their priors. That is, the "prior variance" of policy attitudes, if we could measure it, might be similar across individuals.

These experiments show that when people are exposed to information, that they do indeed update their views accordingly. However, one way in which these findings might not generalize to non-experimental contexts is if people selectively avoid counter-attitudinal information. Prior (2007) and Arceneaux and Johnson (2013) find that many individuals, if given the choice, simply consume entertainment rather than news information, thereby selecting out of both pro- and counter-attitudinal information in one stroke. On the other hand, (Bakshy, Messing and Adamic 2015) shows that while partisan Facebook users do

appear to prefer pro-attitudinal news stories, they are exposed to and do consume a large amount of counter-attitudinal information. Future research should consider the conditions under which individuals could be induced to seek out larger or smaller doses of information with which they disagree.

A reasonable objection to these findings is that while subjects may not polarize when reading relatively sterile descriptions of academic studies, individuals may do so when actually arguing about some proposition with an opponent. Political disputes linger and do not easily resolve when new information comes to light. One possibility is that, under some conditions, individuals' likelihood functions may not follow the monotone likelihood ratio property, the restriction that everyone has compatible expectations of the meaning of evidence. I speculate that in contentious political environments, in which opposing sides routinely insult the other (or much worse), the introduction of evidence could induce a divergence in attitudes. Perhaps in such antagonistic contexts, individuals become distrustful of counter-attitudinal arguments. These and other conditions under which attitude polarization might occur are the subject of Chapter 3.

3 The Rule, not the Exception: Bayesian Reasoning in Contentious Environments

In the previous chapter, I showed, contra the attitude polarization hypothesis, that individuals exhibit Bayesian reasoning when encountering counter-attitudinal arguments and evidence; that is, they update their beliefs in the direction of evidence. An intriguing possibility is that the ***conditions*** *for attitude polarization were lacking in the relatively sterile survey experimental environment. This chapter explores this possibility by investigating individuals' response to information in a wider variety of contexts. I randomize, along with the presentation of attitude-consistent and -inconsistent evidence, a contextual factor hypothesized to trigger motivated reasoning: ad hominem discourse. I show that insulting language does not inhibit individuals' capacity to incorporate new information. I conclude from these results that even in contentious environments, information processing consistent with Bayesian reasoning remains possible.*

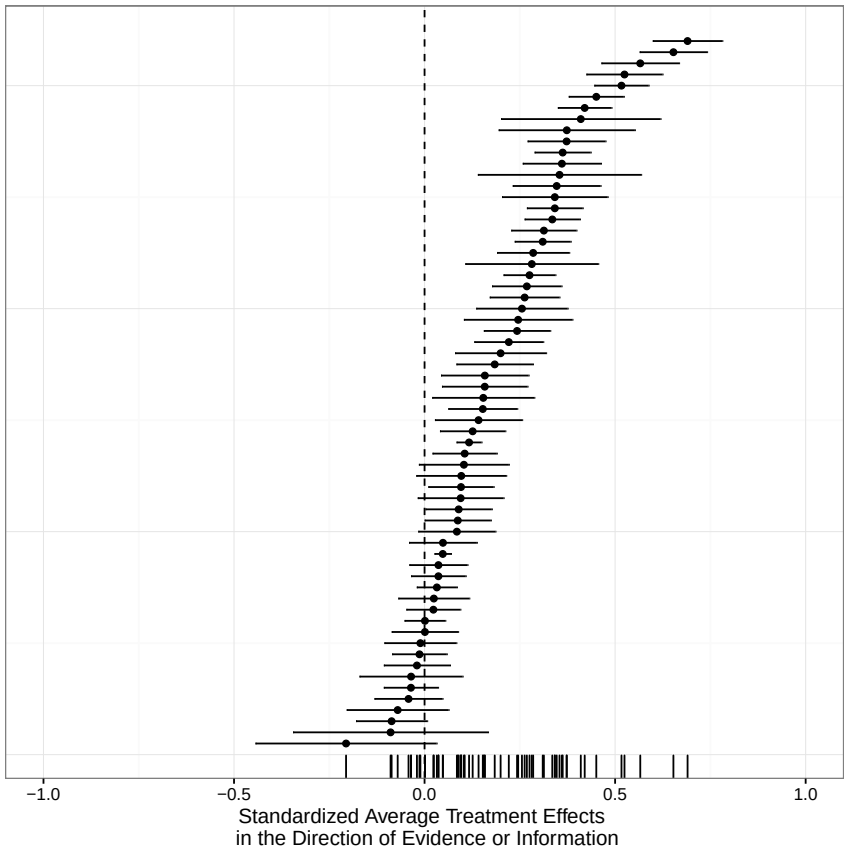
This chapter draws on joint work with Andrew Guess.

The suite of theories falling under the heading "motivated reasoning" (Kunda 1990) is frequently offered as a theoretical explanation for attitude polarization or "backlash"(e.g., Taber and Lodge 2006). Attitude polarization is a causal process in which individuals, in response to some stimulus or by engaging in some activity, move away from the "center" of opinion and toward the extremes. Describing the result of some treatment as attitude polarization presumes a one-dimensional attitude space, with some individuals holding a positive attitude toward a stimulus object and others holding a negative attitude. Attitude polarization occurs when negative attitudes become more negative and positive attitudes become more positive.

Backlash, at best, appears to be a very rare phenomenon. In chapter 1, I showed how on a large range of issues, individuals update in parallel. This pattern, first observed at the macro level in (Page and Shapiro 1992), also obtained at the micro level in a series of 20 survey experiments. In chapter 2, I forwarded a theory of Bayesian rationality that predicts these changes. Individuals have prior beliefs, prior uncertainty, and a likelihood function they use to process information. Individuals are allowed any prior belief or uncertainty. Given a small restriction on the likelihood function (the minimum likelihood ratio property), individuals are predicted to update either in the direction of evidence or not at all, but are never predicted to update away from evidence.

Figure 3.1 shows that backlash does not occur on average. Treatment effect estimates and 95% confidence intervals from the 20 survey experiments discussed in Chapter 1 are plotted on the horizontal axis. The effects are ordered on the vertical axis by magnitude. The polarity of the treatment effects matches the "direction of information." By the direction of information, I mean that differences are scored relative to the sign of the hypothesized effect. In most cases, this direction is easy to discern, e.g., pro-capital punishment information should move subjects in a positive direction. The figure shows that the majority of the treatment effects are estimated to be positive. Among those that are estimated to be negative, not a single treatment effect estimate is statistically significant. At least on average, no "incorrectly" signed statistically significant effects were observed in this set of 20 experiments.

Figure 3.1: Average Treatment Effects in 20 Survey Experiments



While the average effects of information treatments appear to be correctly signed in most cases, there remains the possibility that under some conditions and for some subgroups, a particular information treatment might cause a negative effect. In this chapter, I attempt to *induce* backlash effects by assigning subjects to reason in contentious information environments. In an uncontentious control condition, subjects are randomly exposed to evidence as in a standard information experiment. In a "contentious" condition, subjects are insulted with ad hominem discourse designed to reproduce the contentious environments in which political discourse often takes place – environments in which political observers often lament the ubiquity of polarized debate, motivated reasoning, and backlash (Mutz and Reeves 2005).

To preview the results, I find strong evidence that insults do not cause backlash. Subjects in both processing conditions respond to evidence in the same way: small but perceptible shifts in the direction of evidence.

3.1 Theory

To briefly recapitulate the theory presented in Chapter 2, Bayesian reasoning predicts that individuals will update their attitudes and beliefs in accordance with the evidence provided to them. As discussed in the previous chapter, there are three dimensions along which individuals can vary: their prior beliefs, prior uncertainties, and their likelihood functions. Individuals' posterior beliefs are analogous to weighted averages of their priors and evidence, where the weights are determined by their prior uncertainties and likelihood functions. Individual with high prior uncertainty will be persuaded relatively more than those with low prior uncertainty. Individuals whose likelihood functions lead them to regard evidence with skepticism will be persuaded less than those with more forgiving likelihood functions. In this framework, unless their likelihood functions are somehow "backwards," (for example, they interpret a positive jobs report as evidence of a flagging economy), people will either update their beliefs in the direction of evidence or not update at all. Individuals are never predicted to update *away* from the evidence presented to them.

These predictions stand in contrast to those made in the literature surrounding attitude

polarization and backlash (Taber and Lodge 2006; Lord et al. 1979). Motivated reasoning is the label given to a class of models of human information processing in which biases due to directional goals inhibit individuals' capacity for dispassionate evaluation of evidence. Many phenomena can be explained either using a model of motivated reasoning or Bayesian reasoning; some phenomena predicted under motivated reasoning, however, are not predicted by Bayesian reasoning. Attitude polarization or backlash is one such phenomenon.

Under one version of motivated reasoning, backlash occurs in the following way: Individuals encounter counter-attitudinal evidence. They then spend cognitive energy "counter-arguing" the evidence presented to them. This process results in individuals becoming more convinced of their original position than they would be had no evidence been presented at all. In sum, the treatment effect of counter-attitudinal evidence is predicted to be *negative*.

The empirical literature on the effects of information has afforded few unambiguous examples of information backlash. One incontrovertible example is that pro-vaccine information leads anti-vaccine subjects to report lower intention-to-vaccinate (Nyhan et al. (2014), result replicated in Nyhan and Reifler (2015)), though those studies also found positive effects on belief in vaccine efficacy for all subjects (See Appendix C for details of these studies). Many other demonstrations of backlash contain design and analysis choices (lack of random assignment, conditioning on post-treatment variables, multiple comparisons) that render results difficult to interpret.

Ad hominem remarks or insulting language in general may make subjects less likely to receive and accept persuasive messages. There are at least two possible mechanisms by which this effect may occur. The first is that ad hominem remarks may trigger "hot cognition," a hypothesis placing primacy on affect and post-hoc rationalization of preexisting feelings toward an attitude object – in other words, inducing a strong directional motivation in subjects' reasoning (Redlawsk 2002). By this account, insulted subjects might be more reluctant to back down or change their minds in response to evidence given the perceived stakes and the nature of the opposition. The second possible channel operates via the likelihood function. An insult may color how subjects perceive the source of the insult and the caliber of the information that source later produces. Subjects may reason that an

insulting source may not have good information, or may be attempting to manipulate or mislead them.

There exists some empirical evidence for the hypothesized effects of ad hominem discourse. Abelson and Miller (1967) reports the results of an early field experiment conducted with confederates in Washington Square Park in New York City. Subjects were invited to take part in a survey on racial attitudes; a confederate posing as another survey respondent (randomly) either insulted the subjects for their views or did not insult them. This experiment is extraordinary enough that it is worth quoting Abelson and Miller's description of the insult treatment at length:

> In the Insulting remarks variation [the confederate] preceded all of his statements with insults to the subject. He had a list of five standard insults from which he could draw the one which seemed the most appropriate for each statement, provided that he used the majority of them for each subject. They were: (1) 'That's ridiculous'; (2) 'That's just the sort of thing you'd expect to hear in this park'; (3) 'That's obviously wrong'; (4) 'That's terribly confused'; and (5) 'No one really believes that.' In practice, he used all five insults for all subjects, for it was found that they were general enough that they could be said in a natural way almost regardless of what the subject had said.

Pre-treatment, the subjects all expressed pro-racial equality attitudes. In both the insult and control conditions, the confederate expressed anti-equality views; the main dependent variable was the extent to which the subject then reported agreeing with the confederate's statements. In the control condition, the average agreement was 16 on a 30 point agreement scale; in the insults condition, average agreement was 7.5. My reading of this experiment is that subjects were less likely to report agreeing with the confederate when they liked him less, i.e., when they report a lower agreement score, they are registering their dislike of the confederate, not necessarily their honest evaluation of his views. However, the evidence from this experiment could provide support for either of the proposed mechanisms by which ad hominem discourse is supposed to impede persuasion.

3.1.1 Predictions

Figure 3.2 displays the predictions of the theories of Bayesian and motivated reasoning for the patterns of treatment versus control response we should see, by information processing condition. The plot contains four facets. The top row of facets presents the predictions of Bayesian reasoning while the bottom row shows the predictions of motivated reasoning. The left column corresponds to processing conditions that are "uncontentious."

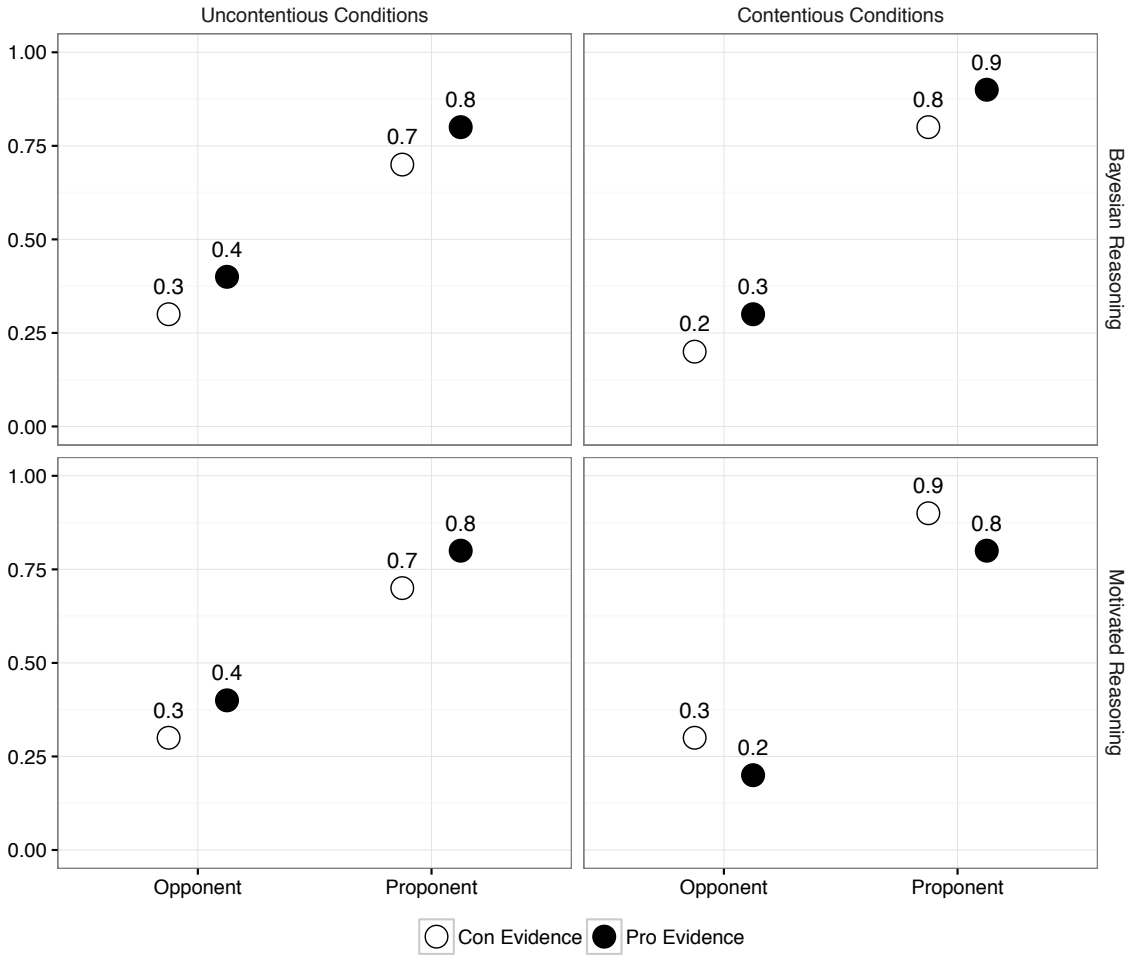
Beginning with the top left panel, the Bayesian theory of information processing predicts that under uncontentious conditions, subjects will update their views in the direction of evidence: while opponents and proponents hold different baseline views of the issue, the average outcome when treated with con evidence is lower for both groups than when treated with pro evidence. This is the pattern of evidence that has obtained in Chapters 1 and 2: subjects update their views in line with the information presented to them. However, those who assert that reasoning is motivated by directional goals might say that theories of motivated reasoning would predict the same pattern. Because the information environment is uncontentious, the impediments to the standard incorporation of evidence may not have been present. For this reason, the predictions in the lower left panel, corresponding to the outcomes predicted by motivated reasoning under uncontentious conditions, are identical to those presented in the upper left panel.

The distinction between the two theories, then, comes when the information processing environment becomes contentious. Under Bayesian reasoning (the top right panel), the difference between pro evidence and con evidence is still positive for both types of subjects, proponents and opponents alike. However, under motivated reasoning, the ordering changes: the difference is negative for both types.

Figure 3.2 shows that both theories are consistent with a direct effect of the contentious information environment on attitudes. For example, average opinions are higher for proponents and lower for opponents under contentious conditions compared with uncontentious conditions. The crucial distinction is how the information environments interact with subjects' ability to process information: under Bayesian reasoning, subjects update in the same direction regardless of the processing condition, whereas under motivated reasoning,

treatment effects are positive in uncontentious environments and negative in contentious environments.

Figure 3.2: Predicted Outcomes by Subject Type, Treatment, and Theory



3.2 Experimental Design

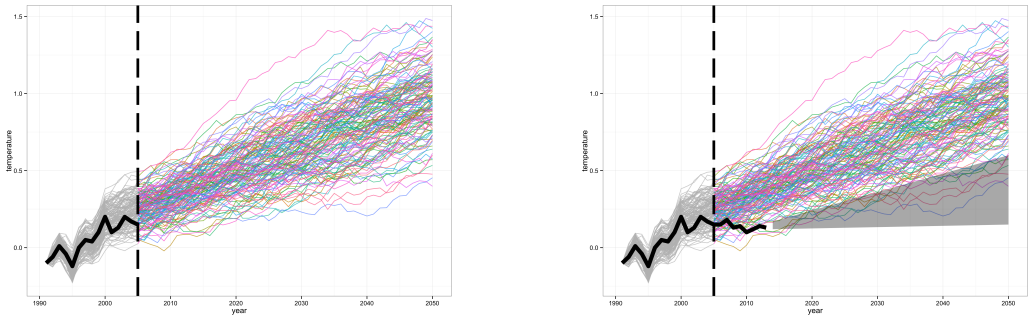
The overall design of the experiment is a two-arm information trial under two processing conditions. A key feature of this design is a pre-treatment measure of subjects' "believer" or "skeptic" status. Prior to randomization, I asked all subjects, "How strongly do you believe in the scientific evidence that global temperatures are rising due to human activ-

ity?" Subjects were given two response options: those who answered "Very strongly" were classified as believers, and those answering "Not very strongly" as skeptics.

3.2.1 Main Manipulation

The main information treatment consisted of an over time graph of global temperature predictions with accompanying explanatory text. In the control condition, subjects were presented with a graph of predictions (Figure 3.3a), showing general agreement among the world's climate scientists about the probable extent of temperature increases in the near term. In the "Haitus" treatment, subjects were shown a graph indicating how the observed temperature did not match predictions (Figure 3.3). These graphs are very similar to those used in Study 3 but are not identical. The full text of these treatments is included in Appendix A.

Figure 3.3: Main Manipulation



(a) Control

(b) Hiatus Treatment

3.2.2 Processing Condition Manipulations

In addition to the main manipulation, subjects were randomly assigned to one of two processing conditions:

- Control. In this condition, subjects were shown their main treatment graph with explanatory text. This condition corresponds to an "uncontentious" processing environment.
- Insults. In this condition, after answering the above pre-treatment question, subjects were exposed to an ad hominem remark. Believers were told: "Some people think

that your position makes you an alarmist supporter of big government. How do you respond?" Skeptics were told: "Some people think that your position makes you an unthinking opponent of science. How do you respond?" Subjects were given a text box in which to respond to these insults.

3.2.3 Subjects

Table 3.1 summarizes the number of believers and skeptics in each experimental condition. The sample contained relatively more believers than skeptics, though this in and of itself does not present a threat to inference. Because of the larger sample size for that group, estimates for believers will be more precise than those for skeptics.

Table 3.1: Number of Subjects in Each Condition

	Believers		Skeptics	
	Control	Insults	Control	Insults
Control Graph	210	216	77	51
Hiatus Graph	207	188	63	70

3.2.4 Dependent Variables

I used two main dependent variables, *Belief* and *Degrees*

- *Belief*: Since 1850, there is a well-documented increase in global temperatures. Some attribute this increase to human activity while others attribute it to natural causes. Which comes closest to your view?

  1. I believe that the increase is entirely due to natural causes
  2. I believe that the increase is mostly due to natural causes
  3. I believe that the increase is somewhat due to natural causes
  4. I believe that the increase is equally due to human activity and natural causes
  5. I believe that the increase is somewhat due to human activity
  6. I believe that the increase is mostly due to human activity
  7. I believe that the increase is entirely due to human activity

- *Degrees*: How many degrees do you believe the earth will warm over the next 100 years? (-1 to +3, slider)

These questions were of each subject twice, once immediately after the experimental manipulation and a second time approximately 10 days later in a follow-up survey.

3.3 Results

This section will present a series of analyses of increasing complexity, beginning with the estimates of average effects of the information treatment and ending with estimates of the conditional average treatment effects by skeptic/believer status and processing condition.

3.3.1 Average Treatment Effects

Table 3.2 presents estimates of the average treatment effect of the hiatus treatment on the *Belief* and *Degrees* dependent variables, both with and without covariate adjustment. The covariates include the processing conditions (insults or control) to which subjects were randomly assigned as well as standard pre-treatment measures of demographics. In the first column, I estimate that the hiatus treatment decreases the belief that global warming is due to human activity by 0.17 scale points using difference-in-means estimation ($p < 0.10$) The covariate-adjusted estimate, shown in the second column, is smaller in magnitude (0.10 scale point decrease, no longer statistically significant). I conclude from these results that the average effect on *Belief* is likely to be in the correct direction but small.

The effects of treatment on the *Degrees* variable are similar. Without covariate adjustment, the effect is estimated to be a difference of 0.11 degrees; with adjustment it is estimated to be 0.10 degrees. Both estimates are statistically significant at $p < 0.05$.

On average, the hiatus treatment causes subjects to become (marginally) more skeptical of the idea that global warming is due to human activity and to predict a (slightly) smaller increase in temperatures over the next 100 years.

3.3.2 Conditional Average Treatment Effects by Prior Beliefs

Many of the theoretical distinctions outlined above depend on subjects' priors before encountering evidence. Tables 3.3 and 3.4 present the experimental results for believers and skeptics separately. In columns 1 and 2 of Table 3.3, both the unadjusted and adjusted effects of the hiatus and insult treatments are presented. While covariate adjustment does decrease the standard errors slightly, it does not help a great deal. As above, the effect of the hiatus treatment is negative as is the effect of the insult treatment.

Table 3.2: Average Treatment Effects on Belief and Degrees

	Belief		Degrees	
	(1)	(2)	(3)	(4)
Hiatus Treatment	−0.173*	−0.095	−0.114**	−0.098**
	(0.094)	(0.069)	(0.049)	(0.043)
Constant (Control)	5.061	3.592	1.747	1.402
	(0.065)	(0.231)	(0.034)	(0.138)
Covariates	No	Yes	No	Yes
N	1,082	1,082	1,082	1,082
R^2	0.003	0.478	0.005	0.274

*p < .1; **p < .05; ***p < .01
Robust standard errors are in parentheses.

Columns 3 and 4 present the effects among skeptics. The effect of the hiatus treatment, while negative, is not statistically significant. The effect appears larger among skeptics than believers, but the difference between these coefficients (as tested with an interaction model) is not itself statistically significant, so we cannot make confident claims about the difference in effects for these two groups.

Table 3.3: Conditional Average Treatment Effects: Belief

	Belief in Human-Caused Global Warming			
	Believers	Believers	Skeptics	Skeptics
	(1)	(2)	(3)	(4)
Hiatus Treatment	−0.100	−0.077	−0.169	−0.082
	(0.071)	(0.070)	(0.191)	(0.187)
Insult Treatment	−0.076	−0.116*	−0.290	−0.228
	(0.071)	(0.070)	(0.191)	(0.183)
Constant (Control)	5.632	5.440	3.405	4.061
	(0.061)	(0.198)	(0.165)	(0.511)
Covariates	No	Yes	No	Yes
N	821	821	261	261
R^2	0.004	0.066	0.013	0.133

*p < .1; **p < .05; ***p < .01
Robust standard errors are in parentheses.

A similar pattern in presented in Table 3.4: The effects of the hiatus treatment are negative for both believers and skeptics but insignificant; the insult treatment does not exert effects that are large enough to be distinguished from zero.

Table 3.4: Conditional Average Treatment Effects: Degrees

	Predicted Increase in Global Temperatures			
	Believers	Believers	Skeptics	Skeptics
	(1)	(2)	(3)	(4)
Hiatus Treatment	−0.067	−0.059	−0.163	−0.157
	(0.047)	(0.046)	(0.105)	(0.107)
Insult Treatment	0.038	0.031	−0.122	−0.071
	(0.046)	(0.046)	(0.104)	(0.103)
Constant (Control)	1.926	2.087	1.135	1.847
	(0.040)	(0.129)	(0.087)	(0.287)
Covariates	No	Yes	No	Yes
N	821	821	261	261
R^2	0.003	0.047	0.017	0.087

*p < .1; **p < .05; ***p < .01
Robust standard errors are in parentheses.

3.3.3 Conditional Average Treatment Effects by Processing Condition and Prior Beliefs

The foregoing analysis considered the effects of the insult condition relative to the control condition, averaging over the values of the hiatus treatment. However, the main purpose of the study was to test for an *interaction* between a contentious reasoning environment and the treatment information. Allowing for differences by prior beliefs and processing condition requires a three-way interaction model: information by processing by prior beliefs. Such models can be difficult to interpret when they are presented in regression tables, so I have opted to present a series of conditional means in graphs in Figure 3.4, which follows a similar format to Figure 3.2 for easy comparison.

In Figure 3.4, each panel represents the conditional mean of skeptics and believers, by information condition and dependent variable. Most obviously, the baseline opinions of skeptics and believers differ dramatically: skeptics' average belief hovers around 3 where

as the value is approximately 5.5 among believers. Similar differences are apparent for the predictions about future warming. Turning now to the differences by information, we see that the differences are consistently in the correct direction. In no cases is the control mean higher than the hiatus treatment mean.

The first column of facets, corresponding to the control processing condition, looks very similar to the column of facets corresponding to the insults processing condition; formal statistical tests confirm this visual intuition. There do not appear to be any differences in subjects' response to information by processing condition – the insult preceding the allocation of treatment information did not alter how these subjects interpreted the global warming evidence.

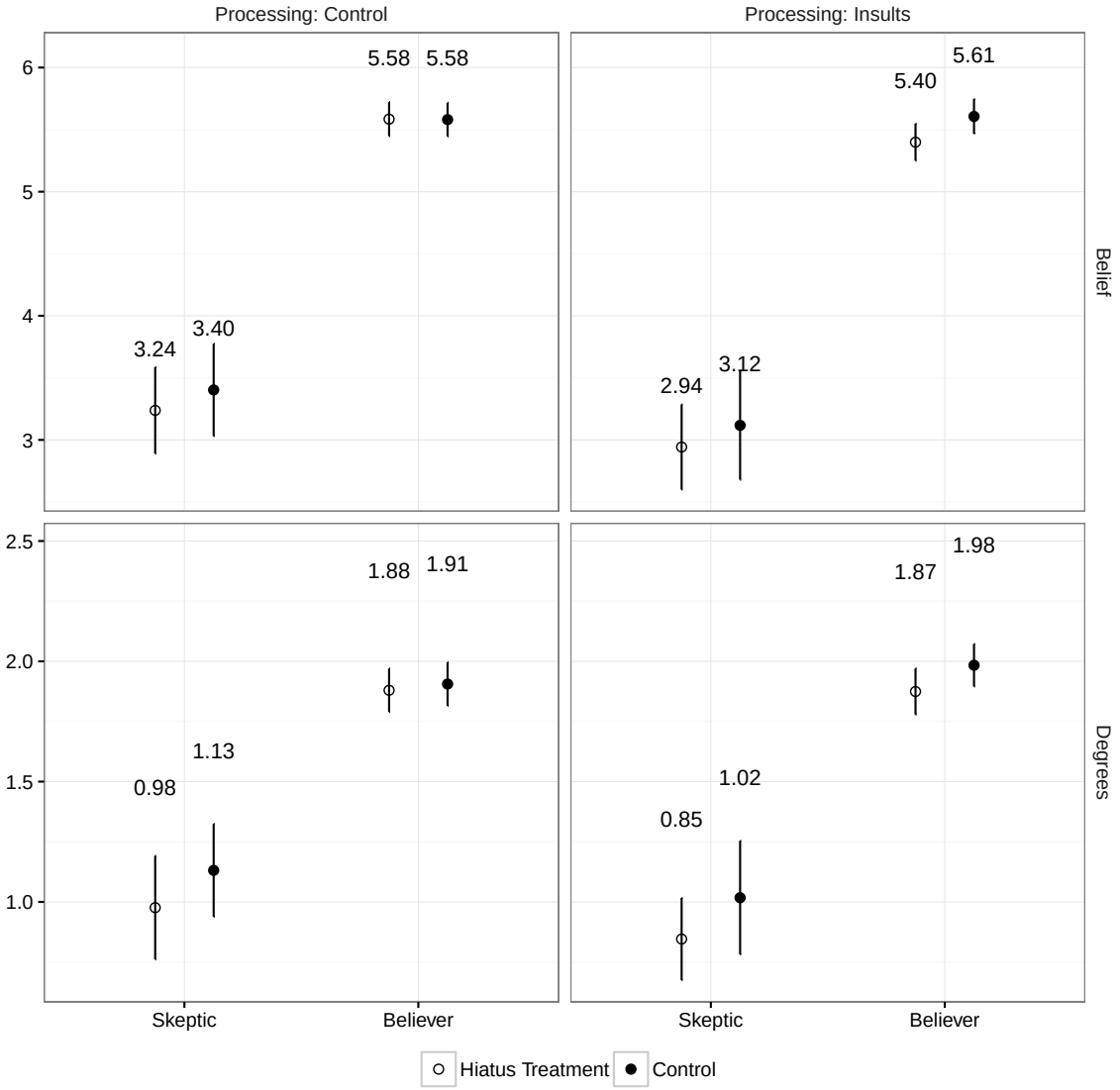
3.4 Conclusion

The premise that "all reasoning is motivated" (Taber and Lodge 2006, p. 756) might lead one to despair of the democratic process and of individuals' capacity for reason. Some worry that scientific evidence is counter-productive and that individuals, knowingly or not, might become further entrenched when presented with evidence they do not like. The results of this experiment show that this effect is difficult to uncover. As in the experiments reported in Chapters 1 and 2, individuals – regardless of their initial position *or the processing condition to which they were randomly assigned* – appeared to update their beliefs in the direction of evidence.

The results of this experiment provide further evidence that if backlash does occur, it occurs infrequently. The pattern of evidence appears to support the conclusion that "correct"responses are far and away the most common reaction to information treatments. As a description of everyday, commonplace responses to the sorts of persuasive messages that are prevalent in the media, Bayesian reasoning performs remarkably well.

The past three chapters have demonstrated that persuasive effects are generally small, positive, and homogeneous. The next two chapters explore the implications of these findings for two questions of methodological interest. First, how well do survey experimental estimates generalize from sample to sample? If effects are homogeneous, as these persuasive

Figure 3.4: Effects on Belief and Degree, by Type, Treatment, and Processing



effects appear to be, there are fewer dangers associated with extrapolating from one sample to another. Second, how long do these effects last? By conducting additional rounds of measurement, I am able to offer some preliminary results on the extent to which different treatments endure.

4 The Generalizability of Survey Experimental Results

To what extent do survey experimental treatment effect estimates generalize to other populations and contexts? Survey experiments conducted on convenience samples have often been criticized on the grounds that subjects are sufficiently different from the public at large to render the results of such experiments uninformative more broadly. In the presence of moderate treatment effect heterogeneity, however, such concerns may be allayed. I provide evidence from a series of 12 survey experiments that results derived from convenience samples like Amazon's Mechanical Turk are similar to those obtained from national samples. These results suggest that either the treatments deployed in these experiments cause similar responses for many subject types or convenience and national samples do not differ much with respect to treatment effect moderators. Using evidence of limited within-experiment heterogeneity, I show that the former is likely to be the case. Despite a wide diversity of background characteristics across samples, the effects uncovered in these experiments appear to be relatively homogeneous.

The generalizability of an experiment is the extent to which it generates knowledge about causal relationships that is applicable beyond the narrow confines of the research site. How much further beyond such confines results generalize is often a matter of great scientific importance, and vigorous discussion of the generalizability of findings occurs across all social scientific fields of inquiry (in economics: Levitt and List (2007); psychology: Sears (1986); Henrich, Heine and Norenzayan (2010), education: Tipton (2012); sociology: Lucas (2003); and political science: McDermott (2002)).

The generalizability of an experiment is closely related to its replicability. If an experiment is successfully replicated on a new sample that is drawn from a different population, we learn that the results of the original experiment generalize to the new population. In Clemens' (2015) typology of replications, such a demonstration is an "extension." If an experiment is successfully replicated on a new sample drawn from the same population (a "reproduction" in Clemens' terms), we have confirmed that the original results were less likely to be a fluke of sampling variability, but we have not learned much about the generalizability of the finding itself. Both extensions and replications involve the same treatments and outcome measures. By contrast, a "conceptual" replication alters both, thereby possibly changing the estimand. If a conceptual replication is successful, the extent to which we have learned that a finding generalizes depends heavily on the strength of the analogy between the treatments and outcome measures across studies. If a finding fails to replicate in repeated extensions and reproductions, we can conclude that the finding does not generalize (though we cannot, on this basis alone, conclude that the original finding was false or incorrect).

Concerns about the generalizability (and therefore replicability) of findings usually fall into one of two categories: criticisms of the experimental context and criticisms of the experimental subjects. The first concern is a worry that an effect measured under experimental conditions would not obtain in the "real world." For example, a classic critique of survey experiments that investigate priming is that the effect, while "real" in the sense that priming has been shown to happen in the lab, is unimportant for the study of politics because primes have fleeting effects and because political communication takes place in a

competitive space where the marginal impact of a prime is likely to be canceled by a counterprime (Chong and Druckman 2010). A similar question arises concerning the extent to which laboratory behavior generalizes to the field. Because subjects in the laboratory may respond to contextual clues and the obtrusive nature of the measurement, findings across lab and field might be expected to diverge. Contrary to this expectation, an analysis of the published record of paired lab and field studies finds a strong correlation of findings across settings (Coppock and Green 2015).

Within political science, a point of some contention has been the increase in the use of online convenience samples of experimental subjects, particularly samples obtained via Amazon's Mechanical Turk (MTurk).[1] Mechanical Turk respondents often complete dozens of academic surveys over the course of a week, leading to concerns that such "professional" subjects are particularly savvy or susceptible to demand effects (Chandler, Paolacci, Peer, Mueller and Ratliff in press). These worries have been tempered somewhat by empirical studies that find that the Mechanical Turk population responds in ways similar to other populations (e.g., Berinsky, Huber and Lenz (2012*b*)), but concerns remain that subjects on Mechanical Turk differ from other subjects in both measured and unmeasured ways.

In the present study, I contribute to the growing literature on the replicability of survey experiments across platforms, following in the footsteps of two closely-related studies. Mullinix et al. (2015) replicate 20 experiments and find a high degree of correspondence between estimates obtained on Mechanical Turk and on national probability samples, with a cross-sample correlation of 0.75 (XXX when corrected for measurement error).[2] Krupnikov and Levine (2014) find that when treatments are expected to have different effects by subgroups, cross-sample correspondence is weaker. The correlation between the MTurk and YouGov estimates is 0.41 (0.84 corrected).[3] To preview the findings presented below, I find

[1] As noted by Krupnikov and Levine (2014), criticisms of MTurk are often made on blogs rather than in academic journals. See, e.g., Kahan (2013).

[2] This figure obtained from private correspondence with the authors.

[3] I am grateful to Yanna Krupnikov for sharing the replication data for this study, from which this correlation was calculated. In the course of reanalyzing the data, it was discovered that the statistical routine originally used to compare samples overstated the confidence with which many of the YouGov/MTurk

a strong degree of correspondence between national probability samples and Mechanical Turk, with a cross-study correlation of 0.66 (0.83 corrected).

The remainder of this article will proceed as follows. First, I will present a theoretical framework that shows how the extent of treatment effect heterogeneity determines the generalizability of findings to other places and times. I will then present replication results from 12 studies, showing that in large part, original findings are replicated on both convenience and probability samples. I attribute this strong correspondence to the overall lack of treatment effect heterogeneity in these experiments; I conduct formal tests for unmodeled heterogeneity to bolster this claim.

4.1 Treatment Effect Heterogeneity and Generalizability

Findings from one site are generalizable to another site if the subjects, treatments, contexts, and outcome measures are the same – or would be the same – across both sites (Cronbach, Shapiro et al. 1982; Coppock and Green 2015). I define a *site* as the time, place, and group of units at and among which a causal process may hold. The most familiar kind of site is the research setting, with a well defined group of subjects, a given research protocol, and implementation team. Typically, the purpose of an experiment conducted at a given site is to generalize knowledge to other sites where no experiment has been conducted. For example, after an experiment conducted in one school district finds that a new curriculum is associated with large increases in student learning, we use that knowledge to infer both what would happen if the new curriculum were implemented in a different district, and what must have happened in the places where the curriculum is already in place.

The inferential target of a survey experiment conducted on a national probability sample of respondents is (often) the average treatment effect in the national population at a given moment in time, the Population Average Treatment Effect (PATE). The site of such an experiment might be an online survey administered to a random sample of adult Americans in 2012. The inferential target of the same experiment conducted on a convenience sample is a Sample Average Treatment Effect (SATE), where the sample in question is not drawn

differences could be deemed significant.

at random from the population. The SATE and the PATE are not, in general, equal to one another. If a treatment engenders *heterogeneous* effects such that the distribution of treatment effects among those in the convenience sample is different from the distribution in the population, then the SATE will likely be different from the PATE.

When treatments, contexts, and outcome measures are held constant across sites, the generalizability of results obtained from one site to other sites depends on the degree and nature of treatment effect heterogeneity. If treatments have constant effects (that is, treatment effects are homogeneous), then the peculiarities of the experimental sample are irrelevant: what is learned from any subgroup can be generalized to any other population of interest. When treatments have heterogeneous effects, then the experimental sample might be very consequential. In order to assert that findings from one site are relevant for another site, a researcher must have direct or indirect knowledge of the distribution of treatment effects in both sites.

To illustrate the relationship between generalizability and heterogeneity, consider Figure 4.1 below. It displays the potential outcomes and treatment effects for an entire population in two different scenarios. In the first scenario (represented in the left column of panels), treatment effects are heterogeneous, whereas in the second scenario, treatment effects are constant. The top row of panels displays the treated and untreated potential outcomes of the subjects and the bottom row displays the individual-level treatment effects (the difference between the treated and untreated potential outcomes).

An unobserved characteristic (U) of subjects is plotted on the horizontal axes of Figure 4.1. This characteristic represents something about subjects that is related to both their willingness to participate in survey experiments and their political attitudes. In both scenarios, U is negatively related to the untreated potential outcome (Y0): higher values are associated with lower values of Y0. This feature of the example represents how convenience samples may have different baseline political attitudes. Subjects on Mechanical Turk, for example, have been shown to hold more liberal values than the public at large (Berinsky et al. 2012*b*).

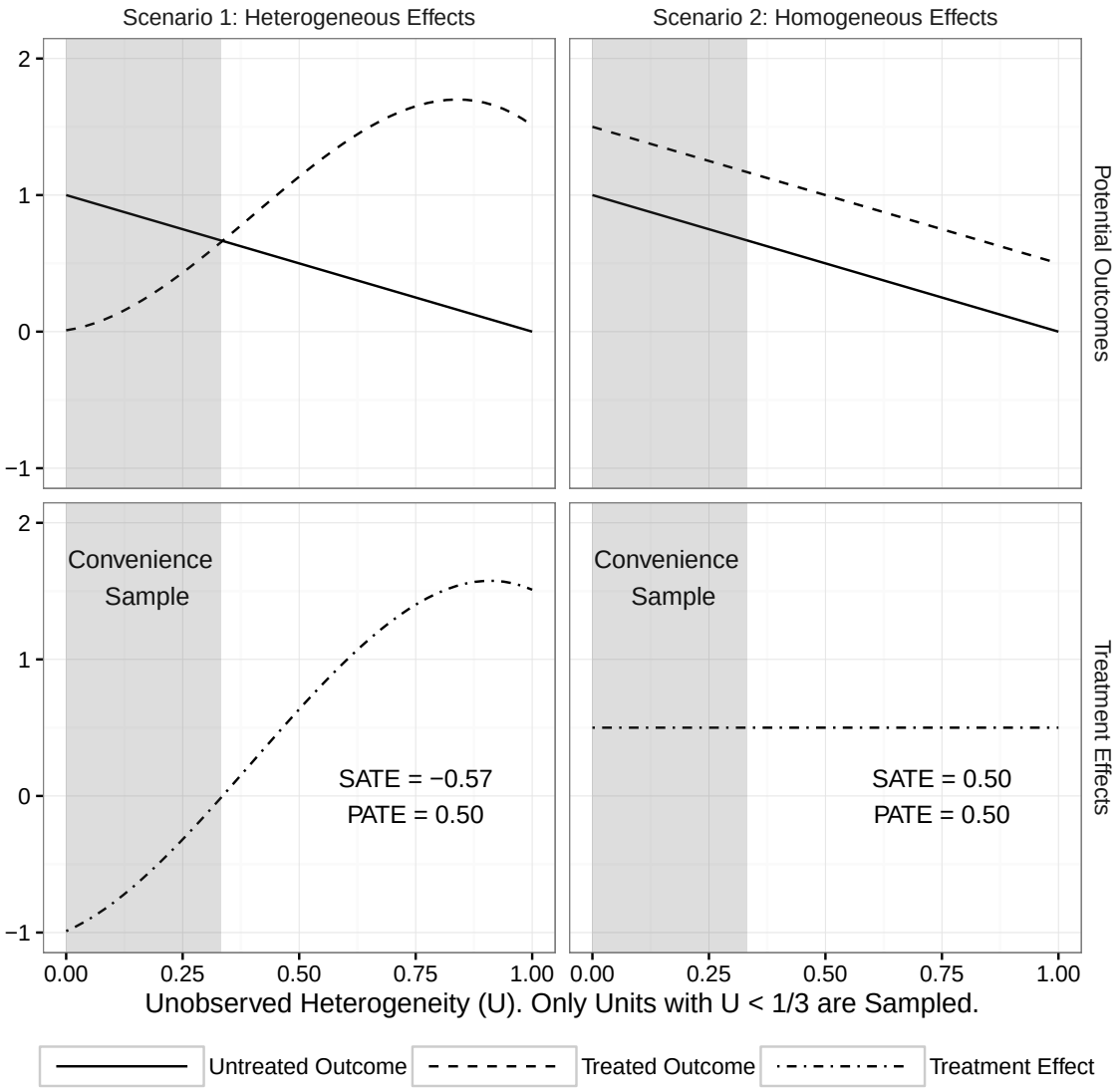
The untreated potential outcomes do not differ across scenarios, but the treated potential

outcomes do. In scenario 1, effects are heterogeneous: higher values of U are associated with higher treatment effects, though the relationship depicted here is nonlinear. In scenario 2, treatment effects are homogeneous, so the unobserved characteristic U is independent of the differences in potential outcomes.

In both scenarios, imagine that two studies are conducted: one that samples from the entire population and a second that uses a convenience sample indicated by the shaded region. The population-level study targets the PATE, whereas the study conducted with the convenience sample targets a SATE. In scenario 1, the SATE and the PATE are different: generalizing from one research site to the other would lead to incorrect inferences. Note that this is a two-way street: with only an estimate of the PATE in hand, a researcher would make poor inferences about the SATE and vice-versa. In scenario 2, the PATE and SATE are the same: generalizing from one site to another would be appropriate.

Figure 4.1 illustrates four points that are important to the study of generalizability. First, the fact that subjects may self-select into a study does not on its own mean that a study is not generalizable. Generalizability depends on whether treatment effect heterogeneity is independent of the (observed and unobserved) characteristics that determine self-selection. Second, even when *baseline* outcomes (Y0) are different in a self-selected sample from some population, the study may still be generalizable. The relevant theoretical question concerns the differences between potential outcomes (the treatment effects), not their levels. Thirdly, when the characteristics that distinguish the population from that sample are unobserved, the PATE may not be informative about the SATE, i.e., experiments that target the PATE should not be privileged over those that use convenience samples unless the PATE is truly the target of inference. Finally, it is important to distinguish systematic heterogeneity from idiosyncratic heterogeneity. When treatment effects vary systematically with measurable subject characteristics, then the generalizability problem reduces to measuring these characteristics, estimating conditional average treatment effects, then reweighting these estimates by post-stratification. This reweighting can lead to estimates of either the SATE or the PATE, depending on the relevant target of inference. However, when treatment effects vary according to some unobserved characteristic of sub-

Figure 4.1: Implications of Treatment Effect Heterogeneity for Generalizability from Convenience Samples



jects (that also correlates with the probability of participation in a convenience sample), then no amount of post-stratification will license the generalization of convenience sample results to other sites.

4.2 Results I: Replications of 12 Survey Experiments

The approach I adopt in this chapter is to replicate surveys originally conducted on probability samples on Amazon's Mechanical Turk. The experiments selected for replication came in two batches. The first set of five (Haider-Markel and Joslyn 2001; Peffley and Hurwitz 2007; Transue 2007; Chong and Druckman 2010; Nicholson 2012) were selected for three reasons. First, as evidenced by their placement in top journals, these studies addressed some of the most pertinent political science questions. Second, these studies all had replication data available, either posted online, available on request, or completely described in summaries published in the original article. Finally, they were all conducted on probability samples drawn from some well-defined population. As shown in Table 4.1, the target population was not always the U.S. national population. For example, in Haider-Markel and Joslyn (2001) the target of inference is the PATE among adult Kansans in 1999.

The second set of seven replications grew out of a collaboration with Time-Sharing Experiments for the Social Sciences (TESS), an NSF-supported organization that funds online survey experiments conducted on a national probability sample administered by GfK. These studies are of high quality due in part to the peer-review of study designs prior to data collection and to the TESS data-transparency procedures, by which raw data are posted one year after study completion. I selected seven studies, four of which (Brader 2005; Nicholson 2012; McGinty, Webster and Barry 2013; Craig and Richeson 2014) I replicated on Mechanical Turk, and three of which (Hiscox 2006; Hopkins and Mummolo 2015; Levendusky and Malhotra 2015) I replicated both on Mechanical Turk and TESS/GfK. By and large, the replications were conducted with substantially larger samples than the original studies. All replications were conducted between January and September 2015.

The studies cover a wide range of issue areas – gun control, immigration, the death penalty, the Patriot Act, home foreclosures, mental illness, free trade, health care, and

Table 4.1: Study Manifest

			Replications N	
Citation	Sampling Frame	Original N	MTurk	TESS/GfK
Haider-Markel and Joslyn (2001)	Kansans in 1999 (RDD)	518	1,010	
Brader (2005)	Americans in 2003 (TESS/KN)	354	2,138	
Peffley and Hurwitz (2007)	Black and White Americans in 2000-2001 (SRC)	1,182	1,176	
Transue (2007)	White Americans in the Twin Cities area in 1998 (MMIS)	343	366	
Chong and Druckman (2010)	Americans in 2009 (Bovitz)	1,302	1,820	
Nicholson (2012)	Americans in 2008 (YouGov)	1,096	1,503	
McGinty et al. (2013)	Americans in 2012 (TESS/GfK)	2,952	2,985	
Craig and Richeson (2014)	Americans in 2012 (TESS/GfK)	620	847	
Johnston and Ballard (2014)	Americans in 2013 (TESS/GfK)	2,394	2,985	
Hiscox (2006)	Americans in 2003 (TESS/CSR)	1,578	2,972	2,084
Hopkins and Mummolo (2015)	Americans in 2011 (TESS/GfK)	3,269	2,972	3,180
Levendusky and Malhotra (2015)	Americans in 2012 (TESS/GfK)	1,587	2,972	2,115

polarization – and generally employ framing, priming, or information treatments designed to persuade subjects to change their political attitudes. The particulars of each study's treatments and outcome measures are detailed in Appendix A. In most cases, the studies employ multiple treatment arms and consider effects on a single dependent variable. In some cases, however, a single treatment versus control comparison is considered with respect to a range of dependent variables. All replications followed the original protocols with respect to question wording and number of items. In some cases, not all dependent variables were reported in the original publications; in such cases, I used the dependent variables as reported. In other cases, (e.g., Brader (2005)), many dependent variables were measured; I used my best judgment to decide which were the "main" outcome variables. I acknowledge that these choices introduce some subjectivity into the analysis. Mullinix et al. (2015) address this problem by focusing the analysis on the "first" dependent variable in each study, as determined by the temporal ordering of the dependent variables in the original TESS protocol. Their approach has the advantage of being automatically applicable across all studies but is no less subjective.

A wide range of analysis strategies were used in the original publications, including difference-in-means, difference-in-differences, ordinary least squares with covariate adjustment, subgroup analysis, and mediation analysis. To keep the analyses comparable, I reanalyzed all the original experiments using difference-in-means without conditioning on subgroups or adjusting for background characteristics.[4] Survey weights were incorporated where available. In all cases, I employed the Neyman variance estimator, equivalent to a standard variant of heteroskedasticity-robust standard errors (Samii and Aronow 2012). I used the identical specifications across each version of a study.

The study-by-study results are presented in Figure 4.2. On the horizontal axis of each facet, I have plotted the point estimate of the ATE with 95% confidence intervals. The scale of the horizontal axis is different for each study, for easy inspection of the within-study correspondence across samples. On the vertical axes, I have plotted each treatment

[4]In the case of factorial designs, I used OLS to obtain a single set of coefficients for each factor, essentially averaging over the other margins. In one case, I adjusted for blocks to account for the randomization scheme.

versus control comparison, by study version. The top nine facets compare two versions of each study (original and MTurk), while the bottom three compare three versions (original, MTurk, and TESS/GfK). The coefficients are ordered by the magnitude of the original study effects from most negative to most positive. This plot gives a first indication of the overall success of the replications: in no cases do the replications directly contradict one another, though there is some variation in the magnitudes of the estimated effects.

All together, I estimated 47 original-MTurk pairs of coefficients. Of the 28 coefficients that were originally significant, 19 were significant in the MTurk replications, all with the correct sign. Of the 19 coefficients that were not originally significant, 15 were not significant in the MTurk replications either, for an overall replication rate (narrowly defined) of $(19 + 15)/(28 + 19) = 72\%$. In zero cases did two versions of the same study return statistically significant coefficients with opposite signs. The match rate of the sign and statistical significance of coefficients across studies is a crude measure of correspondence, as it conflates the power of the studies with their correspondence. For example, if all pairs of studies had a power of 0.05, i.e., they only had a 5% chance of correctly rejecting a false null hypothesis of no average effect, then the match rate across pairs of studies would be very high because nearly all coefficients would be deemed statistically insignificant.

A better measure of the "replication rate" is the correlation of the standardized coefficients. Rather than relying on the artificial distinction between significant and insignificant, the correlation is a straightforward summary of the extent to which larger effects in the original studies are associated with larger effects in the replications. One obstacle is that in the presence of sampling error, correlation coefficients are biased towards zero. In an attempt to combat this, I disattentuate the correlations using the procedure described in Chapter 1. In the case of these 12 pairs of studies (47 coefficients), the raw correlation between the MTurk and original coefficients is 0.81. Corrected, this correlation increases to 0.84. This correlation can be visually inspected in Figure 4.3, which plots each coefficient with 95% confidence intervals for both the original and MTurk versions.

For the three studies replicated in parallel on MTurk and on fresh TESS/GfK samples (Hiscox 2006; Hopkins and Mummolo 2015; Levendusky and Malhotra 2015), the replica-

Figure 4.2: Original and Replication Results of 12 Studies

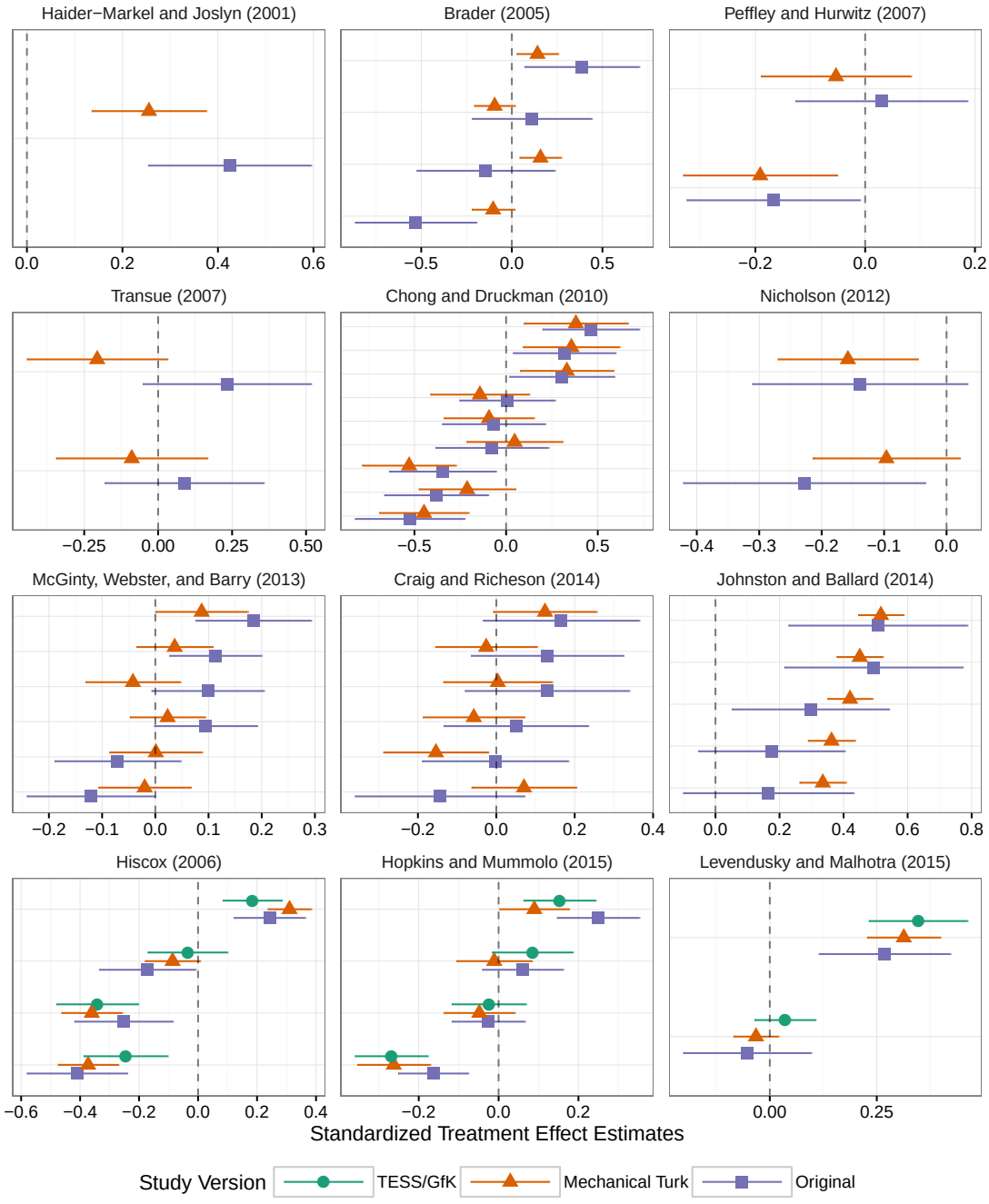
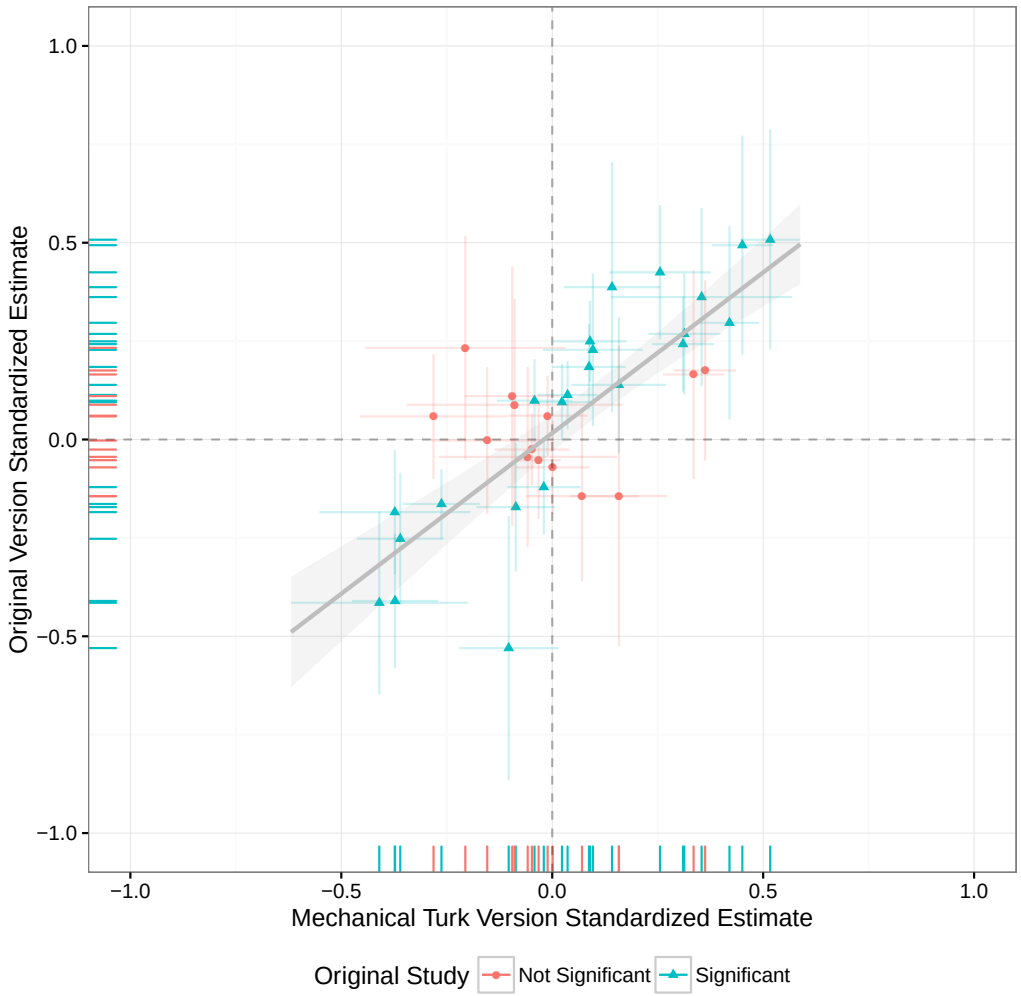


Figure 4.3: Original Coefficient Estimates Versus Estimates Obtained on Mechanical Turk



tion picture is even rosier. The raw correlation of the MTurk and original estimates is 0.94; TESS/GfK estimates with the MTurk estimates, 0.96; TESS/GfK with the original estimates, 0.90. Corrected, all these correlations exceed 1.00.

All in all, these results show a strong pattern of replication across samples, lending credence to the idea that at least for the sorts of experiments studied here, estimates of causal effects obtained on Mechanical Turk samples in 2015 generalize to the national population in 2015.

4.3 Results II: Testing the Null of Treatment Effect Homogeneity

What can explain the strong degree of correspondence between the Mechanical Turk and probability sample estimates of average causal effects? It could be that effects are heterogeneous within each sample – but this heterogeneity works out in such a way that the average effects across samples are approximately equal. This explanation is not out of the realm of possibility, and future work should consider whether the conditional average treatment effects (CATEs) estimated within well-defined demographic subgroups (e.g., race, gender, and partisanship) also correspond across samples.

In this section, I will report the results of empirical tests of a second theoretical explanation for the generalizabilty of these results across research sites: the treatments explored in this series of experiments have constant effects. If effects are homogeneous across subjects, then the differences in estimates obtained from different samples would be entirely due to sampling variability.

Building on advances in Fisher permutation tests, Ding et al. (2015) propose a test of treatment effect heterogeneity that compares treated and control outcomes with a modified KolmogorovSmirnov (KS) statistic. The traditional KS statistic is the maximum observed difference between two cumulative distribution functions (CDFs), and is useful for summarizing the overall difference between two distributions. The modified KS statistic compares the CDFs of the de-meaned treated and control outcomes, thereby removing the estimated average treatment effect from the difference between the distributions and increasing the

sensitivity of the test statistic to treatment effect heterogeneity.

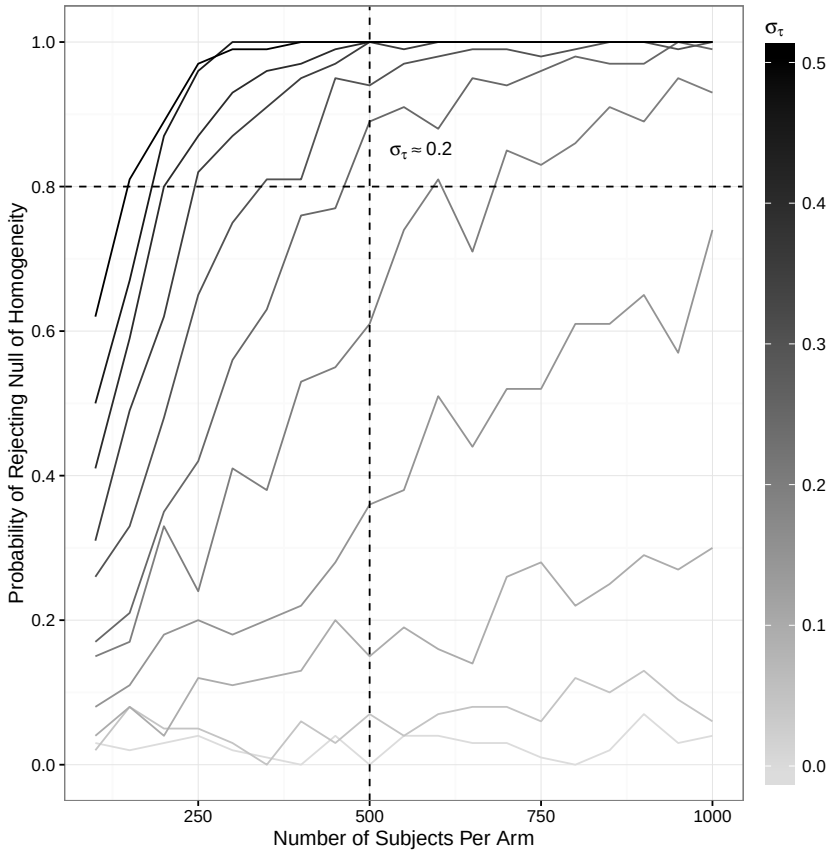
The permutation test compares the observed values of the modified KS statistic to a simulated null distribution. This null distribution is constructed by imputing the missing potential outcomes for each subject under the null of constant effects, then simulating the distribution of the modified KS statistic under a large number of possible random assignments. One difficulty is choosing *which* constant effect to use for imputation. A natural choice is to use the estimated ATE; however, as Ding et al. (2015) show, this approach may lead to incorrect inferences. Instead, they advocate repeating the test for all plausible values of the constant ATE and reporting the maximum p-value. In practice, set of "all plausible" values of the ATE is approximated by the 99.99% confidence interval around the estimated ATE.

This test, like other tests for treatment effect heterogeneity (Gerber and Green 2012, pp. 293-294), can be low-powered. To gauge the probability of correctly rejecting a false null hypothesis, I conducted a small simulation study that varied two parameters: the number of subjects per treatment arm and the degree of treatment effect heterogeneity. Equation 4.1 shows the potential outcomes function for subject i, where Z_i is the treatment indicator and can take values of 0 or 1, σ_τ is the standard deviation of the treatment effects, and X_i is an idiosyncratic characteristic, drawn from a standard normal distribution. The larger σ_τ, the larger the extent of treatment effect heterogeneity, and the more likely the test is to reject the null of constant effects. To put σ_τ in perspective, consider a treatment with enormous effect heterogeneity: large positive effects of 0.5 standard units for half the sample, but large negative effects of 0.5 standard units for the other half. In this case the standard deviation of the treatment effects would be equal to 0.5. While the simple model of heterogeneity used for this simulation study does not necessarily reflect how the test would perform in other scenarios, it provides a first approximation of the sorts of heterogeneity typically envisioned by social scientists.

$$Y_i = 0 \cdot Z_i + \sigma_\tau \cdot X_i \cdot Z_i + X_i \tag{4.1}$$

The results of the simulation study are presented in Figure 4.4. The MTurk versions of

Figure 4.4: Simulation Study: Power of the Randomization Test



the experiments studied here typically employ 500 subjects per treatment arm, suggesting that we would be well powered (power ≈ 0.8) to detect treatment effect heterogeneity on the scale of 0.2 standard deviations, and moderately powered for 0.1 standard deviations (power ≈ 0.6).

The main results of the heterogeneity test applied to the present set of 24 studies are displayed in Table 4.2. Among the original 12 studies, just 1 of the 55 treatment versus control comparisons revealed evidence of effect heterogeneity.[5] Among the Mechanical Turk replications, 7 of 55 treatments were shown to have heterogeneous effects. On the TESS/GfK replications, a similar proportion (2 of 15) tests were significant. In only one case (Hiscox 2006), did the same treatment versus control comparisons return a significant test statistic across samples. In order to guard against drawing false conclusions due to conducting many tests, I present the number of significant tests after correcting the p-values by the Holm correction (Holm 1979) in the last column of Table 4.2.[6]

Table 4.2: Tests of Treatment Effect Heterogeneity

Study Site	N Comparisons	N Significant	N Significant (Holm)
Original	55	1	0
Mechanical Turk	55	7	5
TESS/GfK	15	2	2

Far more often than not, we fail to reject the null of treatment effect homogeneity. Because we are relatively well-powered to detect politically-meaningful differences in treatment response, I conclude from this evidence that the treatment-effect-homogeneity explanation

[5]The number of comparisons increases from 47 to 55 because this test cannot easily accommodate averaging over the margins of a factorial design. In the case of a 2x2 factorial design, I would have estimated two coefficients above but conducted three tests here.

[6]The Holm correction controls the family-wise error rate under the same assumptions as the more familiar Bonferroni correction, but is strictly more powerful (Holm 1979). To implement the correlation, order the m uncorrected p-values within a "family" from smallest to largest. Identify the smallest p-value for which the following condition holds: $p_k \leq \frac{k}{m}\alpha$, where k indexes the p-values, and α is the target significance level. The smallest p-value that meets this condition is insignificant, as are all larger p-values. All smaller p-values are significant.

of the correspondence across experimental sites is plausible.

4.4 Discussion

Levitt and List (2007, p. 170) remind us that "Theory is the tool that permits us to take results from one environment to predict in another[.]" When the precise nature of treatments varies across sites, we need theory to distinguish the meaningful differences from the cosmetic ones. When the contexts differ across sites – public versus private interactions, field versus laboratory observations – theory is required to generalize from one to the other. When outcomes are measured differently, we rely on theory to predict how a causal process will express itself across sites.

In the results presented above, I have not made theoretical claims about the differences in treatments, contexts, or outcome measures across sites – they were all held constant by design. The only remaining impediment to the generalizability of the survey experimental findings from convenience samples to probability samples is the composition of the subject pools. If treatments have heterogeneous effects, researchers have to be careful not to generalize from a sample that has one distribution of treatment effects to populations that have different distributions of effects.

Before this replication exercise (and others like it such as Mullinix et al. (2015) and Krupnikov and Levine (2014)), social scientists had a limited empirical basis on which to develop theories of treatment effect heterogeneity for the sorts of treatments explored here, which by and large attempt to persuade subjects of policy positions. Within political science, some theories predict heterogeneity by partisanship, political knowledge, and need for cognition. Others, such as the theory of Bayesian updating offered in Chapter 2, predict limited heterogeneity in persuasive treatment effects. Both within and across samples, the treatments studied here have exhibited muted treatment effect heterogeneity. Whatever differences (measured and unmeasured) there may be between the Mechanical Turk population and the population at large, they do not appear to interact with the treatments employed in these experiments in politically meaningful ways. In my view, it is this lack of heterogeneity that explains the overall correspondence across samples.

I echo the concerns of Mullinix et al. (2015), who caution that the pattern of strong probability/convenience sample correspondence does not imply that all survey experiments can be conducted with opt-in Internet samples with no threats to inference more broadly. Indeed, the crucial question is the extent of treatment effect heterogeneity. The weaker correspondence findings of Krupnikov and Levine (2014) can be reconciled with those reported here by noting that those authors had strong expectations of treatment effect heterogeneity for some of their tests. Additionally, their experiments were conducted on relatively small samples, implying that the observed correlation between MTurk and the national sample treatment effect estimates (0.17) may be especially attenuated by measurement error.

Finally, it is worth reflecting on the remarkable robustness of the experiments replicated here. Concerns over p-hacking (Simonsohn, Nelson and Simmons 2014), fishing (Humphreys et al. 2013), data snooping (White 2000), and publication bias (Franco, Malhotra and Simonovits 2014; Gerber, Malhotra, Dowling and Doherty 2010) have lead many to express a great deal of skepticism over the reliability of results published in the scientific record (Ioannidis 2005). An effort to replicate 100 papers in psychology (Open Science Collaboration 2015) was largely viewed as a failure, with only one-third to one-half of papers replicating, depending on the measure. The raw Pearson correlation of original and replication was 0.51. By contrast, the empirical results in this set of experiments were largely confirmed. The success of these replications could be due to a number of factors. First, a condition for inclusion in my set of replications was the availability of the datasets (or sufficient summary statistics). It is possible that the availability of data post-publication is a marker for study or research quality, which may be correlated with replicability. Second, the TESS studies replicated here all went through pre-implementation peer review, which may have increased quality on the front end. Finally, because all these studies have a similar design, maintaining fidelity to the original protocols may have been easier than it was for the 100 psychology replications.

5 The Persistence of Survey Experimental Treatment Effects

Do treatments deployed in survey experiments have persistent effects? The prototypical survey experiment collects background information, delivers experimental treatments, and records outcomes all in the space of a few minutes. Survey experimental treatments may cause real, underlying attitude change, or they may simply induce a temporary increase in subjects' probability of choosing one response over another. Further, some treatments may have more persistent effects than others. I hypothesize that priming and framing treatments that make certain considerations more accessible have fleeting effects whereas treatments that make considerations newly applicable or provide new information persist longer. Results from eighteen panel survey experiments show that on average, survey experimental treatment effects persist after 10 days, albeit at approximately half their original magnitudes. This suite of evidence also provides strong support for the theory that information treatments have more persistent effects than framing treatments.

The previous chapter examined the extent to which survey experiments shed light on real-world political discourse, focusing on the generalizability or "external validity" of survey experimental investigations. This chapter addresses the separate, but related, issue of treatment effect persistence. A common criticism of survey experiments is that they uncover real but fleeting effects. If so, the effects studied in survey experiments may be little more than laboratory curiosities. In the words of Gaines, Kuklinski and Quirk (2007, p. 6) "The implications of survey-experimental results for politics depend crucially on how long the effects last, with relevant periods measured in weeks, or months, not minutes."

Why might survey experimental treatment effects fade? One hypothesis holds that large initial effects are due to experimenter demand: subjects answer how they think researchers want immediately after treatment, but do not face such pressure after some time has passed. Another view is that subjects may simply forget the information provided to them in a persuasive treatment. If it is the information itself that is responsible for short-term shifts in attitudes, the effects should disipate as the memory of the treatments becomes more distant. Framing treatments in particular are hypothesized to have fleeting effects because subjects may naturally encounter counterframes in the days and weeks after treatment.

Understanding the persistence of persuasive effects is especially important in an era of long electoral campaigns. If the effects of even very persuasive messages evaporate within a few hours, then campaigns would be best served by not spending large amounts of money on early advertisements, only to see those effects disappear by election day. If, however, persuasive information can induce durable attitude change, the cumulative effects of two years' campaigning might have a substantial impact on election results.

Previous work has found mixed evidence on persistence. Those who find little or no persistence include de Vreese (2004), who showed that subjects exposed to "strategic" news about the enlargement of the EU reported higher (0.44 scale points, 5-point scale) political cynicism than subjects who saw news focused on the issues surrounding enlargement. This difference was almost non-existent in a follow up conducted one week later (0.02 scale points). Similarly, Druckman and Nelson (2003) find that a special-interests frame increases

support for a bill by 0.71 scale points (7-point scale); 10 days later this effect drops[1] to 0.41. Mutz and Reeves (2005) report that videos featuring uncivil discourse decreased political trust by 0.44 standard deviations relative to civil discourse videos but that the difference (not reported) was no longer statistically significant approximately one month after treatment.

Some survey experiments have found strong evidence of persistence: Tewksbury, Jones, Peske, Raymond and Vig (2000) report that exposure to a pro-regulation news story increased support for the regulation of hog farms by 24.1 percentage points relative to an anti-regulation story; this difference remained at 25.2 percentage points three weeks later. Lecheler and de Vreese (2011) show that a positive economic-benefit frame increased support for the inclusion of Bulgaria and Romania in the EU by 1.1 scale points (7-point scale); this effect was 0.93 a day later, 1.35 a week later, and 0.81 after two weeks.

These divergent results have led scholars to pose more nuanced theoretical questions. Instead of asking if survey experimental treatment effects persist in general, they ask, "Which treatments exhibit stronger or weaker persistence?" Drawing on theories of framing, priming, and persuasion, Baden and Lecheler (2012) hypothesize that three features of an experimental treatment in particular will influence its durability. Some primes and cues operate by making subjects' preexisting knowledge more *accessible*. Treatments that operate primarily though this channel are hypothesized to have fleeting effects. Other treatments may operate by making preexisting knowledge newly *applicable*. For example, drawing a connection between a policy proposal and a social movement might induce subjects to consider how the movement's principles are applicable in the policy context. Such treatments may persist longer. Finally, some treatments supply subjects with *new information*, possibly leading to enduring belief change. Most treatments are likely to be a bundle of all three processes – the idiosyncratic mix in each treatment may influence durability.

The Bayesian theory of persuasion elaborated in earlier chapters predicts that persuasive

[1] Approximately 50% of the sample completed the follow-up, leading to a loss of power to detect treatment effects. The follow-up estimate is not significantly different from zero – or from the immediate estimate of 0.71.

effects will last longer, the smaller the variance of subjects' posterior views. That is, changes in attitudes will last longer, when subjects are more certain of their "new" attitudes. In the Bayesian framework, the process by which treatment effects decay is traced to subjects organically encountering other relevant information in the intervening period between exposure to a persuasive treatment and subsequent waves of measurement. Treatments that make specific, well-reasoned persuasive arguments are predicted to decrease posterior variance and therefore should last longer than treatments that are more oblique. Additionally, this theory predicts individual differences in the persistence of effects. Individuals' posterior variances are functions of their prior variances and their likelihood functions. Those with small prior variances are harder to convince, but once persuaded, are predicted to stay convinced. Those with more discriminating likelihood functions are predicted to have smaller posterior variances as well. The Bayesian predictions for persistence are not at odds with Baden and Lecheler (2012). Indeed, persuasive information should be posterior variance-reducing, whereas the applicability and accessibility mechanisms are unlikely to affect the certainty with which subjects hold their attitudes.

This framework underlines one major complication in the study of persistence: stronger arguments may be associated both with larger initial effects *and* more enduring effects. All else being equal, the stronger the persuasive argument, the larger the initial attitude change and the smaller the posterior variance, suggesting a positive correlation between the size of effects and their durability. However, a far less elaborate theory of persistence could account for why some effects endure and others fade: big effects last and small effects do not. In order distinguish between a Bayesian theory of persistence and this simpler argument, a body of evidence with large effects generated by both strong and weak arguments is required.

The remainder of this chapter will proceed as follows. First, I will describe the theoretical framework alluded to above in greater detail. I will then describe how the set of treatments used to explore durability were selected, the panel experimental design, and the statistical model of persistence used here. I will then present results and interpret them in light of the theoretical framework.

5.1 Theory

The expectancy value model of attitudes (Ajzen and Fishbein 1980; Nelson, Oxley and Clawson 1997; Zaller 1992; Chong and Druckman 2012) posits that subjects have a set of considerations to which subjective weights are attached. When prompted by a survey question, subjects take a weighted average of their considerations and return a response. This framework can accommodate many of the types of treatments that are commonly studied in survey experiments. Emphasis frames (Druckman 2001) are hypothesized to affect responses by changing the weights, not the considerations themselves. Information treatments, by contrast, change responses by furnishing subjects with new considerations to take account of when offering a survey response. The manner in which survey responses vary with framing or information treatments can therefore be modeled as:

$$Y(Z) \sim f(\mathbf{W}, \mathbf{C}, Z, \sigma_Y, \ldots),$$

where Y is the survey response, $\mathbf{W}$ is a vector of weights attached to a vector of considerations $\mathbf{C}$, Z is an indictor for the treatment condition, and σ_Y is a measure of prior uncertainty that subjects have about the issue area. This description of how responses might change is very general and includes "..." to indicate that many other factors influence attitudes. This model allows for complex interactions between Z, $\mathbf{W}$, and $\mathbf{C}$. Unfortunately the flexibility of this model renders it difficult to work with when making predictions about the relative durability of different treatments.

I will turn instead to a much simplied version of the general model above:

$$Y(Z) = \sum_{1}^{j} w_j(Z) \cdot c_j,$$

where Y is a simple weighted average of considerations. The weights w_j are themselves responsive to Z, indicating that the weights attached to considerations might differ depending on the treatment condition. In this model, there is no requirement that the weights w_j sum to 1. This choice reflects the theoretical possibility that the treatments themselves change the amount of "attention" that a subject gives to a survey response. If a respondent has been asked to carefully weigh a series of arguments for and against a proposition, the total

amount of attention (weight) given to that response may be higher than if the subject were solicited for a response without any additional context.

To illustrate how this model helps draw distinctions among treatments, suppose that all subjects' survey responses are a function of just three considerations, c_A, c_B, and c_C, and that the treatment indication Z can take three values: control, an emphasis frame or an information treatment.

$$\text{Response under control: } Y(Z=0) = 0.5 \cdot c_A + 0.2 \cdot c_B$$

$$\text{Response under emphasis frame: } Y(Z=1) = 0.8 \cdot c_A + 0.2 \cdot c_B$$

$$\text{Response under information treatment: } Y(Z=2) = 0.5 \cdot c_A + 0.2 \cdot c_B + 0.3 \cdot c_C$$

The baseline (control) responses are a weighted average of considerations c_A and c_B, where the weights are 0.5 and 0.2. The responses under an emphasis are also a weighted average of these two considerations, but c_A is given a new weight, 0.8. Under an information treatment, the weights on c_A and c_B remain unchanged, but a new consideration, c_C, with an associated weight of 0.3 is added.

With this notation in hand, we can define two estimands, the average treatment effect framing and the average treatment effect of information. In this context, the expectation operator $\mathbb{E}[\cdot]$ denotes the true sample mean.

$$\begin{aligned}\text{Average Framing Effect} &= \mathbb{E}[Y(Z=1) - Y(Z=0)] \\ &= (0.8 * c_A + 0.2 * c_B) - (0.5 * c_A + 0.2 * c_B) \\ &= 0.3 * c_A\end{aligned}$$

$$\begin{aligned}\text{Average Information Effect} &= \mathbb{E}[Y(Z=2) - Y(Z=0)] \\ &= (0.5 * c_A + 0.2 * c_B + 0.3 * c_C) - (0.5 * c_A + 0.2 * c_B) \quad = 0.3 * c_C\end{aligned}$$

The question under study here is, to what extent do these effects endure? In order to answer this question, we must add a time dimension to our model. An attitude measured at time t can be written:

$$Y_t(Z) = \sum_{1}^{j} w_{j,t}(Z) \cdot c_j$$

In this model, the weights are subscripted by consideration j and time t, and may be different depending on treatment status (Z). Using this setup, I define the average persistence of an effect as the ratio of the ATE at time 2 to the ATE at time 1:

$$\frac{\text{ATE Time 2}}{\text{ATE Time 1}} = \frac{\mathbb{E}[Y_{t=2}(Z=1) - Y_{t=2}(Z=0)]}{\mathbb{E}[Y_{t=1}(Z=1)Y_{t=1}(Z=0)]}$$

It should be emphasized that while this estimand is well defined, it is not equal to the average of all individual-level ratios of time 1 effect to the time 2 effect because the expectation of a ratio is not, in general, equal to the ratio of expectations. This non-equality is due to the possible covariance between individuals' treatment effects at time 1 with their treatment effect at time 2. I recognize that this definition introduces a theoretical inconsistency: the psychological processes hypothesized to be responsible for persistence take place at the individual level but the persistence estimand is defined at the group level. Ultimately, this inconsistency is unavoidable because of the fundamental problem of causal inference: we can no more identify the covariance between effects at the individual level than we can identify the effects themselves.

The average persistence estimand would be misleading in some scenarios. Imagine, for example, that half the sample has a treatment effect of 1 at time 1 but 0 at time 2; the other half has treatment effects of 0 at time 1 but 1 at time 2. For individuals in the first group, the true persistence is $\frac{0}{1} = 0$; for individuals in the second group, the true persistence is undefined: $\frac{1}{0}$. The average of these individual level persistences is likewise undefined. The average persistence, however, is $\frac{0.5}{0.5} = 1$. By contrast, if treatment effects were equal to 1 for the whole sample at time 1 and equal to 0.5 at time 2, then the average of each individual's effect ratio would be equal to the ratio of the average effects: $\frac{1}{N}\sum_1^N \frac{0.5}{1} = \frac{1}{N}\sum_1^N 0.5 / \frac{1}{N}\sum_1^N 1 = 0.5$.

With this caveat in mind, I define the following persistence estimands:

$$\begin{aligned}\text{Framing Effect Persistence} &= \frac{\mathbb{E}[Y_{t=2}(Z=1) - Y_{t=2}(Z=0)]}{\mathbb{E}[Y_{t=1}(Z=1) - Y_{t=1}(Z=0)]} \\ &= \frac{(0.6 * c_A + 0.2 * c_B) - (0.5 * c_A + 0.2 * c_B)}{(0.8 * c_A + 0.2 * c_B) - (0.5 * c_A + 0.2 * c_B)} \\ &= \frac{0.1 * c_A}{0.3 * c_A} \\ &= \frac{1}{3}\end{aligned}$$

$$\begin{aligned}\text{Information Effect Persistence} &= \frac{\mathbb{E}[Y_{t=2}(Z=2) - Y_{t=2}(Z=0)]}{\mathbb{E}[Y_{t=1}(Z=2) - Y_{t=1}(Z=0)]} \\ &= \frac{(0.5 * c_A + 0.2 * c_B + 0.2 * c_C) - (0.5 * c_A + 0.2 * c_B)}{(0.5 * c_A + 0.2 * c_B + 0.3 * c_C) - (0.5 * c_A + 0.5 * c_B)} \\ &= \frac{0.3 * c_C}{0.5 * c_C} \\ &= \frac{2}{3}\end{aligned}$$

It is no accident that in the foregoing toy example, the information effect showed stronger persistence than the framing effect. The assumption that changes in weights are less durable than the addition of new considerations is central to the theory of treatment effect persistence employed here.

This model can accommodate the theoretical predictions of treatment effect persistence put forth by Baden and Lecheler (2012). *Accessibility* treatments are hypothesized to operate primarily by increasing the weight given to a particular consideration. Because such treatments only affect outcomes through the weighting scheme, they are hypothesized to be fleeting. As an example of an accessibility treatment, consider the study described in Transue (2007), replicated as Study 5 below. In that experiment, subjects were randomly exposed to a treatment question that asked subjects: "How close do you feel to [your ethnic or racial group]/[other Americans]?" The dependent variable was subjects' attitudes toward raising taxes for education; the treatment operates by increasing the weight given to ethnic identity or superordinate identity when answering this question.

Applicability frames also operate by changing the weights given to considerations, but do so in a manner distinct from accessibility frames. In this model, subjects hold a set of considerations in mind but the associated weights are attitude-specific. That is, consideration c_A may be given one weight when a subject is offering an attitude on gun control but a different weight when responding to a question on national security. An applicability frame works by linking the two attitudes, so that the considerations in common are given greater weight. As an example of an applicability treatment, consider the frames employed in Levendusky and Malhotra (2015), replicated below as Studies 12 and 13. Subjects are exposed to news stories depicting the electorate as being either moderate or polarized. The dependent variable is subjects' attitudes toward four policies. The ancillary attitude is the value subjects place on moderation in politics; the relevant considerations are given more weight when responding to the policy questions.

Information treatments operate by adding new considerations. These treatments are sometimes referred to as *belief change* treatments, a term I avoid so as to maintain a clear distinction between the cause (information) and the effect (attitude change), although in some instances, information treatments may persuade by changing beliefs. The treatments in the capital punishment study explored in chapter 2 (Study 1 below) are examples of information treatments: subjects were exposed to research reports on the efficacy of capital punishment before registering their support or opposition to the practice.

Finally, I have cast framing treatments as operating on the weights and information treatments as introducing new considerations, but there is no reason to imagine that a given framing treatment could never introduce some new consideration or that an information treatment does not influence the weights given to pre-existing considerations. Nevertheless, the distinction is important to maintain in order to highlight the processes by which effects may persist or decay.

This framework generates persistence predictions for the set of studies presented below.[2]

[2]The predictions for the 10 TESS replication studies (Studies 8 - 17) were pre-registered and posted at `http://egap.org/design-registration/registered-designs/`. Study 3 did not employ a panel design and is not included in this chapter.

The descriptions below are limited to the bare-bones justifications for the associated predictions. See Appendix A for full descriptions of study design, including treatment stimuli, number of subjects, and wording of dependent variables. Finally, scholars may reasonably disagree about which set of mechanisms is primarily responsible for treatment effects, so I will make the case for my own categorization while recognizing that others are possible as well.

- *Study 1: Capital Punishment*
  Treatments: Pro, con, or null social scientific evidence about the efficacy of capital punishment in deterring crime.
  Mechanisms: *Information.* The evidence was presented in graphs with accompanying explanatory text.
  Prediction: Strong persistence.

- *Study 2: Minimum Wage*
  Treatments: Short web videos, two in favor of raising the minimum wave, two against it, and two placebos.
  Mechanisms: *Information.* The non-placebo treatments offered very detailed information about the economics and politics of raising the minimum wage.
  Prediction: Strong persistence.

- *Study 4: Gun Control*
  Original study: Haider-Markel and Joslyn (2001)
  Treatments: Gun control question framing in terms of public safety or individual rights.
  Mechanisms: *Accessibility.* Safety and rights considerations are triggered by the frame. *Applicability.* For some subjects, these considerations may be newly applicable to the gun control debate. *Information* The questions also contained a small dose information about a concealed carry gun law.
  Prediction: Moderate persistence. With hindsight, I would change this prediction to weak persistence, since the applicability and information channels are so thin. I nevertheless group this study with the "moderate" predictions in the analyses that follow.

- *Study 5: Superordinate Identity*
  Original study: Transue (2007)
  Treatments: Primes that activate either particularistic (racial/ethnic) or superordinate (national) identity.
  Mechanisms: *Accessibility* The treatment makes one identity more salient or accessible, but does not create new associative links or offer new information.
  Prediction: Weak persistence.

- *Study 6: Patriot Act*
  Original study: Chong and Druckman (2010)
  Treatment: Pro or con information about the Patriot Act.
  Mechanisms: *Information.* Subjects are directly given facts about the Patriot Act. *Accessibility.* The information framed to activate some pre-existing considerations such as individual freedom. *Applicability.* The information may also create links between considerations such as surveillance and national security.
  Prediction: Strong persistence.
  Of the replications, this was the only original study to measure outcomes at multiple points in time – the strong persistence prediction, therefore, is made with the benefit of knowing beforehand that the effects did indeed persist for 10 days originally.

- *Study 7: Elite Endorsements*
  Original study: Nicholson (2012)
  Treatments: Policy proposals with partisan endorsement cues.
  Mechanisms: *Accessibility.* These cues may make considerations of partisan loyalty more accessible but do not apply, for example, Democratic party values to new issue areas. *Information.* This treatment may also operate by giving subjects the novel information that someone they respect (or despise) endorses a proposal. This sort of information is qualitatively different from information that, for example, informs subjects of the policy's expected benefits.
  Prediction: Weak persistence.

- *Studies 8 and 9: Free Trade A and B*
  Original study: Hiscox (2006)
  Treatments: Valence frames, which describe why some Americans do or do not support free trade, and an expert frame, which states that professional economists are united in their support for free trade.
  Mechanisms: *Accessibility.* The valence frames prompt subjects to give negative or positive economic considerations extra weight. *Information* Subjects learn that professionals in this complicated area appear united on this question.
  Prediction: Moderate persistence.

- *Studies 10 and 11: Frame Breadth A and B*
  Original study: (Hopkins and Mummolo 2015)
  Treatments: Pro-conservative statements by U.S. senators.
  Mechanisms: *Accessibility.* These treatments operate by making terrorism, crime, health care costs, or wasteful spending more salient when answering budget questions, though the senators do provide information in their arguments.
  Prediction: Weak persistence. In hindsight (but also with the benefit of having seen the results) I would now predict moderate persistence for these treatments in view of their information content, though a weak prediction was pre-registered. In the analysis below, the pre-registered predictions will be used to evaluate the theoretical framework.

- *Studies 12 and 13: Polarization A and B*
  Original study: Levendusky and Malhotra (2015)
  Treatments: Articles depicting politics as polarized or moderate.
  Mechanisms: *Applicability.* If subjects place a value on being "moderate" or "reasonable" members of the electorate, and this value is made more salient via the polarized treatment, then this consideration will be made newly applicable to the policy questions.
  Prediction: Moderate persistence.

- *Study 14: Immigration*
  Original study: Brader (2005)
  Treatments: Positive or negative newspaper stories on immigration featuring either Latino or European immigrants.
  Mechanisms: *Accessibility.* The articles are intended to heighten racial considerations when providing immigration policy preferences. *Information.* The accompanying text also includes some information about, variously, the positive or negative economic effects of immigration.
  Prediction: Moderate persistence.

- *Study 15: System Threat*
  Original study: (Craig and Richeson 2014)
  Treatments: A frame highlighting the increasing population share of minorities.
  Mechanisms: *Accessibility.* The treatment may trigger respondents' racial anxieties.
  Prediction: Weak persistence.

- *Study 16: Expert Economists*
  Original study: Johnston and Ballard (2014) Treatments: Subjects are provided the consensus view among economists for various policy proposals. Mechanisms: *Information.* Subjects are given information about expert views on these five policies. *Accessibility.* Subjects are more or less told which answer is the "correct" one. If the first mechanism is at work, the treatment ought to have enduring effects. If effects are primarily due to the second mechanism, effects should be fleeting. Overall, I view this treatment as operating through the accessibility channel: the "correct" answer is more accessible at the moment of providing a survey response.
  Prediction: Weak persistence.

- *Study 17: Mental Illness*
  Original study: McGinty et al. (2013)
  Treatments: Newspaper vignettes depicting a mass shooting and providing policy proposals.
  Mechanisms: *Accessiblity.* One treatment (the story depicting the shooting) operates by highlighting considerations about gun violence. *Information.* The second treatment (the policy proposals) operates by providing information about various gun policies.
  Prediction: Moderate persistence.

- *Study 18: Contentious Global Warming*
  Treatments: Graphs that showed a scientific consensus about warming trends or evidence of a warming "hiatus".
  Mechanisms: *Information*. The treatment explicitly furnishes subjects with new information about warming trends.
  Prediction: Strong persistence.

- *Study 19: Newspapers*
  Treatments: Newspaper op-ed pieces on five policy areas.
  Mechanisms: *Information*. The op-eds make extended arguments in favor of a particular policy.
  Prediction: Strong persistence.
  In this experiment, three waves of post-treatment measurement were conducted, so I will provide a more detailed explanation of this study below. I also note that the design of this study was strongly influenced by (Hopkins and Mummolo 2015), replicated here as Studies 10 and 11.

5.2 Research Approach

The 18 studies enumerated above share many common design elements. The majority of the studies (15) were conducted on Amazon's Mechanical Turk, with the remainder (3) conducted on a nationally-representative survey administered by GfK. Four of the studies conducted on Mechanical Turk are original: the capital punishment and minimum wage experiments featured in Chapter 2, the contentious global warming experiment discussed in Chapter 3, and the op-eds experiment explored in chapter 4. The remaining fourteen experiments are replications of other authors' work. Four (Haider-Markel and Joslyn 2001; Transue 2007; Chong and Druckman 2010; Nicholson 2012) were chosen for two main reasons. First, as evidenced by their placement in top political science journals, they are well-crafted and theoretically important. Second, because they employ text-based treatments, they are directly replicable. Many other prominent survey experiments (e.g., Nelson, Clawson and Oxley (1997); Mutz and Reeves (2005); Brader (2005)) present subjects with short videos that, for example, depict political debates or advertisements. The original materials for such experiments are sometimes no longer available and further, may appear out-of-date to contemporary audiences.

Seven studies originally conducted by other authors on the Time-Sharing Experiments

in the Social Sciences (TESS) platorm were also selected for replication. I chose these seven studies because they met the following criteria:

1. The original experiments showed precisely estimated and substantively large immediate average treatment effects. In the absence of so-called "sleeper effects" (Hovland, Lumsdaine and Sheffield 1949; Hovland and Weiss 1951; Cook and Flay 1978), unambiguous immediate effects are required for the study of persistence.
2. The experiments did not employ vignettes or hypothetical candidate scenarios. Asking "Do you prefer Candidate A or Candidate B" after 10 days has no meaning without re-displaying the stimuli, i.e., re-treating subjects.
3. The frames were sufficiently relevant to present-day (2015) politics. Experiments that relied on subjects' presumed familiarity with, for example, the 2008 presidential election, could not be directly replicated.
4. The original authors provided a clear study description or write-up on the TESS website. I did not pursue the few studies lacking this information. In their study of the TESS database, Franco et al. (2014) showed that null results were far less likely to be written up, suggesting that the studies excluded for this reason would probably have been excluded for point 1 had more information been available in any case.

Three of the TESS studies (Hiscox 2006; Hopkins and Mummolo 2015; Levendusky and Malhotra 2015) were replicated in parallel on Mechanical Turk and GfK. The other four TESS replications (Brader 2005; Craig and Richeson 2014; Johnston and Ballard 2014; McGinty et al. 2013) were conducted on Mechanical Turk only.[3]

The experiments all followed a similar panel survey experimental design.[4] In wave 1, subjects answered a battery of demographic questions, were exposed to their randomly-assigned treatments, and then responded to questions whose answers constituted the outcome variables. The average time spent on wave 1 surveys was between 5 and 10 minutes.

Wave 2 occurred approximately 10 days after treatment and consisted only of the dependent variables. The Mturk studies were executed with the assistance of the `MTurkR` package for `R` (Leeper 2015). In an effort to guard against a possible artifact in which subjects

[3]See chapter 4 for an exploration of the extent to which convenience and probability samples yield similar results.

[4]Studies 1 and 2 also included a pre-treatment survey before wave 1.

simply remember how they previously answered questions, in some studies, innocuous features of the question presentation were manipulated. For example, in Studies 8-14, whether follow-up response options were displayed vertically or horizontally was randomly varied. Similarly, in one study (Study 19: Newspapers), half the sample was exposed to an additional op-ed treatment in order to defend against a demand effect in which subjects merely remember what they think the experimenters want to hear. The results of these manipulations (not reported here) are negligible, suggesting that these alternative explanations for persistence are not at work.

5.2.1 Measuring Persistence

When discussing the duration of treatment effects, it is common to dichotomize effects into those that persist and those that do not (Druckman and Nelson 2003; de Vreese 2004; Mutz and Reeves 2005). For example, if a treatment caused a statistically significant shift in outcomes at time 1, but the time 2 estimate is not statistically significant, the treatment effect is said to "not persist." While this categorization may have some heuristic value, the approach taken here is to measure the percentage of the treatment effect measured at time 1 still present at time 2.

In the expressions below, $Y_{T1,i}$ and $Y_{T2,i}$ represent the outcome for subject i at time 1 and 2, respectively. Z_i is subject i's treatment assignment; the average treatment effect (ATE) at time 1 is represented by α_1 and the ATE at time 2 by β_1. The error terms ϵ_i and η_i represent idiosyncratic variation in the outcome variables not accounted for by Z_i.

$$Y_{T1,i} = \alpha_0 + \alpha_1 Z_i + \epsilon_i$$
$$Y_{T2,i} = \beta_0 + \beta_1 Z_i + \eta_i$$

The ratio of β_1 to α_1 is our metric of treatment effect persistence, γ, that is, the percentage of the ATE at time 1 still evident at time 2.

$$\gamma = \frac{\beta_1}{\alpha_1}$$

I estimate α_1 and β_1 separately by an ordinary least squares (OLS) regression of outcomes on the treatment indicator. I estimate γ by calculating the ratio of $\hat{\beta_1}$ to $\hat{\alpha_1}$, and obtain standard errors and confidence intervals via the nonparametric bootstrap. Because γ is a ratio, when $\hat{\alpha_1}$ is very close to zero, persistence estimates may be correspondingly very large, and further, are very unstable, swinging from extremely large positive values to extremely large negative values. To combat this problem, I trim off the 0.5^{th} and 99.5^{th} percentiles of the bootstrap samples before computing standard errors. This operation has a very small influence on the resulting confidence interval estimates, but constrains estimates of standard errors to a reasonable range of values. In practice, these standard errors are slightly larger than those that result from two-stage least squares, a parametric model that corresponds to this setup.

Another complication attending to the estimation of treatment effect persistence is panel attrition. If not all subjects respond to the survey at time 2, then estimates of β_1 (and therefore γ) may be biased. I address this problem with an assumption about the sample distribution of subject types. Subjects have a set of potential outcomes R that may differ depending on their treatment assignment. For example $R_i(Z_i = 1)$ describes the subject i's reporting status when i is in treatment ($Z_i = 1$). If this quantity is equal to 1, i reports in the second round and not otherwise. Subjects who never report in the second round, regardless of treatment status ($R_i(Z_i) = 0$) are called "Never-reporters." Subjects who always report, regardless of treatment status ($R_i(Z_i) = 1$) are called "Always-reporters." Similarly, types such as "If-treated reporters" and "If-untreated reporters" are theoretically possible. These experiments, however, all have more than two treatment arms, so the possible number of subject types quickly becomes unwieldy. In the analyses that follow, subjects are assumed to be either Never-reporters or Always-reporters.[5] The estimates of over-time persistence only pertain to the Always-reporters. For this reason, α_1 (the immediate effect of treatment) will be estimated among Always-reporters only as well.

[5]See Aronow, Coppock, Gerber, Green and Kern (2015) for an alternative estimation strategy for Studies 12 and 13 (Polarization) that makes fewer assumptions.

5.3 Results

In this section, I present summaries of the persistence of effects for all studies. To ease interpretation, all outcome variables were standardized by subtracting off the control group mean and dividing by the control group standard deviation. In Study 7 (Elite Endorsements), the ordinal dependent variable is rescaled to take numeric values between -1 and 1 in order to apply this standardization procedure.

Figure 5.1 plots standardized time 1 treatment effect estimates on the horizontal axis and time 2 estimates on the vertical axis, separately for each study. Points lying on the 45 degree line represent perfect persistence: the time 2 estimate is the same as the time 1 estimate. Points above the 45 degree line indicate that the effect strengthened over time and points below indicate that the treatment effect decayed over time. Figure 5.1 shows a great deal of variation in persistence from study to study. Study 1 (Capital Punishment) shows a very strong degree of persistence: effects that are strongly positive in time 1 are strongly positive in time 2. Study 4 (Gun Control) featured a very robust time 1 estimate; this effect evaporates completely by time 2.

Figure 5.2 plots estimates of γ, the percentage of the time 1 effect still present at time 2 separately for each treatment/dependent variable combination. When the estimate of α, the effect of treatment at time 1, is very small (or imprecisely estimated), estimates of γ become very unstable. In some cases, the estimate is far outside the -1.5 to 1.5 range displayed on the plot. The vertical line represents the precision weighted average of the effects in each study, and the vertical shading displays the 95% confidence interval implied by the standard error[6] of the weighted average (Gerber and Green 2012, p. 356). In all 18 studies, the weighted average of the persistence estimates is positive and the 95% confidence interval does not cross zero in 10 of these studies.

[6]The calculation of this standard error relies on the assumption that the observations are independent of one another; they are not. In the case of multiple treatment arms, effects are all assessed relative to the same control group. In the case of multiple outcome variables, estimates compare the same subjects on multiple dimensions, necessarily introducing dependence among them. Both of these features will produce a positive covariance, implying that the variance estimates presented in the graph are slightly anti-conservative.

Figures 5.1 and 5.2 display the same persistence estimates in two related ways. The points in Figure 5.2 can be thought of as the ratio of the y-axis to the x-axis in Figure 5.1. The slope of the points in Figure 5.1 is a summary of the overall persistence within a study, similar to the weighted average summary represented by the vertical lines in Figure 5.2.

Table 5.1 presents persistence estimates of γ and associated standard errors for each treatment effect. These estimates are identical to the estimates that would be obtained by dividing the y-axis by the x-axis in Figure 5.1. As indicated by the sometimes large standard errors, these estimates can be quite imprecise. A meta-analytic summary of all studies is obtained by taking a weighted average of the estimates, where the weights are the inverse of the squared standard errors (Borenstein, Hedges, Higgins and Rothstein 2009, p. 65). The overall estimate is 44% with a standard error of 2%.

Table 5.1: Average Persistence of 18 Studies

	Study	Average Persistence	SE	Prediction
Study 7	Elite Endorsements	-0.03	0.48	Weak
Study 16	Expert Economists	0.15	0.04	Weak
Study 10	Frame Breadth GfK	0.21	0.14	Weak
Study 15	System Threat	0.33	0.34	Weak
Study 11	Frame Breadth Mturk	0.60	0.15	Weak
Study 5	Superordinate Identity	0.99	1.14	Weak
Study 4	Gun Control	-0.06	0.32	Moderate
Study 8	Free Trade GfK	0.12	0.17	Moderate
Study 13	Polarization Mturk	0.22	0.09	Moderate
Study 12	Polarization GfK	0.28	0.12	Moderate
Study 9	Free Trade Mturk	0.49	0.08	Moderate
Study 14	Immigration	0.62	0.36	Moderate
Study 17	Mental Illness	0.87	0.31	Moderate
Study 19	Newspapers	0.52	0.03	Strong
Study 6	Patriot Act	0.61	0.17	Strong
Study 2	Minimum Wage	0.62	0.06	Strong
Study 18	Contentious Global Warming	0.84	0.44	Strong
Study 1	Capital Punishment	0.93	0.08	Strong
	All Studies	0.44	0.02	

To what extent are the predictions of the theoretical framework presented above borne out? Using the study-specific summaries presented in Table 5.1, I take the average of

Figure 5.1: Standardized Average Treatment Effect Estimates: Time 1 versus Time 2

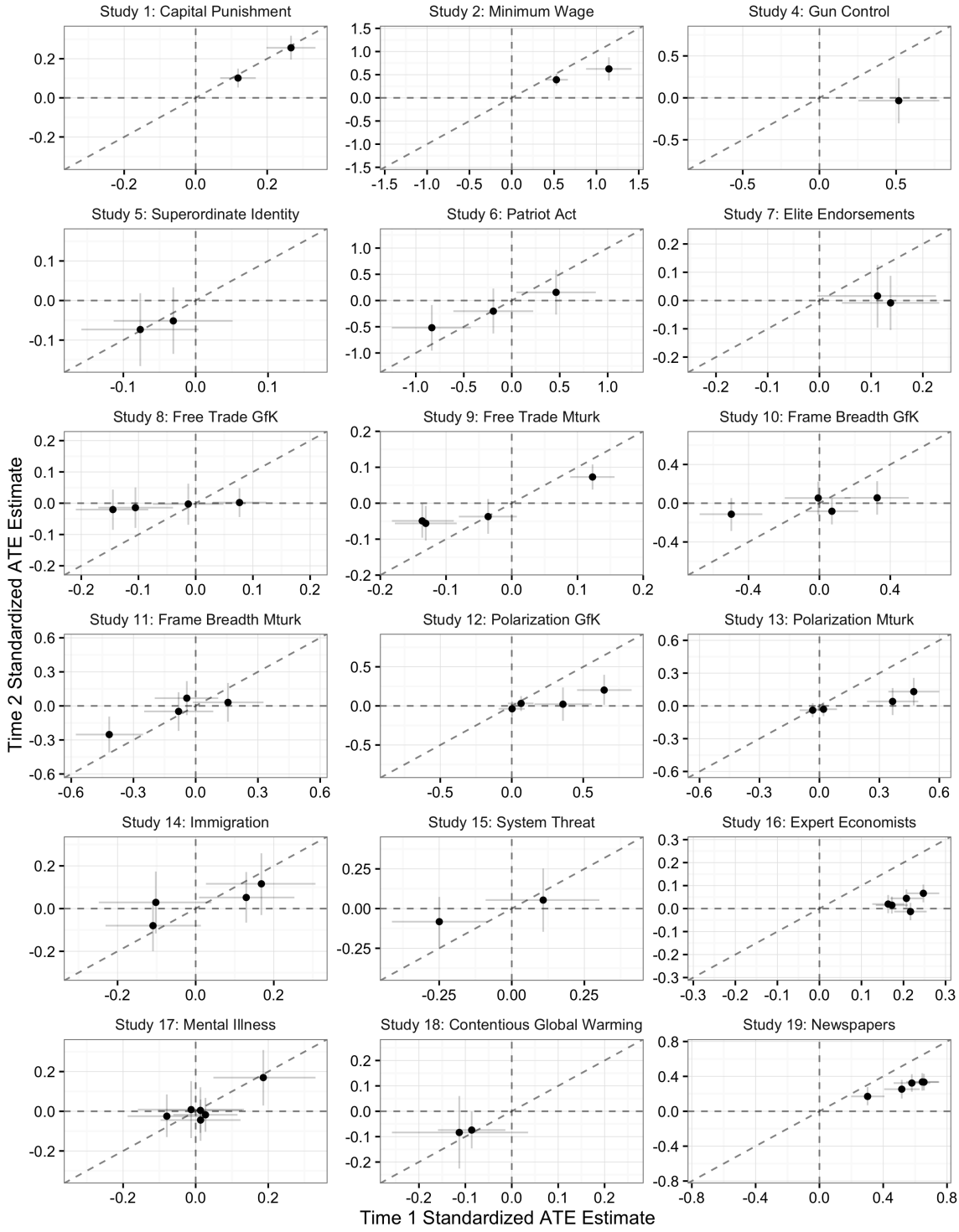
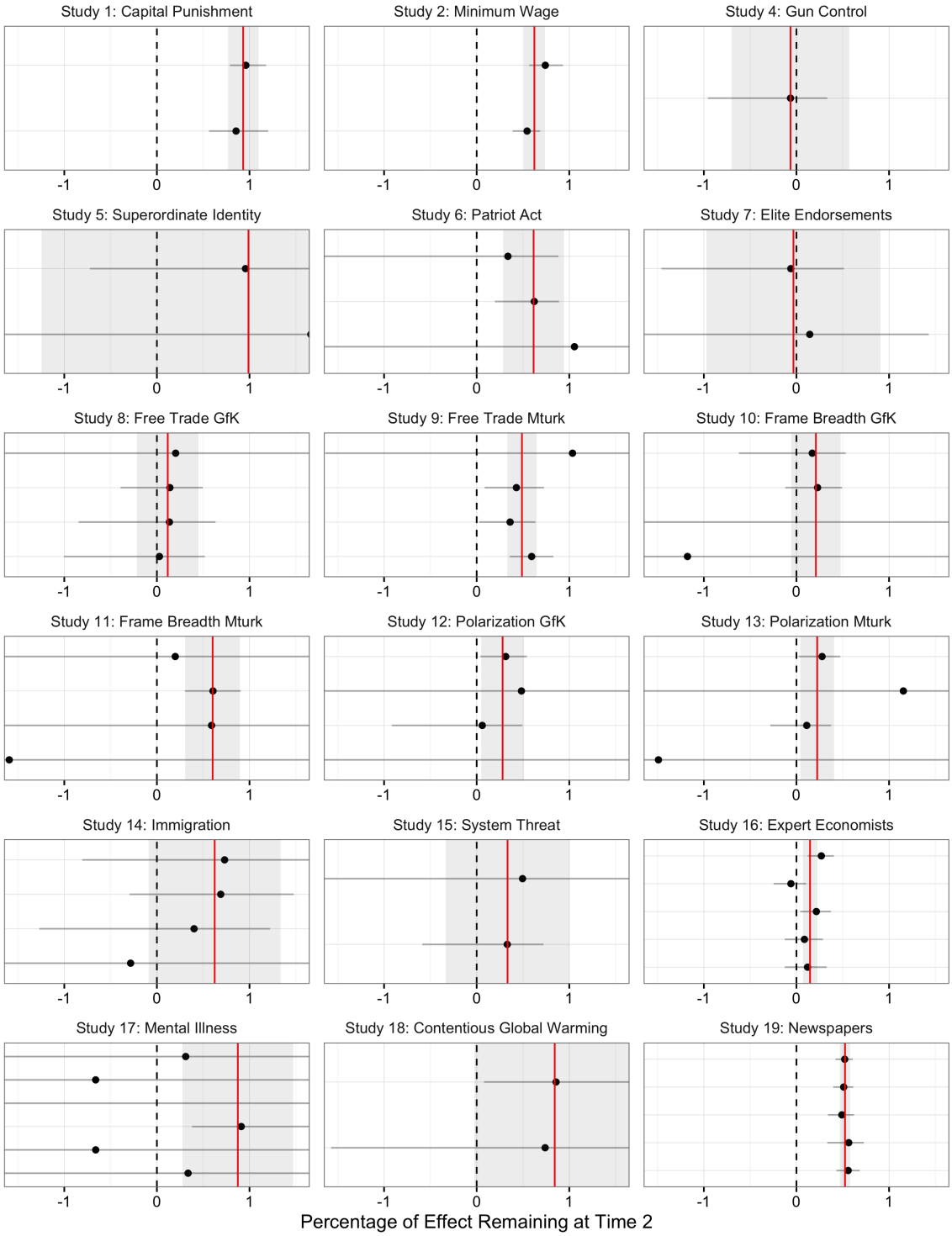


Figure 5.2: Persistence Estimates for 18 Studies



studies in each prediction category (weak, moderate, strong), weighting each study's average persistence by the inverse of its squared standard error. The table shows that on average, studies predicted to have weak persistence were at 18% of their original strength after 10 days. Those predicted to have moderate persistence were at 34% the original magnitude, while studies with a strong persistence prediction were at 58% of the time 1 effect at time 2. The differences across the prediction groups are all statistically significant at $p < 0.01$. Broadly speaking, these results confirm the predictions generated by the theoretical framework described above.

Table 5.2: Average Persistence, Pooled by Prediction

Prediction	Average Estimate	SE
Weak	0.18	0.04
Moderate	0.34	0.05
Strong	0.58	0.02

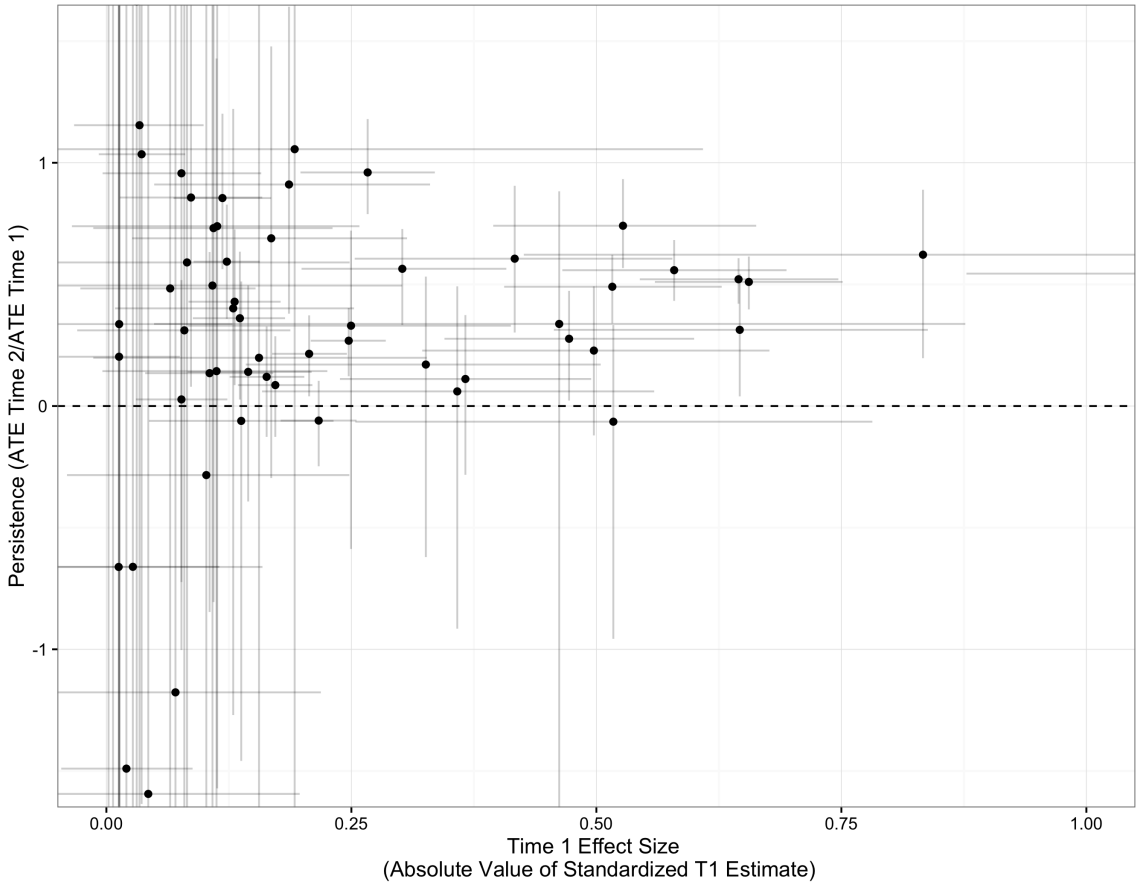
5.3.1 Alternative Explanation: Big Effects Persist?

To what extent can these results be explained by the more parsimonious theory that "big effects persist; small ones do not?" Figure 5.3 shows that this theory fails to provide meaningful insight into why some effects persist and others do not. On the x-axis, the effect size estimated at time 1 (in absolute value) is presented; the persistence estimate is on the y-axis. If the "big effects" theory of persistence were true, we would observe a positive correlation in these data. The relationship between these two variables, however, is essentially flat. The deviations from this flat relationship are highly heteroskedastic – the variance of the persistence estimates gets larger as the time 1 estimate gets weaker. Stated another way, the variance of a ratio estimate increases as the denominator gets closer to zero.

5.4 Long Term Stability

The persistence results presented thus far suggest that treatment effects decline to approximately half their original strength after 10 days. If effects continue to decay at the same

Figure 5.3: Magnitude of T1 ATE Versus Persistence



rate, we might project that they dissipate entirely after 20 days. Alternatively, effects might exhibit some measure of proportional decay, resulting in 25% strength after 20 days, 12.5% strength after 30 days, and so on. In order to trace a fuller picture of the rate of decay, we need to measure outcomes at more points in time.

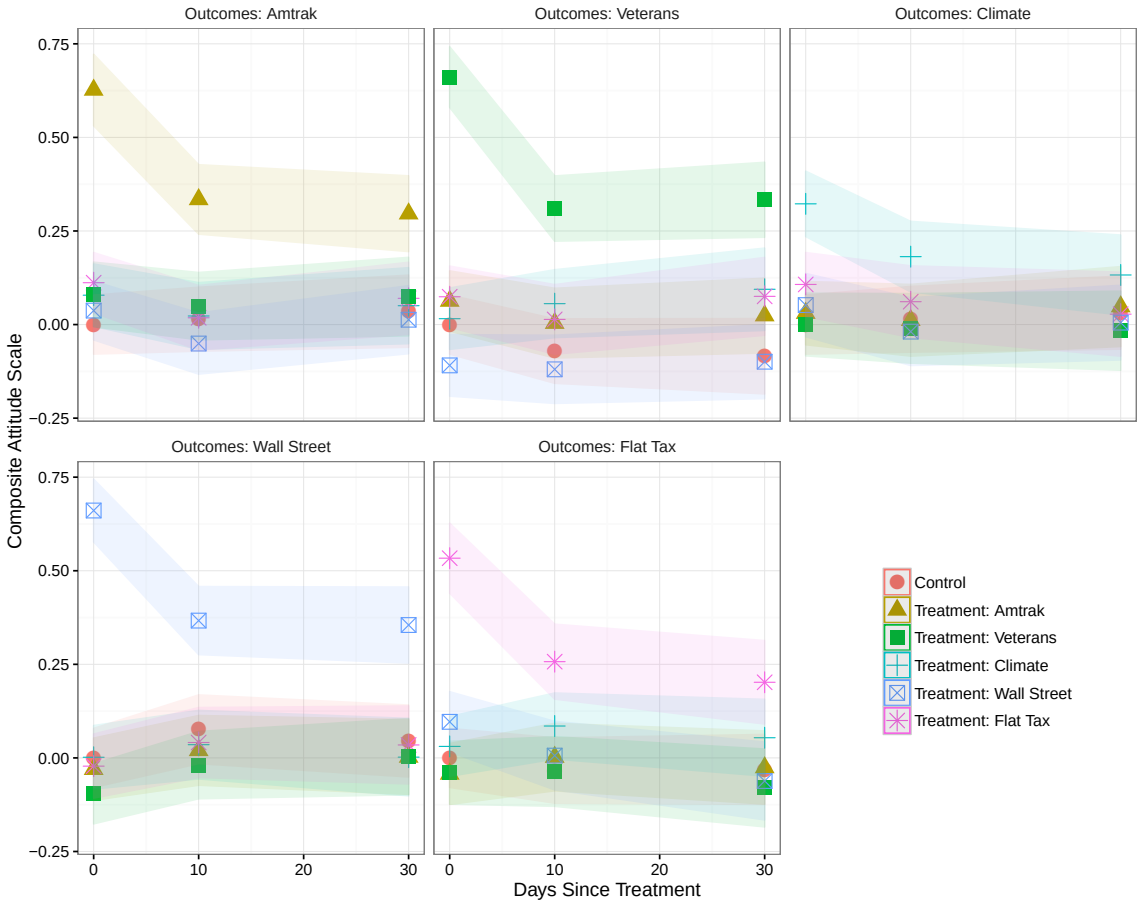
One of the studies (Study 17: Newspapers) included in the batch of 18 analyzed above included a third wave of post-treatment measurement after 30 days. That study employed a 6-group design: 5 groups that read an op-ed on a particular policy area and one control group. All subjects answered a series of questions in each issue area. These questions were combined into a composite scale using principal components analysis. Question wordings and details of scale construction are available in Appendix A.

Figure 5.4 displays the results of this experiment. In each panel, the average outcomes for a single policy area are shown on the vertical axis. The horizontal axis is the number of days since treatment, and displays three measurements for each of the six randomly-formed groups: immediately after reading the op-eds, 10 days after treatment, and 30 days after treatment.

The pattern of results is extraordinarily consisent across all five issue areas. The group that was treated with the issue-specific op-ed has much higher outcomes immediately after zero; the remaining groups cannot be distinguished from control. After 10 days, as we saw above, treatment effects decay to 52.1% (SE=2.4%) their original magnitudes. Remarkably, there appears to be very little further decay appears to take place between 10 days and 30 days after treatment. After 30 days, average persistence declines to 50.0% (SE = 2.4%), indicating that the change in treatment effect persistence between day 10 and day 30 is not statistically significant.

These results suggest a "hockey stick" pattern of decay: after an initial decline, subsequent decreases are smaller. This pattern can be reconciled with the larger theoretical setup described above: when new considerations are introduced, they arrive with artificially high weights attached. Over time, these weights "settle," but the consideration itself remains in the mind of the subject. This conjecture should be subjected to further empirical testing in experiments that track subjects' responses to treatment over longer periods of time.

Figure 5.4: Long Term Effects of Op-eds



5.5 Discussion

The results from this diverse set of studies yields a striking summary conclusion: on average, survey experimental treatment effects are approximately half their immediate size after 10 days. The challenge set by Gaines et al. (2007) was to demonstrate that survey experimental estimates of effects are politically relevant by showing that such effects endure beyond the few minutes it takes to complete a public opinion survey. By that standard, I have demonstrated that persuasive treatments of the kind studied here likely do have important and long-lasting impacts on political discourse.

The theoretical framework above predicted that treatments that operate primarily by making considerations more accessible would have fleeting effects, and treatments that operate by creating new associations or providing new information would last longer. This pattern was confirmed: the treatments in studies 1, 2, 6, 18 and 19 provided subjects with new information and all had strongly persistent effects. Studies 7 and 15 were predicted to increase the accessibility of certain considerations; the effects of these treatments appear not to have persisted after 10 days.

These results also offer an opportunity to revisit the large literature in psychology on the so-called "sleeper effect" (Hovland et al. 1949; Hovland and Weiss 1951; Cook and Flay 1978), in which initially null effects would blossom into strong effects in time. The proposed mechanism is that subjects would forget why they initially discounted some new piece of information; when the discounting falls away, the information would exert some persuasive influence. Attempts to document the existence of the sleeper effect have usually failed. Consistent with this line of evidence, none of the studies reported here saw an initially null result that became statistically significant. I do grant, however, that a determined advocate of the sleeper effect would rightly point out that the required theoretical conditions are probably not present in these experiments.

These results indicate that the psychological mechanisms of *accessibility*, *applicability*, and *new information* may indeed play a role in the persistence of treatment effects; they also may play a role in the strength of treatment effects measured immediately. However, the treatments in this set of studies do not only vary along these dimensions. For example,

each one is concerned with a different issue area. Further, some deliver treatments as questions and others as statements. In order to determine the extent to which each of these mechanisms influences persistence, it would be preferable to hold all other features of the treatment constant and randomly assign the mechanisms. I leave this project to future research.

References

Abelson, Robert P. and James C. Miller. 1967. "Negative Persuasion Via Personal Insult." *Journal of Experimental Social Psychology* 3(4):321–333.

Achen, Christopher H. 1992. "Social Psychology, Demographic Variables, and Linear Regression: Breaking the Iron Triangle in Voting Research." *Political Behavior* 14(3):195–211.

Ajzen, Icek and Martin Fishbein. 1980. *Understanding Attitudes and Predicting Social Behavior*. Vol. 278 Englewood Cliffs, NJ: Prentice Hall.

Allport, Gordon. W. 1935. Attitudes. In *Handbook of Social Psychology*, ed. C. Murchison. Worcester, MA: Clark University Press pp. 798 – 844.

Althaus, Scott L. 2003. *Collective Preferences in Democratic Politics*. United Kingdom: Cambridge University Press.

Arceneaux, Kevin and Martin Johnson. 2013. *Changing Minds Or Changing Channels?: Partisan News in an Age of Choice*. University of Chicago Press.

Aronow, Peter M., Alexander Coppock, Alan S. Gerber, Donald P. Green and Holger Kern. 2015. "Double Sampling for Missing Outcome Data in Randomized Experiments." *Unpublished Manuscript* pp. 1–20.

Baden, Christian and Sophie Lecheler. 2012. "Fleeting, Fading, or Far-Reaching? A Knowledge-Based Model of the Persistence of Framing Effects." *Communication Theory* 22(4):359–382.

Bakshy, Eytan, Solomon Messing and Lada A Adamic. 2015. "Exposure to ideologically diverse news and opinion on Facebook." *Science* 348(6239):1130–1132.

Bartels, Larry. 2002. "Beyond the Running Tally: Partisan Bias in Political Perceptions." *Political Behavior* 24(2):117–150.

Bartels, Larry M. 1994. "The American Publics Defense Spending Preferences in the Post-Cold-War Era." *Public Opinion Quarterly* 58(4241):479–508.

Benoît, Jean-Pierre and Juan Dubra. 2014. "A Theory of Rational Attitude Polarization." *Available at SSRN 2529494* .

Berinsky, Adam, Gregory Huber and Gabriel Lenz. 2012*a*. "Evaluating Online Labor Markets for Experimental Research: Amazon.com's Mechanical Turk." *Political Analysis* 20(3).

Berinsky, Adam J., Gregory A. Huber and Gabriel S. Lenz. 2012*b*. "Evaluating Online Labor Markets for Experimental Research: Amazon.com's Mechanical Turk." *Political Analysis* 20(3):351–368.

Borenstein, Michael, Larry V. Hedges, Julian P.T. Higgins and Hannah R. Rothstein. 2009. *Introduction to Meta-Analysis*. Hoboken, NJ: John Wiley & Sons.

Brader, Ted. 2005. "Striking a Responsive Chord: How Political Ads Motivate and Persuade Voters by Appealing to Emotions." *American Journal of Political Science* 49(2):388–405.

Brader, Ted, Nicholas A Valentino and Elizabeth Suhay. 2008. "What Triggers Public Opposition to Immigration? Anxiety, Group Cues, and Immigration Threat." *American Journal of Political Science* 52(4):959–978.

Brooks, Clem and Jeff Manza. 2013. "A Broken Public? Americans' Responses to the Great Recession." *American Sociological Review* 78:727–748.

Buhrmester, Michael D., Tracy Kwang and Samuel D. Gosling. 2011. "Amazon's Mechanical Turk: A New Source of Inexpensive, yet High-Quality, Data?" *Perspectives on Psychological Science* 6(1):3–5.

Bullock, John G. 2009. "Partisan Bias and the Bayesian Ideal in the Study of Public Opinion." *The Journal of Politics* 71(03):1109–1124.

Campbell, Angus, Philip E. Converse, Warren E. Miller and Donald E. Stokes. 1960. *The American Voter*. New York: Wiley.

Chandler, Jesse, Gabriele Paolacci, Eyal Peer, Pam Mueller and Kate Ratliff. in press. "Non-Naïve Participants Can Reduce Effect Sizes." *Psychological Science* .

Chipman, Hugh A., Edward I. George and Robert E. McCulloch. 2007. "Bayesian Ensemble Learning." *Advances in Neural Information Processing Systems* 19:265.

Chipman, Hugh A., Edward I. George and Robert E. McCulloch. 2010. "BART: Bayesian Additive Regression Trees." *Annals of Applied Statistics* 4(1):266–298.

Chong, Dennis and James N. Druckman. 2010. "Dynamic Public Opinion: Communication Effects over Time." *American Political Science Review* 104(04):663–680.

Chong, Dennis and James N. Druckman. 2012. Dynamics in Mass Communication Effects Research. In *The Sage Handbook of Political Communication*, ed. Holli Semetko and Maggie Scammell. Los Angeles, CA: Sage Publications pp. 307–323.

Clemens, Michael A. 2015. "The Meaning of Failed Replications: A Review and Proposal." *Center for Global Development Working Paper* 399.

Clifford, Scott and Jennifer Jerit. 2014. "Is There a Cost to Convenience? An Experimental Comparison of Data Quality in Laboratory and Online Studies." *Journal of Experimental Political Science* .

Cohen, Jacob. 1992. "A Power Primer." *Psychological Bulletin* 112(1):155.

Cook, Thomas D and Brian R Flay. 1978. "The Persistence of Experimentally induced Attitude Change." *Advances in Experimental Social Psychology* 11:1–57.

Coppock, Alexander and Donald P. Green. 2015. "Assessing the Correspondence Between Experimental Results Obtained in the Lab and Field: A Review of Recent Social Science Research." *Political Science Research and Methods* 3(1):113–131.

Corner, Adam, Lorraine Whitmarsh and Dimitrios Xenias. 2012. "Uncertainty, Scepticism and Attitudes towards Climate Change: Biased Assimilation and Attitude Polarisation." *Climatic Change* 114(3-4):463–478.

Craig, Maureen A. and Jennifer A. Richeson. 2014. "More Diverse Yet Less Tolerant? How the Increasingly Diverse Racial Landscape Affects White Americans' Racial Attitudes." *Personality and Social Psychology Bulletin* 40(6):750–761.

Cronbach, Lee J, Karen Shapiro et al. 1982. *Designing Evaluations of Educational and Social Programs*. Jossey-Bass.

de Finetti, Bruno. 1937. La Prévision: ses Lois Logiques, ses Sources Subjectives. In *Annales de l'institut Henri Poincaré*. Vol. 7 Presses universitaires de France pp. 1–68.

de Vreese, Claes. 2004. "The Effects of Strategic News on Political Cynicism, Issue Evaluations, and Policy Support: A Two-Wave Experiment." *Mass Communication and Society* 7(2):191–214.

Degeorges, Adrien and Frédéric Gonthier. 2012. "'Plus ça change, plus c'est la même chose': The Evolution and the Structure of Attitudes Toward Economic Liberalism in France Between 1990 And 2008." *French Politics* 10(3):233–268.

DiMaggio, Paul, John Evans and Bethany Bryson. 1996. "Have American's Social Attitudes Become More Polarized?" *American Journal of Sociology* 102(3):690.

Ding, Peng, Avi Feller and Luke Miratrix. 2015. "Randomization Inference for Treatment Effect Variation." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* .

Dixit, Avinash K and Jörgen W Weibull. 2007. "Political Polarization." *Proceedings of the National Academy of Sciences of the United States of America* 104(2):7351–7356.

Druckman, James N. 2001. "The Implications of Framing Effects for Citizen Competence." *Political Behavior* 23(3):225–256.

Druckman, James N., Erik Peterson and Rune Slothuus. 2013. "How Elite Partisan Polarization Affects Public Opinion Formation." *American Political Science Review* 107(01):57–79.

Druckman, James N. and Kjersten R. Nelson. 2003. "Framing and Deliberation: How Conversations Limit Elite Influence." *American Journal of Political Science* 47(4):729–745.

Druckman, James N., Matthew S. Levendusky and Audrey McLain. 2015. "No Need to Watch: How the Effects of Partisan Media Can Spread via Inter-Personal Discussions." *Unpublished Manuscript* .

Edwards, Kari and Edward E. Smith. 1996. "A Disconfirmation Bias in the Evaluation of Arguments." *Journal of Personality and Social Psychology* 71(1).

Edwards, Ward, Harold Lindman and Leonard J Savage. 1963. "Bayesian Statistical Inference for Psychological Research." *Psychological Review* 70(3):193.

Edwards, Ward and Lawrence D Phillips. 1964. "Man as Tranducer for Probabilities in Bayesian Command and Control Systems.".

Eichenberg, Richard C. and Richard J. Stoll. 2012. "Gender Difference or Parallel Publics? The Dynamics of Defense Spending Opinions in the United States, 1965–2007." *Journal of Conflict Resolution* 56(2):331–348.

Ellis, Christopher R. and Joseph Daniel Ura. 2011. United We Divide?: Education, Income, and Heterogeneity in Mass Partisan Polarization. In *Who Gets Represented?*, ed. Peter K. Enns and Christopher Wlezien. New York: Russell Sage.

Enns, Peter K. 2007. The Uniform Nature of Mass Opinion PhD thesis University of North Carolina at Chapel Hill.

Enns, Peter K. and Christopher Wlezien. 2011. Group Opinion and the Study of Representation. In *Who Gets Represented?*, ed. Peter K. Enns and Christopher Wlezien. New York: Russell Sage pp. 1–25.

Enns, Peter K. and Gregory E. McAvoy. 2012. "The Role of Partisanship in Aggregate Opinion." *Political Behavior* 34(4):627–651.

Enns, Peter K. and Julianna Koch. 2013. "Public Opinion in the U.S. States: 1956 to 2010." *State Politics and Policy Quarterly* 13(3):349–372.

Enns, Peter K. and Paul M. Kellstedt. 2008. "Policy Mood and Political Sophistication: Why Everybody Moves Mood." *British Journal of Political Science* 38(03):433–454.

Erikson, Robert S., Michael B. MacKuen and James A. Stimson. 2002. *The Macro Polity*. New York: Cambridge University Press.

Franco, Annie, Neil Malhotra and Gabor Simonovits. 2014. "Publication Bias in the Social Sciences: Unlocking the File Drawer." *Science* 345(6203):1502–1505.

Gaines, Brian J., James H. Kuklinski and Paul J. Quirk. 2007. "The Logic of the Survey Experiment Reexamined." *Political Analysis* 15:1–20.

Gerber, Alan and Donald Green. 1999. "Misperceptions About Perceptual Bias." *Annual Review of Political Science* 2(1):189–210.

Gerber, Alan S. and Donald P. Green. 2012. *Field Experiments: Design, Analysis, and Interpretation.* New York: W.W. Norton.

Gerber, Alan S., Neil Malhotra, Conor M. Dowling and David Doherty. 2010. "Publication Bias in Two Political Behavior Literatures." *American Politics Research* 38(4):591–613.

Gillion, Daniel Q., Jonathan M. Ladd and Marc Meredith. 2015. "Education, Party Polarization and the Origins of the Partisan Gender Gap." *Unpublished Manuscript* pp. 1–61.

Goethals, George R and Richard F Reckman. 1973. "The Perception of Consistency in Attitudes." *Journal of Experimental Social Psychology* 9(6):491–501.

Gonthier, Frédéric. 2015. "Qui Bouge Quand L'Opinion Bouge ?" *Revue Française de Science Politique* 65(1):61.

Green, Donald, Bradley Palmquist and Eric Schickler. 2002. *Partisan Hearts and Minds.* New Haven, Conn.: Yale University Press.

Green, Donald P. and Holger L. Kern. 2012. "Modeling Heterogeneous Treatment Effects in Survey Experiments with Bayesian Additive Regression Trees." *Public Opinion Quarterly* 76(3):491–511.

Haider-Markel, Donald P. and Mark R. Joslyn. 2001. "Gun Policy, Opinion, Tragedy, and Blame Attribution: The Conditional Influence of Issue Frames." *The Journal of Politics* 63(2):520–543.

Henrich, Joseph, Steven J. Heine and Ara Norenzayan. 2010. "The Weirdest People in the World?" *Behavioral and Brain Sciences* 33:61–83.

Hill, Jennifer L. 2011. "Bayesian Nonparametric Modeling for Causal Inference." *Journal of Computational and Graphical Statistics* 20(1):217–240.

Hiscox, Michael J. 2006. "Through a Glass and Darkly: Attitudes Toward International Trade and the Curious Effects of Issue Framing." *International Organization* 60(03):755–780.

Hlavac, Marek. 2014. *stargazer: LaTeX/HTML Code and ASCII text for Well-formatted Regression and Summary Statistics Tables.* Cambridge, USA: Harvard University.

Holm, Sture. 1979. "A Simple Sequentially Rejective Multiple Test Procedure." *Scandinavian Journal of Statistics* pp. 65–70.

Hopkins, Daniel J. 2012. "Whose Economy? Perceptions of National Economic Performance During Unequal Growth." *Public Opinion Quarterly* 76(1):50–71.

Hopkins, Daniel J. and Jonathan Mummolo. 2015. "Assessing the Breadth of Framing Effects." *Unpublished Manuscript* .

Houston, David A. and Russell H. Fazio. 1989. "Biased Processing as a Function of Attitude Accessibility: Making Objective Judgments Subjectively." *Social Cognition* 7(1):51–66.

Hovland, Carl I., Arthur A. Lumsdaine and Fred D. Sheffield. 1949. *Experiments on Mass Communication, Vol. 3.* Princeton, NJ: Princeton University Press.

Hovland, Carl I and Walter Weiss. 1951. "The Influence of Source Credibility on Communication Effectiveness." *Public Opinion Quarterly* 15(4):635–650.

Humphreys, Macartan, Raul Sanchez de la Sierra and Peter van der Windt. 2013. "Fishing, Commitment, and Communication: A Proposal for Comprehensive Nonbinding Research Registration." *Political Analysis* 21(1):1–20.

Huxster, Joanna K., Jason Carmichael and Robert J. Brulle. 2015. "A Macro Political Examination of the Partisan and Ideological Divide in Aggregate Public Concern over Climate Change in the U.S. between 2001 and 2013." *Environmental Management and Sustainable Development* 4(1):1–15.

Ioannidis, John P. A. 2005. "Why Most Published Research Findings are False." *PLoS Medicine* 2(8):e124.

Johnson, Tyler and Paul M. Kellstedt. 2014. "Media Consumption and the Dynamics of Policy Mood." *Political Behavior* 36(2):377–399.

Johnston, Christopher D. and Andrew O. Ballard. 2014. "Economists and Public Opinion: Expert Consensus and Economic Policy Judgments." *Unpublished Manuscript* .

Kahan, Dan M. 2013. "Fooled Twice, Shame on Who? Problems with Mechanical Turk Study Samples, Part 2.".

Kahan, Dan M. 2015. "The Politically Motivated Reasoning Paradigm." *Emerging Trends in Social & Behavioral Sciences, Forthcoming* .

Karlin, Samuel and Herman Rubin. 1956. "The Theory of Decision Procedures for Distributions with Monotone Likelihood Ratio." *The Annals of Mathematical Statistics* 27(2):272–299.

Kellstedt, Paul M. 2003. *The Mass Media and the Dynamics of American Racial Attitudes.* New York: Cambridge University Press.

Kellstedt, Paul M., David A. M. Peterson and Mark D. Ramirez. 2010. "The Macro Politics of a Gender Gap." *Public Opinion Quarterly* 74(3):477–498.

Kelly, Nathan J. and Peter K. Enns. 2010. "Inequality and the Dynamics of Public Opinion: The Self-Reinforcing Link Between Economic Inequality and Mass Preferences." *American Journal of Political Science* 54(4):855–870.

Kruglanski, Arie W. and Donna M. Webster. 1996. "Motivated Closing of the Mind: 'Seizing' and 'Freezing'." *Psychological Review* 103:263–283.

Krupnikov, Yanna and Adam Seth Levine. 2014. "Cross-sample Comparisons and External Validity." *Journal of Experimental Political Science* 1(01):59–80.

Kuhn, Deanna and Joseph Lao. 1996. "Effects of Evidence on Attitudes: Is Polarization the Norm." *Psychological Science* 7(2):115–120.

Kunda, Ziva. 1990. "The Case for Motivated Reasoning." *Psychological Bulletin* 108(3):480–498.

Kyburg, H.E. and H.E. Smokler, eds. 1964. *Studies in Subjective Probability*. John Wiley & Sons.

Lau, R.R. and D.P. Redlawsk. 2006. *How Voters Decide: Information Processing During Election Campaigns*. Cambridge University Press.

Lecheler, Sophie and Claes H. de Vreese. 2011. "Getting Real: The Duration of Framing Effects." *Journal of Communication* 61(5):959–983.

Leeper, Thomas J. 2015. *MTurkR: Access to Amazon Mechanical Turk Requester API via R*. R package version 0.6.5.1.

Leeper, Thomas J. and Rune Slothuus. 2014. "Political Parties, Motivated Reasoning, and Public Opinion Formation." *Advances in Political Psychology* 35(SUPPL.1):129–156.

Levendusky, Matthew and Neil Malhotra. 2015. "Does Media Coverage of Partisan Polarization Affect Political Attitudes ?" *Political Communication* .

Levitt, Steven D. and John A. List. 2007. "What do Laboratory Experiments Measuring Social Preferences Reveal about the Real World?" *The journal of economic perspectives* pp. 153–174.

Lord, Charles S., L. Ross and M. Lepper. 1979. "Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence." *Journal of Personality and Social Psychology* 37:2098–2109.

Lucas, Jeffrey W. 2003. "Theory-testing, Generalization, and The Problem of External Validity." *Sociological Theory* 21(3):236–253.

Mayer, Jeremy D. 2004. "Christian Fundamentalists and Public Opinion Toward the Middle East: Israel's New Best Friends?" *Social Science Quarterly* 85(3):695–712.

McDermott, Rose. 2002. "Experimental Methodology in Political Science." *Political Analysis* 10(4):325–342.

McGinty, Emma E., Daniel W. Webster and Colleen L. Barry. 2013. "Effects of News Media Messages about Mass Shootings on Attitudes Toward Persons with Serious Mental Illness and Public Support for Gun Control Policies." *American Journal of Psychiatry* 170(5):494–501.

Miller, Arthur G., John W. McHoskey, Cynthia M. Bane and Timothy G. Dowd. 1993. "The Attitude Polarization Phenomenon: Role of Response Measure, Attitude Extremity, and Behavioral Consequences of Reported Attitude Change." *Journal of Personality and Social Psychology* 64(4):561–574.

Morgan, Stephen and Minhyoung Kang. 2015. "A New Conservative Cold Front? Democrat and Republican Responsiveness to the Passage of the Affordable Care Act." *Sociological Science* 2:502–526.

Mullinix, Kevin J., Thomas J. Leeper, James N. Druckman and Jeremy Freese. 2015. "The Generalizability of Survey Experiments." *Journal of Experimental Political Science* 2:109–138.

Munro, Geoffrey D. and Peter H. Ditto. 1997. "Biased Assimilation, Attitude Polarization, and Affect in Reactions to Stereotype-Relevant Scientific Information." *Personality and Social Psychology Bulletin* 23(6):636–653.

Mutz, Diana C. and Byron Reeves. 2005. "The New Videomalaise: Effects of Televised Incivility on Political Trust." *American Political Science Review* 99(1):1–15.

Nelson, Thomas E., Rosalee A. Clawson and Zoe M. Oxley. 1997. "Media Framing of a Civil Liberties Conflict and Its Effect on Tolerance." *American Political Science Review* 91(3):567–583.

Nelson, Thomas E., Zoe M. Oxley and Rosalee A. Clawson. 1997. "Toward a Psychology of Framing Effects." *Political Behavior* 19(3):221–246.

Nicholson, Stephen P. 2012. "Polarizing Cues." *American Journal of Political Science* 56(1):52–66.

Nincic, Miroslav. 1997. "Domestic Costs, the U.S. Public, and the Isolationist Calculus." *International Studies Quarterly* 41(4):593–610.

Nincic, Miroslav and Donna J. Nincic. 2002. "Race, Gender, and War." *Journal of Peace Research* 39(5):547–568.

Nisbett, Richard E and Lee Ross. 1980. "Human Inference: Strategies and Shortcomings of Social Judgment.".

Nyhan, Brendan and Jason Reifler. 2010. "When Corrections Fail: The Persistence of Political Misperceptions." *Political Behavior* 32(2):303–330.

Nyhan, Brendan and Jason Reifler. 2015. "Does Correcting Myths about the Flu Vaccine Work? An Experimental Evaluation of the Effects of Corrective Information." *Vaccine* 33(3):459–464.

Nyhan, Brendan, Jason Reifler, Sean Richey and Gary L. Freed. 2014. "Effective Messages in Vaccine Promotion: A Randomized Trial." *Pediatrics* .

Open Science Collaboration. 2015. "Estimating the Reproducibility of Psychological Science." *Science* 349(6251).

Pacheco, Julianna. 2014. "Measuring and Evaluating Changes in State Opinion Across Eight Issues." *American Politics Research* 42(6):986–1009.

Page, Benjamin I. and Robert Y. Shapiro. 1992. *The Rational Public: Fifty Years of Trends in Americans' Policy Preferences*. Chicago: University of Chicago Press.

Peffley, Mark and Jon Hurwitz. 2007. "Persuasion and Resistance: Race and the Death Penalty in America." *American Journal of Political Science* 51(4):996–1012.

Peterson, Cameron R and Alan J Miller. 1965. "Sensitivity of Subjective Probability Revision." *Journal of Experimental Psychology* 70(1):117.

Phillips, Lawrence D. and Ward Edwards. 1966. "Conservatism in a Simple Probability Inference Task." *Journal of Experimental Psychology* 72(3):346–354.

Price, Vincent and John Zaller. 1993. "Who Gets the News? Alternative Measures of News Reception and Their Implications for Research." *Public Opinion Quarterly* 57(2):133–164.

Prior, Markus. 2007. *Post-Broadcast Democracy: How Media Choice Increases Inequality in Political Involvement and Polarizes Elections*. Cambridge University Press.

Pyszczynski, Tom, Jeff Greenberg and Kathleen Holt. 1985. "Maintaining Consistency between Self-Serving Beliefs and Available Data: A Bias in Information Evaluation." *Personality and Social Psychology Bulletin* 11(2):179–190.

R Core Team. 2015. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Ramsey, Frank P. 1931. Truth and Probability. In *The Foundations of Mathematics and other Logical Essays*, ed. R.B. Braithwaite. Harcourt, Brace and Company pp. 156–198.

Redlawsk, David P. 2002. "Hot Cognition or Cool Consideration? Testing the Effects of Motivated Reasoning on Political Decision Making." *The Journal of Politics* 64(4):1021–1044.

Redlawsk, David P., Andrew J. W. Civettini and Karen M. Emmerson. 2010. "The Affective Tipping Point: Do Motivated Reasoners Ever "Get It"?" *Political Psychology* 31(4):563–593.

Ross, Lee. 2012. "Reflections on Biased Assimilation and Belief Polarization." *Critical Review* 24(2):233–245.

Rubin, Donald B. 1974. "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology* 66(5):688.

Samii, Cyrus and Peter M. Aronow. 2012. "On Equivalencies Between Design-based and Regression-based Variance Estimators for Randomized Experiments." *Statistics & Probability Letters* 82(2):365–370.

Sears, David O. 1986. "College Sophomores in the Laboratory: Influences of a Narrow Data Base on Social Psychology's View of Human Nature." *Journal of Personality and Social Psychology* 51(3):515–530.

Shapiro, Robert Y and Yaeli Bloch-Elkon. 2008. "Do the Facts Speak for Themselves? Partisan Disagreement as a Challenge to Democratic Competence." *Critical Review* 20(1-2):115–139.

Simonsohn, Uri, Leif D. Nelson and Joseph P. Simmons. 2014. "P-curve: A Key to the File-drawer." *Journal of Experimental Psychology: General* 143(2):534–547.

Snyder, Jack, Robert Y. Shapiro and Yaeli Bloch-Elkon. 2009. "Free Hand Abroad, Divide and Rule at Home." *World Politics* 61(01):155.

Soroka, Stuart N. and Christopher Wlezien. 2008. "On the Limits to Inequality in Representation." *PS: Political Science & Politics* pp. 319–327.

Strickland, April A., Charles S. Taber and Milton Lodge. 2011. "Motivated Reasoning and Public Opinion." *Journal of Health Politics, Policy and Law* 36(6):935–944.

Taber, Charles S., Damon Cann and Simona Kucsova. 2009. "The Motivated Processing of Political Arguments." *Political Behavior* 31(2):137–155.

Taber, Charles S. and Milton Lodge. 2006. "Motivated Skepticism in the Evaluation of Political Beliefs." *American Journal of Political Science* 50(3):755–769.

Taydas, Zeynep, Cigdem Kentmen and Laura R. Olson. 2012. "Faith Matters: Religious Affiliation and Public Opinion About Barack Obama's Foreign Policy in the "Greater" Middle East." *Social Science Quarterly* 93(5):1218–1242.

Tewksbury, David H., J. Jones, M. W. Peske, A. Raymond and W. Vig. 2000. "The Interaction of News and Advocate Frames: Manipulating Audience Perceptions of a Local Public Policy Issue." *Journalism & Mass Communication Quarterly* 77(4):804–829.

Tipton, Elizabeth. 2012. "Improving Generalizations From Experiments Using Propensity Score Subclassification Assumptions, Properties, and Contexts." *Journal of Educational and Behavioral Statistics* p. 1076998612441947.

Transue, John E. 2007. "Identity Salience, Identity Acceptance, and Racial Policy Attitudes: American National Identity as a Uniting Force." *American Journal of Political Science* 51(1):78–91.

Ura, Joseph Daniel and Christopher R. Ellis. 2008. "Income, Preferences, and the Dynamics of Policy Responsiveness." *PS: Political Science & Politics* 41(04):785–794.

White, Halbert. 2000. "A Reality Check for Data Snooping." *Econometrica* 68(5):1097–1126.

Wickham, Hadley. 2009. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

Wickham, Hadley and Romain Francois. 2015. *dplyr: A Grammar of Data Manipulation*. R package version 0.4.3.

Wlezien, Christopher and Stuart N. Soroka. 2011. Inequality in Policy Responsiveness? In *Who Gets Represented?*, ed. Peter K. Enns and Christopher Wlezien. New York: Russell Sage pp. 285–310.

Zaller, John R. 1992. *The Nature and Origins of Mass Opinion*. New York: Cambridge University Press.

Zechman, Martin J. 1979. "Dynamic Models of the Voter's Decision Calculus: Incorporating Retrospective Considerations into Rational-Choice Models of Individual Voting Behavior." *Public Choice* 34:297–315.

A Study Manifest

A.1 Study 1: Capital Punishment

This study was conducted in collaboration with Andrew Guess. It embeds a replication of the Lord et al. (1979) laboratory demonstration in a larger random-assignment study. The results of the experiment indicate that presenting subjects with pro or con information about capital punishment changed both their beliefs about and support for the policy.

Sample: 683 Mechanical Turk subjects in wave 2, 628 in wave 3. In wave 1, pre-treatment attitudes and beliefs were measured and treatments were allocated in wave 2.

Dates of data collection: May 2014.

Treatments: There were six sets of experimental stimuli: The content of the reports could be Pro, Con, or Null, and the study methodology could be time series or cross sectional. Each report contained two sections: Subjects were first presented study summaries, then study details and criticisms. Subjects saw two sets of evidence. In total, subjects could be assigned to one of 18 combinations of treatments, as shown in the table below. Conditions 7, 8, 11, and 12 correspond to the conditions to which subjects could be assigned in the original Lord et al. (1979) experiment.

Table A.1: Capital Punishment Treatment Conditions

	1st Study		2nd Study				
	Content	Method	Content	Method	Condition	Info Content	N
1	Con	Cross	Con	Time	Con Con	-2	62
2	Con	Time	Con	Cross	Con Con	-2	54
3	Con	Cross	Null	Time	Con Null	-1	31
4	Con	Time	Null	Cross	Con Null	-1	30
5	Null	Cross	Con	Time	Con Null	-1	28
6	Null	Time	Con	Cross	Con Null	-1	26
7	Con	Cross	Pro	Time	Pro Con	0	26
8	Con	Time	Pro	Cross	Pro Con	0	36
9	Null	Cross	Null	Time	Null Null	0	67
10	Null	Time	Null	Cross	Null Null	0	45
11	Pro	Cross	Con	Time	Pro Con	0	23
12	Pro	Time	Con	Cross	Pro Con	0	33
13	Null	Cross	Pro	Time	Pro Null	1	28
14	Null	Time	Pro	Cross	Pro Null	1	32
15	Pro	Cross	Null	Time	Pro Null	1	32
16	Pro	Time	Null	Cross	Pro Null	1	28
17	Pro	Cross	Pro	Time	Pro Pro	2	49
18	Pro	Time	Pro	Cross	Pro Pro	2	53
18	Pro	Pro	Time	Cross	Pro Pro	2	53

A.1.1 Experimental Materials: Study Summaries

- Pro Time

  Kroner and Phillips (2012) compared murder rates for the year before and the year after adoption of capital punishment in 14 states. In 11 of the 14 states, murder rates were **lower after** adoption of the death penalty.

  This research supports the deterrent effect of the death penalty.

- Con Time

  Kroner and Phillips (2012) compared murder rates for the year before and the year after adoption of capital punishment in 14 states. In 11 of the 14 states, murder rates were **higher after** adoption of the death penalty.

  This research opposes the deterrent effect of the death penalty.

- Null Time

  Kroner and Phillips (2012) compared murder rates for the year before and the year after adoption of capital punishment in 14 states. In 5 of the 14 states the murder rate was **lower** after adoption of capital punishment laws. Another 5 of the 14 states

showed the **opposite** pattern. In the remaining states, there was **no change** in the murder rate before and after the adoption of capital punishment laws.

This research is inconclusive regarding the deterrent effect of the death penalty.

- Pro Cross Palmer and Crandall (2012) compared murder rates in 10 pairs of neighboring states with different capital punishment laws. In 8 of the 10 pairs, murder rates were **lower** in the state with capital punishment. This research supports the deterrent effect of the death penalty.
- Con Cross Palmer and Crandall (2012) compared murder rates in 10 pairs of neighboring states with different capital punishment laws. In 8 of the 10 pairs, murder rates were **higher** in the state with capital punishment. This research opposes the deterrent effect of the death penalty.
- Null Cross Palmer and Crandall (2012) compared murder rates in 10 pairs of neighboring states with different capital punishment laws. In 4 of the 10 pairs, murder rates were **lower** in the state with capital punishment. In 4 of the 10 pairs, murder rates were **higher** in the state with capital punishment. In the remaining two pairs, murder rates were the **same** in both states. This research is inconclusive concerning the deterrent effect of the death penalty.

A.1.2 Experimental Materials: Study Details and Criticism

The format of the Study Details and Criticism section was the same in all conditions. Subjects first saw a description of the debate over capital punishment, followed by a description of the research design employed by the fictitious researchers. The next page of the survey presented the conclusions, table, and graph. The final page contained a critique of the design and a summary of a replication study (which always confirmed the original results). There are two main differences between these stimuli and those used by Lord et al.: First, the original study did not include either of the Null study reports. Second, the fictitious publication dates were 1977 for the main study, and 1978 for the replication. I also updated the tables and graphs using modern statistical software.

Pro Time Series Details and Criticism

Can crime be prevented by the death penalty?

In recent years there has arisen a great public debate over whether or not capital punishment (the death penalty) is effective in preventing murders. Those who favor capital punishment claim that the death penalty acts as a deterrent by reminding the potential murderer of the possible consequences, whereas those who oppose capital punishment claim that such officially sanctioned killing by the state sets a violent example that will produce murders that might have been avoided. A recent research effort attempted to shed light on this continuing debate.

The researchers (Kroner and Phillips, 2012) noticed that within the past twelve years fourteen states have passed capital punishment laws or reinstated the death penalty. They reasoned that if capital punishment laws deter murders, the murder rate should have been higher for the year just before the adoption of the death penalty than for the year just after the adoption of death penalty in each of these states.

The results, shown in the table and graph below, were that in 11 of the 14 states the murder rate was lower after adoption of capital punishment laws than before the adoption of capital punishment laws. The researchers concluded that the death penalty does act to deter murderers.

Critics of the study have complained that the crime rate in general dropped significantly in many of the states studied during the two-year period surrounding the legal change, and that they may be able to show that murder rates, when expressed as a percentage of all crimes, actually increased following the adoption of capital punishment (though they present no supporting data).

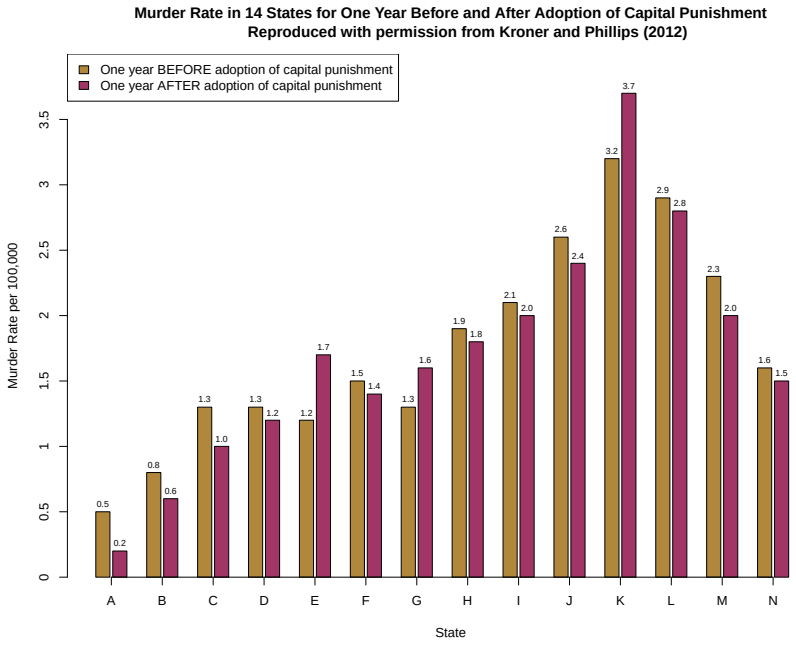
In addition, they pointed out that many other factors may have changed in addition to the adoption of the death penalty, drawing attention especially to the fact that several of the states studied had made parole more difficult for convicted murderers at the same time that they adopted or reinstated capital punishment laws, thus accounting for all or part of the changes.

Kroner and Phillips (2013) have recently replied to this last criticism by stating that several of the states included in their study had also adopted more lenient parole procedures, thus balancing the states that have become stricter on paroles.

Murder Rate in 14 States for One Year Before and After Adoption of Capital Punishment

State	Murder Rate	Year	State	Murder Rate	Year
A	0.5	Before	H	1.9	Before
	0.2	After		1.8	After
B	0.8	Before	I	2.1	Before
	0.6	After		2.0	After
C	1.3	Before	J	2.6	Before
	1.0	After		2.4	After
D	1.3	Before	K	3.2	Before
	1.2	After		3.7	After
E	1.2	Before	L	2.9	Before
	1.7	After		2.8	After
F	1.5	Before	M	2.3	Before
	1.4	After		2.0	After
G	1.3	Before	N	1.6	Before
	1.6	After		1.5	After

Table reproduced with permission from Kroner and Phillips (2012)



Con Time Series Details and Criticism

Can crime be prevented by the death penalty?

In recent years there has arisen a great public debate over whether or not capital punishment (the death penalty) is effective in preventing murders. Those who favor capital punishment claim that the death penalty acts as a deterrent by reminding the potential murderer of the possible consequences, whereas those who oppose capital punishment claim that such officially sanctioned killing by the state sets a violent example that will produce murders that might have been avoided. A recent research effort attempted to shed light on this continuing debate.

The researchers (Kroner and Phillips, 2012) noticed that within the past twelve years fourteen states have passed capital punishment laws or reinstated the death penalty. They reasoned that if capital punishment laws deter murders, the murder rate should have been higher for the year just before the adoption of the death penalty than for the year just after the adoption of death penalty in each of these states.

The results, shown in the table and graph below, were that in 11 of the 14 states the murder rate was **higher after** adoption of capital punishment laws than before the adoption of capital punishment laws. The researchers concluded that the death penalty does not act to deter murderers.

Critics of the study have complained that the crime rate in general dropped significantly in many of the states studied during the two-year period surrounding the legal change, and that they may be able to show that murder rates, when expressed as a percentage of all crimes, actually decreased following the adoption of capital punishment (though they present no supporting data).

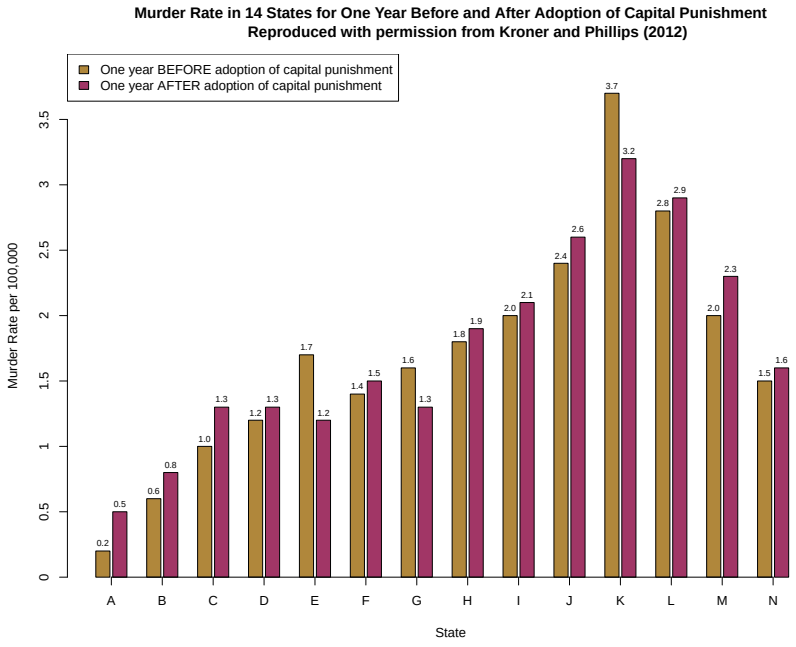
In addition, they pointed out that many other factors may have changed in addition to the adoption of the death penalty, drawing attention especially to the fact that several of the states studied had made parole more difficult for convicted murderers at the same time that they adopted or reinstated capital punishment laws, thus accounting for all or part of the changes.

Kroner and Phillips (2013) have recently replied to this last criticism by stating that several of the states included in their study had also adopted more lenient parole procedures, thus balancing the states that have become stricter on paroles.

Murder Rate in 14 States for One Year Before and After Adoption of Capital Punishment

State	Murder Rate	Year	State	Murder Rate	Year
A	0.2	Before	H	1.8	Before
	0.5	After		1.9	After
B	0.6	Before	I	2.0	Before
	0.8	After		2.1	After
C	1.0	Before	J	2.4	Before
	1.3	After		2.6	After
D	1.2	Before	K	3.7	Before
	1.3	After		3.2	After
E	1.7	Before	L	2.8	Before
	1.2	After		2.9	After
F	1.4	Before	M	2.0	Before
	1.5	After		2.3	After
G	1.6	Before	N	1.5	Before
	1.3	After		1.6	After

Table reproduced with permission from Kroner and Phillips (2012)



Null Time Series Details and Criticism

Can crime be prevented by the death penalty?

In recent years there has arisen a great public debate over whether or not capital punishment (the death penalty) is effective in preventing murders. Those who favor capital punishment claim that the death penalty acts as a deterrent by reminding the potential murderer of the possible consequences, whereas those who oppose capital punishment claim that such officially sanctioned killing by the state sets a violent example that will produce murders that might have been avoided. A recent research effort attempted to shed light on this continuing debate.

The researchers (Kroner and Phillips, 2012) noticed that within the past twelve years fourteen states have passed capital punishment laws or reinstated the death penalty. They reasoned that if capital punishment laws deter murders, the murder rate should have been higher for the year just before the adoption of the death penalty than for the year just after the adoption of death penalty in each of these states.

The results, shown in the table and graph below, were inconclusive: in 5 of the 14 states the murder rate was **lower after** adoption of capital punishment laws. Another 5 of the 14 states showed the **opposite pattern**. In the remaining states, there was **no change** in the murder rate before and after the adoption of capital punishment laws. The researchers concluded that the death penalty has an indeterminate effect on deterrence.

Critics of the study have complained that the crime rate in general dropped significantly in many of the states studied during the two-year period surrounding the legal change, and that they may be able to show that murder rates, when expressed as a percentage of all crimes, actually do respond to the adoption of capital punishment (though they present no supporting data).

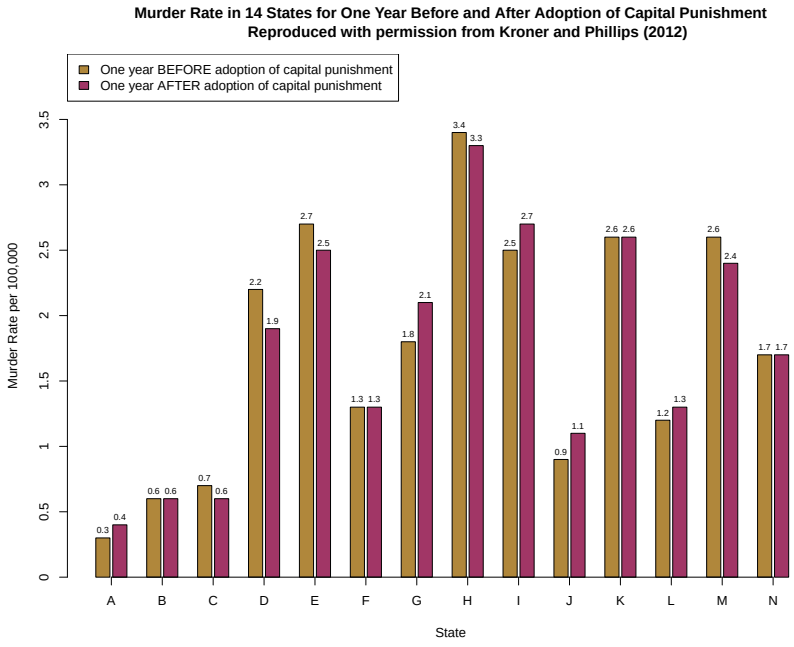
In addition, they pointed out that many other factors may have changed in addition to the adoption of the death penalty, drawing attention especially to the fact that several of the states studied had made parole more difficult for convicted murderers at the same time that they adopted or reinstated capital punishment laws, thus accounting for all or part of the changes.

Kroner and Phillips (2013) have recently replied to this last criticism by stating that several of the states included in their study had also adopted more lenient parole procedures, thus balancing the states that have become stricter on paroles.

Murder Rate in 14 States for One Year Before and After Adoption of Capital Punishment

State	Murder Rate	Year	State	Murder Rate	Year
A	1.7	Before	H	2.9	Before
	0.4	After		3.3	After
B	0.6	Before	I	2.5	Before
	0.6	After		2.7	After
C	0.7	Before	J	0.9	Before
	0.6	After		1.1	After
D	2.2	Before	K	2.6	Before
	1.9	After		2.6	After
E	2.7	Before	L	1.2	Before
	2.5	After		1.3	After
F	1.3	Before	M	2.6	Before
	1.3	After		2.4	After
G	1.8	Before	N	1.7	Before
	2.1	After		1.7	After

Table reproduced with permission from Kroner and Phillips (2012)



Pro Cross Section Details and Criticism

Does Capital Punishment Prevent Crime?

One of the most controversial public issues in recent years has been the effectiveness of capital punishment (the death penalty) in preventing murders. Proponents of capital punishment have argued that the possibility of execution deters people who might otherwise commit murders, whereas opponents of capital punishment denied this and maintain that the death penalty may even produce murders by setting a violent model of behavior. A recent research effort attempted to shed light on this controversy.

The researchers (Palmer and Crandall, 2012) decided to look at the difference in murder rates in states that share a common border but differ in whether their laws permit capital punishment or not. Carefully limiting the states included to those which had capital punishment laws in effect or not in effect for at least five years, they compiled a list of all possible pairs and then selected ten pairs of neighboring states that were alike in the degree of urbanization (percentage of the population living in metropolitan areas), thus controlling for any relationship between the size of urban population and crime per capita. They also limited the capital punishment states to those which had actually used their death penalty statutes, thus controlling for the possibility that the mere existence of the death penalty may not carry the same weight unless capital punishment is known to be a possibility. Using the murder rate (number of willful homicides per 100,000 population) in 2010 as their index, they assembled the table and graph shown on the next page. They reasoned that if capital punishment has a deterrent effect, the murder rates should be lower in the state with capital punishment laws.

The results, as shown in the table and graph below, were that in eight of the ten pairs of states selected for their study the murder rates were **lower** in the state with capital punishment laws than in the state without capital punishment laws. The researchers concluded that the existence of the death penalty does work to deter murderers.

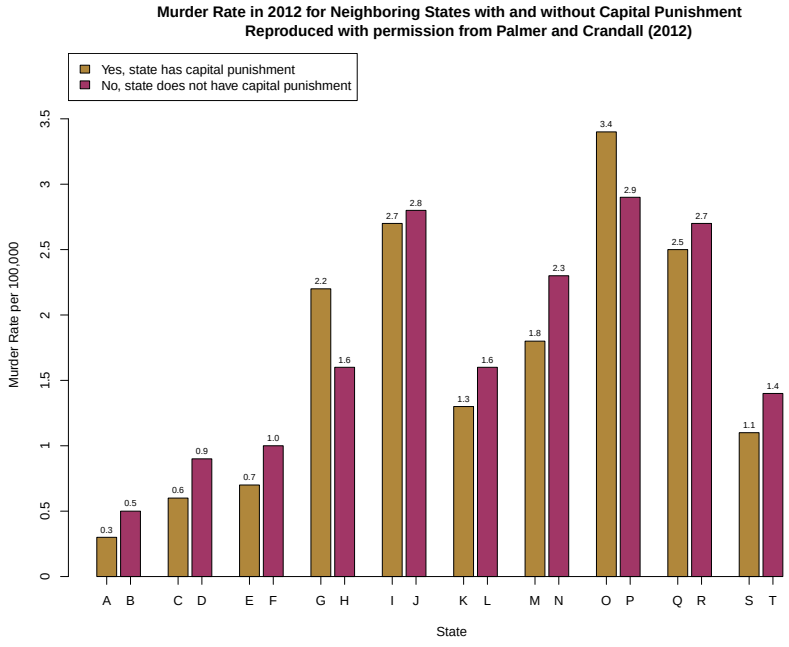
Critics of the study have complained that selection of a different set of ten neighboring states might have yielded a far different, perhaps even the opposite, result.

In replying to this criticism, Palmer and Crandall (2013) have recently reported a replication of their study, using a different set of ten states that share a common border but differ in whether their laws permit capital punishment or not. The results of this second study were essentially the same, murder rates being lower in the capital punishment state for seven of the ten comparisons.

Murder Rate in 2012 for Neighboring States With and Without Capital Punishment

Pair	State	Murder Rate	Capital Punishment	Pair	State	Murder Rate	Capital Punishment
1	A	0.3	Yes	6	K	1.3	Yes
	B	0.5	No		L	1.6	No
2	C	0.6	Yes	7	M	1.8	Yes
	D	0.9	No		N	2.3	No
3	E	0.7	Yes	8	O	3.4	Yes
	F	1.0	No		P	2.9	No
4	G	2.2	Yes	9	Q	2.5	Yes
	H	1.6	No		R	2.7	No
5	I	2.7	Yes	10	S	1.1	Yes
	J	2.8	No		T	1.4	No

Table reproduced with permission from Palmer and Crandall (2012)



Con Cross Section Details and Criticism

Does Capital Punishment Prevent Crime?

One of the most controversial public issues in recent years has been the effectiveness of capital punishment (the death penalty) in preventing murders. Proponents of capital punishment have argued that the possibility of execution deters people who might otherwise commit murders, whereas opponents of capital punishment denied this and maintain that the death penalty may even produce murders by setting a violent model of behavior. A recent research effort attempted to shed light on this controversy.

The researchers (Palmer and Crandall, 2012) decided to look at the difference in murder rates in states that share a common border but differ in whether their laws permit capital punishment or not. Carefully limiting the states included to those which had capital punishment laws in effect or not in effect for at least five years, they compiled a list of all possible pairs and then selected ten pairs of neighboring states that were alike in the degree of urbanization (percentage of the population living in metropolitan areas), thus controlling for any relationship between the size of urban population and crime per capita. They also limited the capital punishment states to those which had actually used their death penalty statutes, thus controlling for the possibility that the mere existence of the death penalty may not carry the same weight unless capital punishment is known to be a possibility. Using the murder rate (number of willful homicides per 100,000 population) in 2010 as their index, they assembled the table and graph shown on the next page. They reasoned that if capital punishment has a deterrent effect, the murder rates should be lower in the state with capital punishment laws.

The results, as shown in the table and graph below, were that in eight of the ten pairs of states selected for their study the murder rates were **higher** in the state with capital punishment laws than in the state without capital punishment laws. The researchers concluded that the existence of the death penalty does not work to deter murderers.

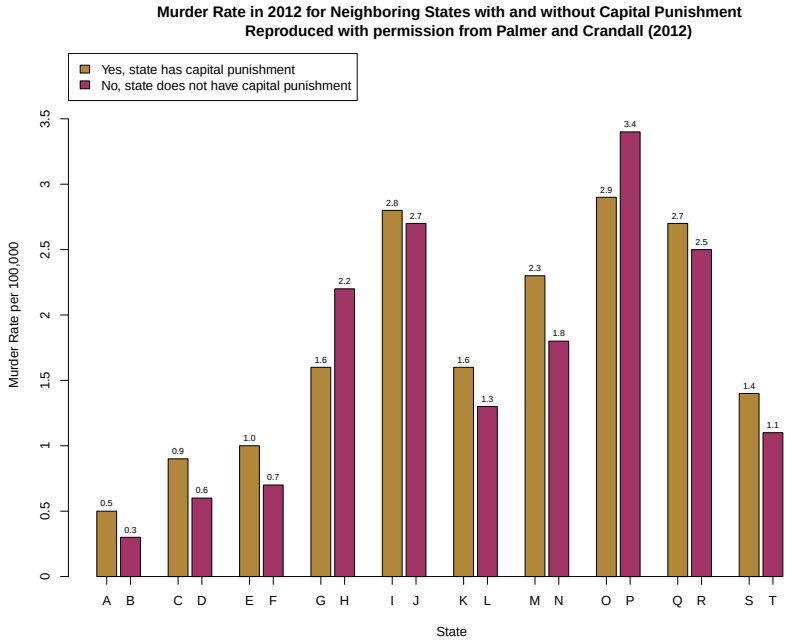
Critics of the study have complained that selection of a different set of ten neighboring states might have yielded a far different, perhaps even the opposite, result.

In replying to this criticism, Palmer and Crandall (2013) have recently reported a replication of their study, using a different set of ten states that share a common border but differ in whether their laws permit capital punishment or not. The results of this second study were essentially the same, murder rates being higher in the capital punishment state for seven of the ten comparisons.

Murder Rate in 2012 for Neighboring States With and Without Capital Punishment

Pair	State	Murder Rate	Capital Punishment	Pair	State	Murder Rate	Capital Punishment
1	A	0.5	Yes	6	K	1.6	Yes
	B	0.3	No		L	1.3	No
2	C	0.9	Yes	7	M	2.3	Yes
	D	0.6	No		N	1.8	No
3	E	1.0	Yes	8	O	2.9	Yes
	F	0.7	No		P	3.4	No
4	G	1.6	Yes	9	Q	2.7	Yes
	H	2.2	No		R	2.5	No
5	I	2.8	Yes	10	S	1.4	Yes
	J	2.7	No		T	1.1	No

Table reproduced with permission from Palmer and Crandall (2012)



Null Cross Section Details and Criticism

Does Capital Punishment Prevent Crime?

One of the most controversial public issues in recent years has been the effectiveness of capital punishment (the death penalty) in preventing murders. Proponents of capital punishment have argued that the possibility of execution deters people who might otherwise commit murders, whereas opponents of capital punishment denied this and maintain that the death penalty may even produce murders by setting a violent model of behavior. A recent research effort attempted to shed light on this controversy.

The researchers (Palmer and Crandall, 2012) decided to look at the difference in murder rates in states that share a common border but differ in whether their laws permit capital punishment or not. Carefully limiting the states included to those which had capital punishment laws in effect or not in effect for at least five years, they compiled a list of all possible pairs and then selected ten pairs of neighboring states that were alike in the degree of urbanization (percentage of the population living in metropolitan areas), thus controlling for any relationship between the size of urban population and crime per capita. They also limited the capital punishment states to those which had actually used their death penalty statutes, thus controlling for the possibility that the mere existence of the death penalty may not carry the same weight unless capital punishment is known to be a possibility. Using the murder rate (number of willful homicides per 100,000 population) in 2010 as their index, they assembled the table and graph shown on the next page. They reasoned that if capital punishment has a deterrent effect, the murder rates should be lower in the state with capital punishment laws.

The results, as shown in the table and graph below, were that in four of the ten pairs of states selected for their study the murder rates were **lower** in the state with capital punishment laws than in the state without capital punishment laws. In another four pairs, the **opposite** was true. In two of the ten pairs, the states had the **same** murder rates. The researchers concluded that the existence of the death penalty has an indeterminate deterrent effect on murderers.

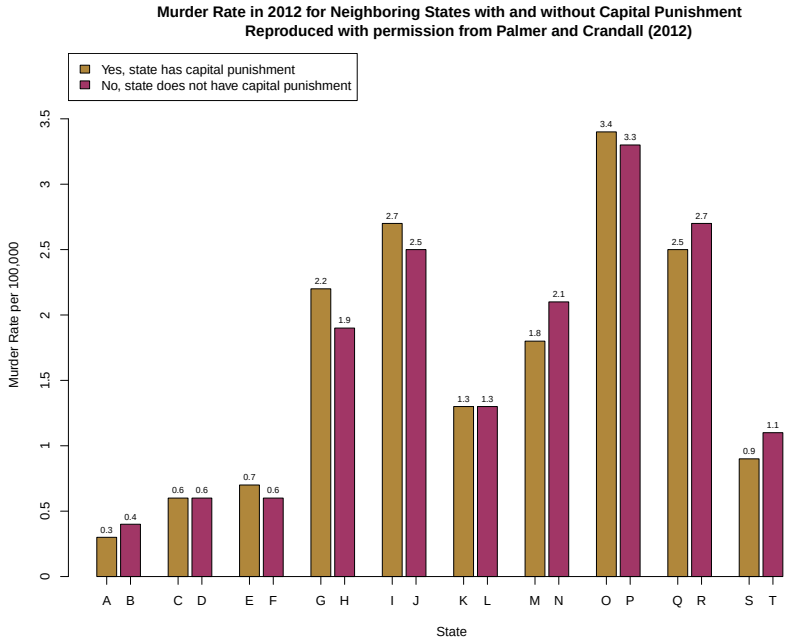
Critics of the study have complained that selection of a different set of ten neighboring states might have yielded a far different result.

In replying to this criticism, Palmer and Crandall (2013) have recently reported a replication of their study, using a different set of ten states that share a common border but differ in whether their laws permit capital punishment or not. The results of this second study were essentially the same, murder rates being lower in the capital punishment state for some comparisons and higher in others.

Murder Rate in 2012 for Neighboring States With and Without Capital Punishment

Pair	State	Murder Rate	Capital Punishment	Pair	State	Murder Rate	Capital Punishment
1	A	0.3	Yes	6	K	1.3	Yes
	B	0.4	No		L	1.3	No
2	C	0.6	Yes	7	M	1.8	Yes
	D	0.6	No		N	2.1	No
3	E	0.7	Yes	8	O	3.4	Yes
	F	0.6	No		P	3.3	No
4	G	2.2	Yes	9	Q	2.5	Yes
	H	1.9	No		R	2.7	No
5	I	2.7	Yes	10	S	1.1	Yes
	J	2.5	No		T	0.9	No

Table reproduced with permission from Palmer and Crandall (2012)



Outcomes:

- **Attitude**: Which view of capital punishment best summarizes your own? (1: I am very much against capital punishment; 7: I am very much in favor of capital punishment)
- **Belief**: Does capital punishment reduce crime? Please select the view that best summarizes your own. (1: I am very certain that capital punishment does not reduce crime; 7: I am very certain that capital punishment reduces crime.)

A.1.3 Results

Tables A.2 and A.3 show the effects of information content on the Attitude and Belief dependent variables, respectively. The effect on attitude is estimated to be 0.08 scale points for a unit change in information content; this effect is estimated with far greater precision when covariates are included. The effect on belief is very strong, at 0.24 scale points for a unit change, which implies that the effect of going from the *Con Con* condition to the *Pro Pro* condition is nearly a full scale point. Both effects are strongly persistent after 10 days.

Table A.2: Effect of Research Reports on Support for Capital Punishment

	Attitude Toward Capitial Punishment			
	Wave 2		Wave 3	
Information Content (-2 to 2)	0.080	0.108***	0.100	0.102***
	(0.065)	(0.025)	(0.068)	(0.025)
Condition: Null Null	−0.335	−0.069	−0.387*	−0.051
	(0.215)	(0.072)	(0.224)	(0.073)
Constant	3.514	0.273	3.598	0.223
	(0.092)	(0.184)	(0.096)	(0.203)
Covariates	No	Yes	No	Yes
N	686	683	628	628
R^2	0.005	0.871	0.008	0.883

*p < .1; **p < .05; ***p < .01
Robust standard errors are in parentheses.
The information content of the Null Null condition is coded 0.
Covariates: T1 Attitude, T1 Belief, age, gender, ideology, race, and education.

Table A.3: Effect of Research Reports on Belief in Deterrent Effect

	Belief in Deterrent Effect			
	Wave 2		Wave 3	
Information Content (-2 to 2)	0.239^{***}	0.258^{***}	0.255^{***}	0.249^{***}
	(0.049)	(0.032)	(0.051)	(0.030)
Condition: Null Null	-0.467^{***}	-0.286^{***}	-0.539^{***}	-0.294^{***}
	(0.163)	(0.090)	(0.171)	(0.086)
Constant	3.601	1.486	3.655	1.277
	(0.069)	(0.233)	(0.071)	(0.245)
Covariates	No	Yes	No	Yes
N	686	683	628	628
R^2	0.043	0.645	0.052	0.722

$^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$
Robust standard errors are in parentheses.
The information content of the Null Null condition is coded 0.
Covariates: T1 Attitude, T1 Belief, age, gender, ideology, race, and education.

A.2 Study 2: Minimum Wage

This study was conducted in collaboration with Andrew Guess. It showed that viewing pro or con videos about the minimum wage moved opinions about raising the minimum wage in the direction of information.

Sample: 1,170 Mechanical Turk subjects in wave 1, 1,019 in wave 2.

Dates of data collection: August 2014.

Treatments: Subjects were assigned to see 2 of 6 videos. A placebo group was assigned to see both placebo videos. The treatment groups were assigned to see any two of the four treatments videos – all orderings of two videos were possible, as shown in the table below.

Table A.4: Minimum Wage Treatment Conditions

	1st Video		2nd Video				
	Content	Tone	Content	Tone	Condition	Info Content	N
1	Con	Old	Con	Young	Con Con	-1	79
2	Con	Young	Con	Old	Con Con	-1	83
3	Con	Old	Pro	Old	Con Pro	0	86
4	Con	Old	Pro	Young	Con Pro	0	91
5	Con	Young	Pro	Old	Con Pro	0	93
6	Con	Young	Pro	Young	Con Pro	0	99
7	Pro	Old	Con	Old	Pro Con	0	79
8	Pro	Old	Con	Young	Pro Con	0	102
9	Pro	Young	Con	Old	Pro Con	0	78
10	Pro	Young	Con	Young	Pro Con	0	93
11	Pro	Old	Pro	Young	Pro Pro	1	90
12	Pro	Young	Pro	Old	Pro Pro	1	104
13	Placebo	Placebo	Placebo	Placebo	Placebo	NA	93

Pro Young

Title: Should We Raise the Minimum Wage?

Presenter: John Green

Length: 3:46

URL: `http://youtu.be/ZI9aDHLptMk`

Transcript:
Good morning Hank, it's Tuesday. So you've started a lot of businesses: Crash Course, Scishow, DFTBA Records, Vidcon, the ceaseless juggernaut that is 2D glasses. And Hank your companies employ dozens of people, none of whom work for the federally-mandated minimum wage of $7.25 per hour.

But Hank, let's imagine that your next project is a fast food restaurant: Corndogs and Sodium. What impact would raising the federal minimum wage have on you and your employees?

At first glance it seems like a no-brainer. Any minimum wage is terrible, both for Corndogs and Sodium and for its employees. The Econ 101 argument goes like this. The free market is going to set wages where they need to be. Like, if you want to pay $5.00 an hour for Corndogs and Sodium employees but no one takes the job for $5.00 an hour, you're gonna have to pay more. You'll increase your wages until you can attract the kind of employees that you need to, you know, batter and fry and serve encased cast-off pig meat. And we know that economies tend to grow less when governments set and control prices, so higher minimum wages restrict economic growth. Plus, unemployment will go up, because if the minimum wage is $10.00 per hour, Corndogs and Sodium can only afford to hire one person. But if there was an unrestricted wage market, then they could attract two people who would be willing to work for $5.00 an hour each. So in the end, setting a minimum wage is an attempt to alleviate poverty that actually increases it.

However Hank, surprisingly enough, it turns out that actual labor markets are a lot more complex than the models of labor markets created by college freshmen. This brings us to a famous study by two economists: David Card and Alan Krueger. So in 1992, the state of New Jersey raised its minimum wage 18.8%. Pennsylvania, right next door, did not raise its minimum wage. Card and Krueger had the bright idea to go to the border of New Jersey and Pennsylvania and do employment surveys on either side of it. And what they found is that restaurant employment in New Jersey actually increased when the minimum wage went up. Since then a bunch of other studies have confirmed Card and Krueger's findings, while some have found that there actually are negative effects to unemployment when you raise the minimum wage though it's surprisingly and consistently mild.

Why? Well, a bunch of reasons. For one, the minimum wage is probably near where the market would set it. But also, low-wage workers tend to spend most of their pay raises, which leads to increased economic activity, which in turn leads to more jobs. And higher wages also means less turnover, which leads to lower costs of training and hiring and firing. On the downside, higher wages are also associated with higher prices on goods and services that rely on low-wage labor which means that your corndogs, Hank, would probably be a bit more expensive.

But Hank the larger question is whether raising the minimum wage actually reduces poverty. And on that front there is consensus that – at least in the medium run – it does. A number of recent studies have shown that raising the minimum wage 10% reduces the number of people in poverty by about 2.5%. Even many opponents of the minimum wage acknowledge this. But it's important to note that, like, that won't always work. At some point, raising the minimum wage will lead to inflation and slower job creation. It's just not clear where that point is. But it's just as disingenuous to call the minimum wage a job killer as it is to say that the minimum wage is going to fix economic inequality. In short Hank, in economics, there's no such thing as a free lunch. But when it comes to reducing poverty without affecting employment, higher minimum wages seem at least to be the cheapest lunch available.

But ultimately, Hank, now that I'm, I guess, an employer, I'm more persuaded by the personal argument. We've found that paying a living wage, which we would do even if we opened Corndogs and Sodium, leads to happier more productive employees. Now I know that's hard to quantify, but it's also what's allowed Vidcon and DFBTA Records to retain employees for years and years and grow sustainably.

Now Hank, obviously I am not an economist (although I did win a bronze medal in economics at the Alabama state academic decathlon tournament in 1993). But our strategy has worked out pretty well for us so far, and it's also working at much larger companies like CostCo.

Hank, the United States is a rich country and I think that there's a growing body of evidence that the U.S. doesn't benefit from having poor workers. Of course raising the minimum wage isn't going to fix that problem, but I hope at least we can begin to have a nuanced conversation about the problem.

Hank – I'll see you on Friday.

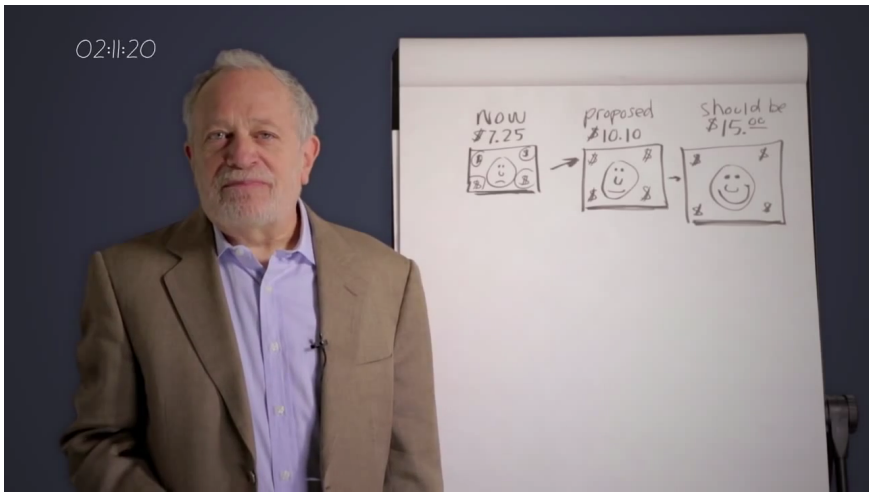
Pro Old

Title: Raise the Minimum Wage to $15/hr (in 2 minutes 30 seconds)

Presenter: Robert B. Reich

Length: 2:41

URL: `http://youtu.be/GOqtl53V3JI`



Transcript:
Democrats are getting ready for a major push to raise the federal minimum wage to $10.10 an hour. Now that's better than nothing. But it is not enough. The federal minimum wage should be raised to $15.00 an hour, incrementally, over the next three years.

Here's seven reasons why.

One. Had the minimum wage of 1968 been adjusted for inflation it would be well above $10.00 an hour today. A typical worker today is also more than twice as productive as back then. Adjusted for inflation and productivity gains therefore, the minimum wage should be at least $15.00 an hour.

Two. $10.10 an hour is not enough to lift all workers and their families out of poverty. This is especially true for millions of low-wage workers who want full time jobs but can only work and find part time work. Most of the workers are not teenagers. They are major breadwinners for their families.

Three. Because some employers don't pay wages that lift their workers out of poverty, the rest of us pay for their Medicaid, food stamps, housing, and other assistance. In effect, subsidizing these low-wage employers. Some, like McDonald's, actually advise their employees to use public programs because their pay is just too low.

Four. Some jobs may be lost if the minimum is raised to $15.00. But many more people will be lifted out of poverty. And because low-wage workers will have more money to spend, their spending will create many more jobs.

Five. Such a wage increase is more likely to come out of profits than be passed on in higher prices, because most employers of low wage workers face intense competition for customers.

Six. Since Republicans will no doubt try to push Democrats to go even lower than their $10.10 proposal, it's doubly important to be clear about what's right in the first place. Democrats should be talking about a bigger increase, not listening to Republican demands for a smaller one.

Seven, and finally. At a time in our nation's history when 95% of all economic gains are going to the top 1%, raising the minimum wage to $15.00 isn't just smart economics, it's the right thing to do.

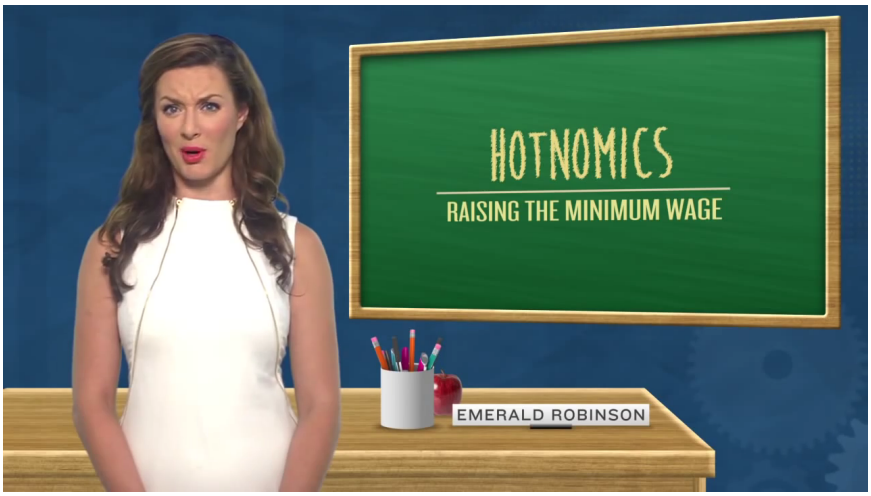
Con Young

Title: Hotnomics – Minimum Wage

Presenter: Emerald Robinson

Length: 2:32

URL: `http://youtu.be/hFG1Ka8AW6Q`



Transcript:
More money more happiness? Or more money more problems? We're cashing in on the effects of raising the minimum wage in this episode of Hotnomics. Raising the minimum wage is a hot topic right now, with President Obama calling on Congress to raise the minimum wage from $7.25 an hour to $10.10 an hour. That's a 40% hike! Who wouldn't take more money, right?

But when it comes to raising the federal minimum wage, it might be a little different than you think. You have to look at the bigger overall picture to really understand what it would mean for workers in our country. Workers in entry level jobs will earn a little more – that is if they even get to keep their job. When asked if they would cut hiring if the federal minimum wage was raised to $10 an hour, the majority of chief financial officers across the country said "Yes."

A recent report from the Congressional Budget Office predicts that 500,000 Americans will lose their job by 2016 if the federal minimum wage is raised is to $10.10 an hour. 500,000 people! That's a lot of out of work teens, Mom and Dad.

Just over half of the minimum wage earners are 16 to 24 years old and tend to live in middle-class households with a family income of over $65,000 dollars per year. Way above the poverty line. In fact, the CBO also estimates that only 20% of minimum wage earners fall below the poverty line. And with raising the federal minimum wage, the skill level that employers expect their employees to have also rises. So the youth from low

income neighborhoods and disadvantaged adults that are supposed to benefit from a higher minimum wage actually get pushed out of jobs.

But you say, "I don't work in an entry-level job and I don't have teenage kids, so what do I care?" Well you might care the next time you go to get that $1.50 hot dog special at your favorite joint on the corner, when it's now $2.00 a hot dog. Yep. Higher minimum wages means that business owners will be forced to up their prices to compensate for the increase in employee wages.

And what about the differences in the cost of living across the country? $10.10 means one thing in Manhattan, as opposed to Birmingham, Alabama.

A recent letter to federal policy makers signed by 500 economists including several Nobel prize winners said that raising the minimum wage is not a silver bullet solution to poverty. Instead these economists support finding solutions that encourage employment, business creation, and boost earnings, rather than across the board mandates that raise the cost of labor.

So in this case, raising the federal minimum wage means more problems for the economy as a whole.

To learn more about how the economic machine works, keep tuning in to Hotnomics. Not your typical economics class.

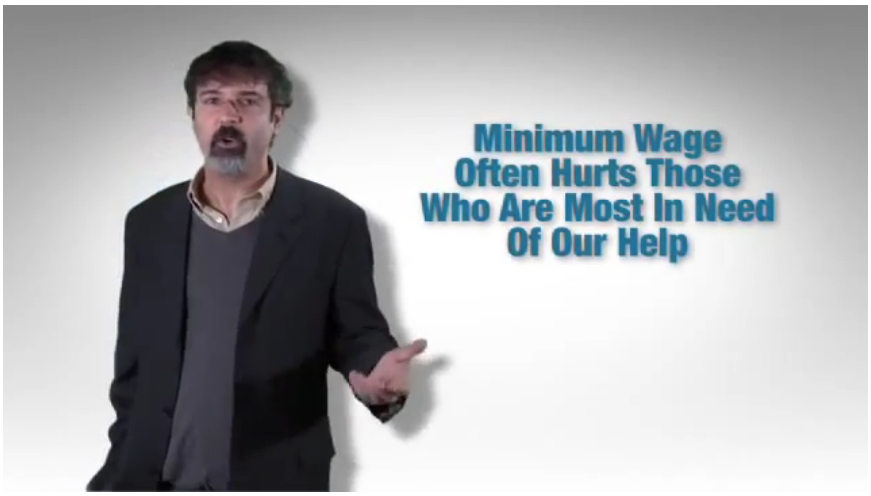
Con Old

Title: Does the Minimum Wage Hurt Workers?

Presenter: Antony Davies

Length: 4:28

URL: `http://youtu.be/Ct1Moeaa-W8`



Transcript:
Some politicians argue that raising the minimum wage helps the poor and disadvantaged. It might seem that way at first. Certainly workers who are earning $8.00 an hour today would be better off if they were earning $12.00 an hour instead. The problem is that this view of the minimum wage overlooks one important detail. A $12.00 minimum wage doesn't force employers to pay $12.00 to every worker. It forces them to pay $12.00 to the workers the employer chooses to keep. The employer pays $0.00 to the workers who get laid off, or who were never hired in the first place.

Let's look at an example. Suppose this guy owns a burger joint. The reason the employer hires a worker is because the worker generates value for the owner. Suppose that, not counting what he pays his worker, the owner makes 10 cents on every burger he sells.

Here's Al. Al can flip 100 burgers an hour. If the owner makes 10 cents on every burger, not counting what he pays Al, then Al generates $10.00 worth of burgers for the owner every hour. If the owner pays Al $8.00 an hour, then the owner makes $2.00 an hour profit: $10.00 an hour on the burgers, minus $8.00 an hour that he pays Al.

Now, suppose that the owner hires three workers of varying abilities: Al, Bob, and Carl. Bob is a faster worker, and can cook 120 burgers an hour. Carl is a slower worker, and can only cook 90 burgers an hour. Since each burger is worth 10 cents to the owner, each hour Al produces $10.00 worth of burgers, Bob produces $12.00, and Carl produces $9.00. Suppose the owner pays Al, Bob, and Carl $8.00 an hour each. After paying the workers, the owner earns $2.00 an hour profit from employing Al, $4.00 an hour profit from

employing Bob, and $1.00 an hour profit from employing Carl. In total, the owner makes $7.00 profit per hour from employing the three workers.

Now, suppose the government imposes a minimum wage of $9.50 an hour. What does this do to the profit of our three workers? Al produces $10.00 worth of burgers per hour. At a cost of $9.50 an hour, Al now generates a profit of only 50 cents an hour for the owner. Bob produces $12.00 worth of burgers an hour. At a cost of $9.50 an hour, Bob now generates a profit of $2.50 an hour. But look at what happens to Carl. Carl produces $9.00 worth of burgers per hour, but he now costs the owner $9.50 per hour in wages. Carl is no longer generating profit for the owner. Carl now generates a loss. In fact, the owner would be 50 cents an hour better off if he fires Carl.

The minimum wage was good for Al and Bob. They're each a $1.50 an hour better off than they were before. But it was devastating for Carl. Carl lost his job, and so is $8.00 an hour worse off than he was before.

Here's the first lesson of the minimum wage. It doesn't help the worker at the expense of the owner. It helps the more productive workers at the expense of the less productive workers. What's even worse is that the more productive workers usually don't need the help. What do you think would have happened over time to Bob, the most productive worker? Either the owner would have rewarded Bob's higher productivity with a raise, or, if the owner didn't reward Bob, one of the owner's competitors would have offered Bob more money to go work for the competitor. Either way, Bob would have ended up earning more anyway.

This is the second lesson of the minimum wage. Many of the workers that it does help would have ended up better off anyway, even if the minimum wage hadn't existed. It works this way in the real world. Increases in the minimum wage have little effect on unemployment among college graduates. Increases in the minimum wage increase unemployment among high school graduates. And among the least skilled, least educated workers, increases in the minimum wage significantly increase unemployment.

The minimum wage may be a well-intentioned policy. But it often hurts the very workers who are in most need of our help.

Placebo 1

Title: Are Workers Who Receive Tips Eligible for Minimum Wage?

Presenter: Daniel Swanson

Length: 0:28

URL: `http://youtu.be/sRMvH2L4Dss`

Transcript:
Workers who receive tips, like waitresses or waiters, are entitled to the minimum wage, but it works a little different for those folks. The employer is required to pay them at least $2.13 an hour, but their tips must equal the amount of the minimum wage, which is $7.25 an hour. If their total compensation does not equal $7.25 an hour, the employer must pay the difference.

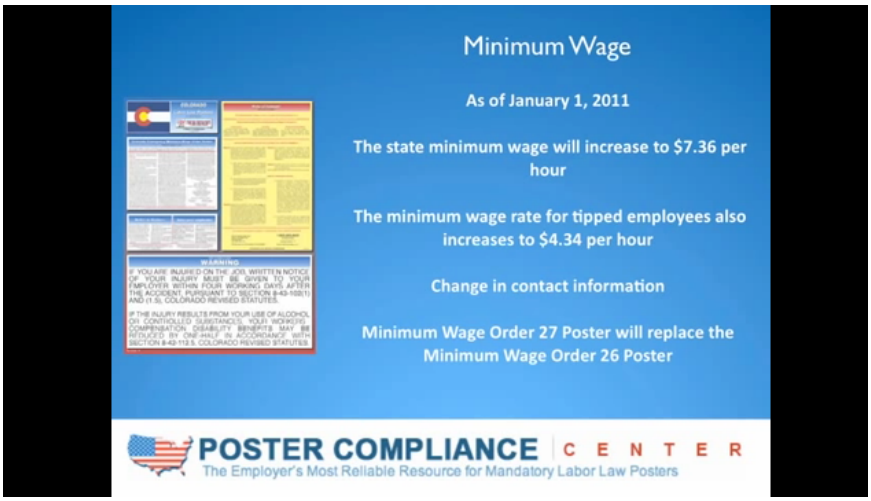
Placebo 2

Title: Labor Law Posters: Colorado Minimum Wage Labor Law Poster

Presenter: Poster Compliance Center

Length: 0:28

URL: `http://youtu.be/xd7RrMW8MwE`



Transcript:
Poster Compliance Center would like to welcome you to our labor law quick update for Colorado. Minimum wage regulations are changing and we would like to inform you of these important changes. As of January 1st, 2011, the state minimum wage will increase to $7.36 per hour. The minimum wage rate for tipped employees also increases to $4.34 per hour. Changes have been made to the contact information listed. The minimum wage order 27 poster will replace the minimum wage order 26 poster. Poster Compliance Center would like to thank you for joining us. We look forward to serving you and all of your labor law needs.

Outcomes

- **Favor Raising**: The federal minimum wage is currently $7.25 per hour. Do you favor or oppose raising the federal minimum wage? (1: Very much opposed to raising the federal minimum wage, 7: Very much in favor of raising the federal minimum wage)
- **Amount**: What do you think the federal minimum wage should be? Please enter an amount between $0.00 and $25.00 in the text box below.

A.2.1 Results

As shown in tables A.5 and A.6, the videos exerted a strong influence on both the Favor Raising and Amount dependent variables. Effects on both dependent variables are strongly persistent, as shown by the third and fourth columns of each table.

Table A.5: Effects of Videos on Favoring Raising the Minimum Wage

	Favoring Raising the Minimum Wage			
	Wave 2		Wave 3	
Information Content (-1 to 1)	0.424***	0.506***	0.314***	0.373***
	(0.093)	(0.064)	(0.102)	(0.063)
Condition: Placebo	0.701***	0.686***	0.502**	0.505***
	(0.205)	(0.132)	(0.228)	(0.139)
Constant	5.067	1.629	4.989	1.172
	(0.056)	(0.258)	(0.061)	(0.294)
Covariates	No	Yes	No	Yes
N	1,170	1,170	1,020	1,020
R^2	0.018	0.638	0.009	0.638

*p < .1; **p < .05; ***p < .01
Robust standard errors are in parentheses.
The information content of the Placebo condition is coded 0.
Covariates: T1 Attitude, T1 Belief, age, gender, ideology, race, and education.

Table A.6: Effects of Videos on Favoring Raising the Minimum Wage

	Amount			
	Wave 2		Wave 3	
Information Content (-1 to 1)	1.138***	1.177***	0.627***	0.636***
	(0.178)	(0.118)	(0.195)	(0.125)
Condition: Placebo	0.687**	0.697***	0.333	0.378**
	(0.345)	(0.163)	(0.378)	(0.187)
Constant	9.685	2.399	9.782	1.791
	(0.103)	(0.459)	(0.115)	(0.409)
Covariates	No	Yes	No	Yes
N	1,170	1,170	1,019	1,019
R^2	0.035	0.674	0.010	0.727

*p < .1; **p < .05; ***p < .01
Robust standard errors are in parentheses.
The information content of the Placebo condition is coded 0.
Covariates: T1 Attitude, T1 Belief, age, gender, ideology, race, and education.

A.3 Study 3: Global Warming

This study was conducted in collaboration with Andrew Guess. It showed that presenting information about a global warming "hiatus" changed whether subjects think global warming is due to natural causes but did not change whether subjects viewed it as a threat to their way of life.

Sample: 1,690 Mechanical Turk subjects.

Dates of data collection: October 2014.

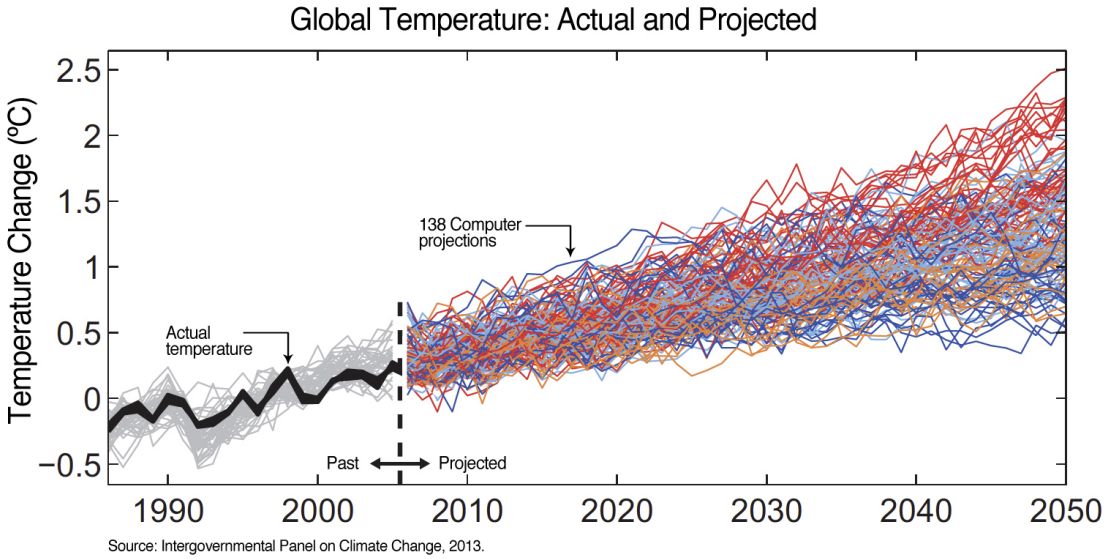
Treatments: Subjects could be assigned to one of two conditions: a control condition and a "hiatus" condition.

Control Condition:

Global warming refers to the increase in the Earth's temperature over the last century. Since 1900, temperatures have increased by 1.4 degrees Fahrenheit. Scientists, politicians, and citizens have been debating over what, if anything, to do to combat global warming.

In order to understand what to do about global warming, we need to know more than how much the Earth has warmed in the past. We need to know how much the Earth will warm in the future.

Climate scientists try to predict how much warming will occur by training statistical models on historical data, then running those models forward. Because any one model can be wrong, the UN's Intergovernmental Panel on Climate Change (IPCC) averages many forecasts, as shown in the figure below.



The colored lines show the models' best guesses for the changes in temperature. The black line shows the actual observed temperature.

There is a great deal of agreement between the models and the observed temperature. This agreement means that we can have greater confidence in the models' projections into the future.

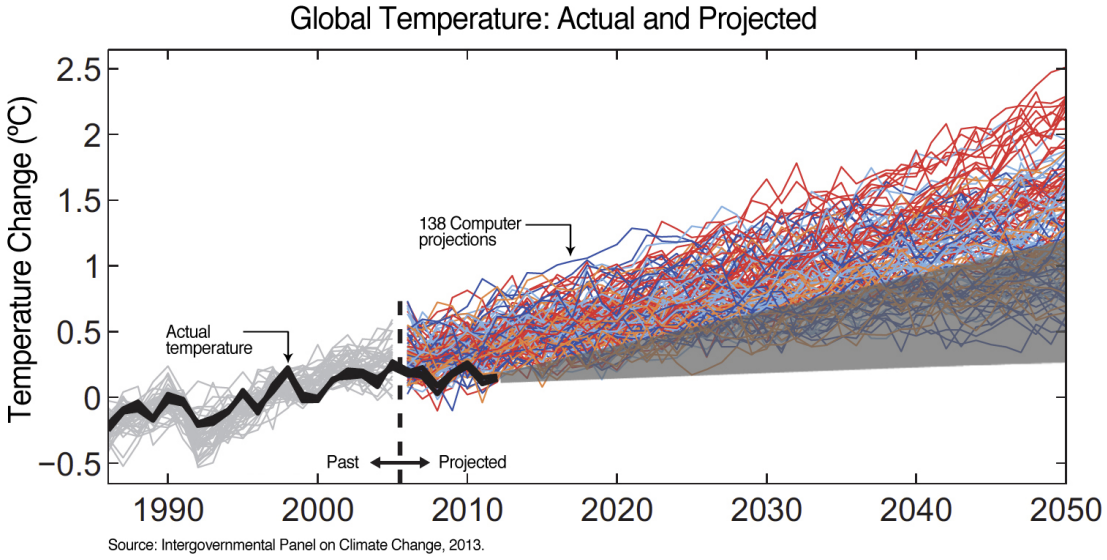
The Earth is warming, and the models predict that the warming trends will continue into the future.

Hiatus Treatment Condition

Global warming refers to the increase in the Earth's temperature over the last century. Since 1900, temperatures have increased by 1.4 degrees Fahrenheit. Scientists, politicians, and citizens have been debating over what, if anything, to do to combat global warming.

In order to understand what to do about global warming, we need to know more than how much the Earth has warmed in the past. We need to know how much the Earth will warm in the future.

Climate scientists try to predict how much warming will occur by training statistical models on historical data, then running those models forward. Because any one model can be wrong, the UN's Intergovernmental Panel on Climate Change (IPCC) averages many forecasts, as shown in the figure below.



The colored lines show the models' best guesses for the changes in temperature. The solid black line shows the actual observed temperature.

The dotted line represents the moment in time when the predictions were made. To the left of the dotted line, the models and the observed temperature agree. But to the right of the dotted line, the actual levels of warming clearly disagree with the predictions of the models.

This disagreement means we cannot be confident of the models' projections into the future. The shaded region describes the updated projections based on more recent data.

Outcomes:

- **Temp Cause**: Some people believe that increases in the Earth's temperature over the last century are due to the effects of pollution from human activities. Others believe that temperature increases are due to natural changes in the environment. Which comes closer to your view? (Temperature increases are mainly due to the effects of pollution from human activities, Temperature increases are mainly due to natural changes in the environment, Temperature increases are due to both human activities and natural changes, I do not believe the Earth's temperature has risen.) The outcome **Natural Cause** is coded 1 if the respondent chose "Temperature increases are mainly due to natural changes in the environment" or "Temperature increases are due to both human activities and natural changes" and 0 otherwise.
- **Threat**: Do you believe global warming will pose a serious threat to your way of life in your lifetime? (Yes, No). The outcome **Not A Threat** is coded 1 if the subject chose "No" and 0 otherwise.

A.3.1 Results

As shown in Table A.7, the hiatus treatment causes subjects to increase their belief that global warming was due to natural causes by 10 percentage points. Any effect that the treatment may have had on belief that global warming will pose a serious threat was statistically indistinguishable from zero.

Table A.7: Effects of Hiatus Treatment

	Natural Cause	Not a Threat
Treatment: Hiatus	0.114***	0.027
	(0.024)	(0.024)
Constant	0.557	0.480
	(0.017)	(0.017)
N	1,690	1,690
R^2	0.014	0.001

*p < .1; **p < .05; ***p < .01
Robust standard errors are in parentheses.

A.4 Study 4: Gun Control

Haider-Markel and Joslyn (2001) report the results of test of experimental primes on attitudes toward gun control, finding that a public safety frame decreased support for a controlled carry law, relative to a citizen's rights frame.

A.4.1 Original Study

Sample: 518 adult residents of Kansas City.

Dates of data collection: March 7 through April 3, 1999.

Treatments:

- **Public Safety Frame**: Concealed handgun laws have recently received national attention. Some people have argued that laws allowing citizens to carry concealed handguns **threaten public safety** because they would allow almost anyone to carry a gun almost anywhere, even onto **school grounds**. What do you think about concealed handgun laws?
- **Citizens' Rights Frame**: Concealed handgun laws have recently received national attention. Some people have argued that law-abiding citizens have the **right** to protect themselves. What do you think about concealed handgun laws?

Outcomes:

- **Support for concealed handgun laws**: 1: Strongly Oppose to 7: Strongly Support.

The data for the original study were not available. However, the original article (p. 526) included histograms of the dependent variable by treatment condition. For the "Reconstructed" version of the dependent variable, I measured the height of each bar of the histogram in pixels and then backed out the number of subjects giving each response. This procedure worked well for the public safety frame – the reconstructed variable had a mean equal to the mean reported in the main text (2.5). However, the reconstructed mean for the citizens' rights frame was lower than reported (3.3 versus 3.5). I then constructed an "Reported" version of the dependent variable that matched this mean that required as few changes to the pixel-reconstructed version as possible. These changes were done by hand and almost certainly do not reflect the true dataset; it is nevertheless close enough for assessing standard errors to a first approximation.

A.4.2 Replication Study

Sample: 1,502 Mechanical Turk respondents in wave 1; 1,311 in wave 2.

Dates of data collection: March 2015.

Treatments: In addition to the **Public Safety Frame** and **Citizens' Rights Frame** treatments above, I included a **No Frame** control condition: Concealed handgun laws, which allow citizens to obtain permits to carry concealed handguns, have recently received national attention. What do you think about concealed handgun laws?

A.4.3 Results

As shown in Table A.8, the public safety frame had strong negative effects on support for concealed handgun laws in the original study (columns 1 and 2) and in the first wave of the MTurk replication (column 3). By the second wave, the strong effects in the first wave had dissipated (column 4).

Table A.8: Gun Control Original and Replication Results

	Support for Concealed Carry Law			
	Reconstructed	Reported	Wave 1	Wave 2
	(1)	(2)	(3)	(4)
Public Safety Frame	-0.806^{***}	-0.978^{***}		
	(0.205)	(0.207)		
Pure Control			0.525^{***}	-0.033
			(0.126)	(0.137)
Constant (Rights Frame)			0.421^{***}	0.112
			(0.131)	(0.141)
Constant	3.332	3.504	3.558	3.820
	(0.155)	(0.158)	(0.092)	(0.101)
Sample	Original	Original	MTurk	MTurk
N	518	518	1,502	1,311
R^2	0.029	0.042	0.013	0.001

$^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$
Robust standard errors are in parentheses.
'Reconstructed' refers to the DV measured by histogram pixels.
'Reported' adjusts the 'Reconstructed' DV to match the reported ATE estimate.

A.5 Study 5: Superordinate Identity

Transue (2007) reports the results of a test of identity frames on support for taxes. These frames had small to negligible effects on tax preferences.

A.5.1 Original Study

Sample: 397 adult residents of Minneapolis-Saint Paul.

Dates of data collection: July and August 1998.

Treatments:

- **Ethnic Identity Frame**: How close do you feel to your ethnic or racial group? Very close, somewhat close, not very close, not at all close.
- **Superordinate Identity Frame**: How close do you feel to other Americans? Very close, somewhat close, not very close, not at all close.

Outcomes:

- **Support Tax Increase (Public Schools)**: Some people have said that taxes need to be raised to take care of pressing national needs. How willing would you be to have your taxes raised to improve education in public schools? Very willing, somewhat willing, not very willing, not at all willing.
- **Support Tax Increase (Minority Opportunity)**: Some people have said that taxes need to be raised to take care of pressing national needs. How willing would you be to have your taxes raised to improve educational opportunities for minorities? Very willing, somewhat willing, not very willing, not at all willing.

The original experiment conditioned on subjects' answers to the treatment questions, arguing that effects of the identity frames are mediated through (and moderated by) the extent to which subjects accept their identities. This procedure is prone to bias, so in my reanalysis of the data, I assessed the treatment effect of the identity frame using difference-in-means.

A.5.2 Replication Study

Sample: 496 Mechanical Turk respondents in wave 1, 447 in wave 2.

Dates of data collection: March 2015.

Treatments: The treatments used the identical text as the original.

Outcomes: I used the same outcome measures with the same wording. In the second wave, I randomly varied whether subjects saw the questions as written above or in a "matrix" style format, in order to thwart subjects "learning" the right way to answer. This manipulation did not appear to affect responses, and it is not presented below.

A.5.3 Results

In the original, the average difference in outcomes by identity frame was not statistically significantly different from zero, either for increasing taxes to support minority education or public schools. In the replication, I obtained results with the opposite sign that, for the public schools dependent variable, are weakly significant. The difference in sign notwithstanding, my interpretation of these results is that the frames, if anything, have a negligible effect on tax preferences, a conclusion supported by the results of both studies.

Table A.9: Superordinate Identity Original and Replication Results

	Minority Opportunity			Public Schools		
	Original	Wave 1	Wave 2	Original	Wave 1	Wave 2
	(1)	(2)	(3)	(4)	(5)	(6)
Other Americans	0.030	−0.028	−0.051	0.080	−0.064*	−0.064*
	(0.047)	(0.040)	(0.043)	(0.050)	(0.038)	(0.038)
Constant (Ethnic)	0.466	0.529	0.561	0.492	0.599	0.599
	(0.032)	(0.027)	(0.031)	(0.036)	(0.027)	(0.027)
Sample	Original	MTurk	MTurk	Original	MTurk	MTurk
N	191	254	232	206	242	242
R^2	0.002	0.002	0.006	0.012	0.012	0.012

*p < .1; **p < .05; ***p < .01
Robust standard errors are in parentheses.

A.6 Study 6: Patriot Act

Chong and Druckman (2010) report the results of an experiment testing the effects of pro and con information on attitudes toward the Patriot Act. This study is unique among those I replicated in being the only study to have included an over time component originally.

A.6.1 Original Study

Sample: 862 adult Americans.

Dates of data collection: December 2009.

Treatments: Subjects could be assigned to receive no treatment, a series of Pro-Patriot Act messages, a series of Con-Patriot Act messages, or both types of messages.

Pro messages:

- The Patriot Act was enacted in the weeks after September 11, 2001 to strengthen law enforcement powers and technology.
- Under the Patriot Act, law enforcement agencies have more tools to prevent new terrorist incidents.
- The Patriot Act gives U. S. security forces the resources they need to identify terrorist plots on American soil and to prevent attacks before they occur.
- The Patriot Act enhances domestic security through counterterrorism funding, surveillance, border protection, and other security policies.
- The Patriot Act includes less known provisions including funding for terrorism victims and their families.
- The Patriot Act enables officials to effectively combat national security threats, and provides prompt aid and compensation to victims in the event of a terrorist attack.

Con messages:

- The Patriot Act was enacted in the weeks after September 11, 2001 to strengthen law enforcement powers and technology.
- The Patriot Act has sparked numerous controversies and been criticized for weakening the protection of citizens' civil liberties.
- Under the Patriot Act, the government has access to citizens' confidential information from telephone and e-mail communications.

- The Patriot Act allows law enforcement officials to search citizens' homes, businesses, and financial records without their permission or knowledge.
- The Patriot Act significantly expands government policing powers without specifying an agency that is responsible for safeguarding citizens' rights.
- Since its passage, the Patriot Act has been challenged in federal courts on the grounds that many of its provisions are unconstitutional.

Subjects who were assigned to Pro, Con, or Both information treatments were also assigned to a processing condition. I will collapse over these categories in all my analyses.

- On-line processing. After reading each statement, subjects are asked: "To what extent does this statement decrease or increase your support for the Patriot Act?"
- Memory-based processing. After reading each statement, subjects are asked: "How dynamic would you say this statement is? (Remember that a statement is more dynamic when it uses more vivid action words.)"
- No instructions. Subjects are asked to read each statement with no further instructions.

Outcomes:

- **Patriot Act Support**: "Do you oppose or support the Patriot Act?" 1: Oppose very strongly to 7: Support very strongly.

A.6.2 Replication Study

Sample: 1161 Mechanical Turk respondents in wave 1, 912 in wave 2.

Dates of data collection: March 2015.

Treatments: The treatments used the identical text as the original.

Outcomes: I used the same outcome measure with the same wording.

A.6.3 Results

In wave 1 of both the original and replication experiments, Pro information increases support for the Patriot Act and Con information decreases it. When both Pro and Con information treatments are presented together, there is no average change in attitudes relative to control. In the original, these effects decline within 10 days and are no longer distinguishable from zero. In the replication, the Con information has a strongly persistent effect on attitudes toward the Patriot Act.

Table A.10: Patriot Act Original and Replication Results

	Patriot Act Support			
	Wave 1	Wave 2	Wave 1	Wave 2
	(1)	(2)	(3)	(4)
Con Information	-0.743^{***}	-0.254	-0.698^{***}	-0.518^{**}
	(0.213)	(0.217)	(0.182)	(0.220)
Pro + Con Information	-0.079	-0.025	-0.099	-0.203
	(0.209)	(0.212)	(0.184)	(0.218)
Pro Information	0.648^{***}	0.184	0.603^{***}	0.156
	(0.207)	(0.211)	(0.186)	(0.218)
Constant (Control)	4.455	4.384	3.530	3.621
	(0.180)	(0.184)	(0.159)	(0.190)
Sample	Original	Original	MTurk	MTurk
N	826	826	1,161	912
R^2	0.091	0.010	0.077	0.022

$^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$
Robust standard errors are in parentheses.

A.7 Study 7: Elite Endorsements

Nicholson (2012) reports the results of a test of partisan cues on support for legislation. Relative to out-party cues, in-party cues increase support for legislation.

A.7.1 Original Study

Sample: 1315 Americans who identify as a Republican or a Democrat, surveyed by YouGov. 915 responded to the foreclosure question and 718 responded to the immigration question.

Dates of data collection: The 2008 Presidential campaign.

Treatments:

- **In Party**: For each of the dependent variables shown below, I define a subject as being in the "In Party" condition if their own party identification matches that of the endorser.
- **Out Party**: For each of the dependent variables shown below, I define a subject as being in the "Out Party" condition if their own party identification is opposite that of the endorser.
- **No Cue**: Subjects in the no cue condition were asked about their support for the two bills without any endorsement.

Outcomes:

- **Support for Foreclosure Bill**: A bill circulating in Congress [supported by Barack Obama/John McCain/George W. Bush/the Democratic Party/the Republican Party] would allow the Federal Housing Administration to guarantee up to $300 billion in new loans to help at-risk homeowners refinance into more affordable mortgages. What is your view of this bill? (-1: I oppose this policy, 0: Not sure, 1: I support this policy)
- **Support for Immigration Bill**: As you know, there has been a lot of talk about immigration reform policy in the news. One proposal [backed by Barack Obama/John McCain/George W. Bush/the Democratic Party/the Republican Party]provided legal status and a path to legal citizenship for the approximately 12 million illegal immigrants currently residing in the United States. What is your view of this immigration reform policy? (-1: I oppose this policy, 0: Not sure, 1: I support this policy)

The original analysis did not group the treatments into In Party and Out Party, but because the experiment is somewhat underpowered, I think that doing so gives a clearer picture of the modest effect that party endorsements in general have on support. In the

analyses I present below, I include an indicator for party, as Republicans and Democrats have different probabilities of assignment to in- and out- party treatments.

A.7.2 Replication Study

Sample: 1,245 Mechanical Turk respondents in wave 1; 1,095 in wave 2.

Dates of data collection: March 2015.

Treatments: Treatments were defined in the same way as the original. I added two treatment conditions, Hillary Clinton and Jeb Bush. Unfortunately, due to a coding error, those assigned to see endorsements by Hillary Clinton or Jeb Bush in the immigration experiment actually saw an endorsement by George W. Bush. This error is addressed when constructing the categorical party match treatment variable.

Outcomes: The wording of the outcome questions was identical to the original.

A.7.3 Results

Table A.11: Elite Endorsements Original and Replication Results: Immigration

	Support for Immigration Bill			Support for Foreclosure Bill		
	Original	Wave 1	Wave 2	Original	Wave 1	Wave 2
	(1)	(2)	(3)	(4)	(5)	(6)
In Party Cue	0.172**	0.112**	0.029	0.191***	0.127***	−0.008
	(0.069)	(0.054)	(0.057)	(0.059)	(0.046)	(0.049)
No Cue	0.118	−0.021	−0.017	0.220***	−0.045	−0.039
	(0.088)	(0.072)	(0.076)	(0.074)	(0.070)	(0.076)
Republican	−0.656***	−0.714***	−0.778***	−0.623***	−0.429***	−0.417***
	(0.062)	(0.056)	(0.059)	(0.054)	(0.049)	(0.053)
Constant	2.009	2.487	2.551	2.227	2.403	2.555
	(0.054)	(0.033)	(0.035)	(0.046)	(0.032)	(0.033)
Sample	Original	MTurk	MTurk	Original	MTurk	MTurk
N	718	1,245	1,095	915	1,245	1,095
R^2	0.130	0.134	0.166	0.130	0.070	0.063

*p < .1; **p < .05; ***p < .01
Robust standard errors are in parentheses.

Relative to the Out Party condition, subjects in the In Party condition report higher

support for both policies. This is true in both the original and in the replication. Unsurprisingly, both the original and the replication studies find a strong negative correlation between party identification and support for the policies, both of which are relatively liberal. Neither of the effects from the first wave of the replication study persist into wave 2.

A.8 Studies 8 and 9: Free Trade

Hiscox (2006) reports the results of a test of positive and negative frames on support for free trade.

A.8.1 Original Study

Sample: 1,578 Americans.

Dates of data collection: The 2008 Presidential campaign.

Treatments: This study employed a 2 X 4 factorial design, where the first factor is the Expert treatment and the second factor is the frame the subject is shown: positive, negative, both, or neither.

- **Expert**: According to the New York Times, almost 100 percent of American economists support increasing trade with other nations. In 1993 over a thousand economists, including all living winners of the Nobel Prize in economics, signed an open letter to the New York Times urging people to support efforts to increase trade between the United States and neighboring countries.
- **Positive**: Many people believe that increasing trade with other nations creates jobs and allows Americans to buy more types of goods at lower prices.
- **Negative**: Many people believe that increasing trade with other nations leads to job losses and exposes American producers to unfair competition.
- **Positive + Negative**: Many people believe that increasing trade with other nations creates jobs and allows Americans to buy more types of goods at lower prices. Others believe that increasing trade with other nations leads to job losses and exposes American producers to unfair competition.
- **Control** (No introduction before asking the free trade question.)

Outcomes:

- **Support for Free Trade**: Do you favor or oppose increasing trade with other nations? (0: oppose; 1: favor)

A.8.2 GfK Replication Study

Sample: 2,084 GfK respondents in wave 1; 1,838 in wave 2.

Dates of data collection: August 2015.

Treatments: Treatments were defined in the same way as the original.

Outcomes: The wording of the outcome questions was identical to the original.

A.8.3 MTurk Replication Study

Sample: 2,972 Mechanical Turk respondents in wave 1; 2,307 in wave 2.

Dates of data collection: July 2015.

Treatments: Treatments were defined in the same way as the original.

Outcomes: The wording of the outcome questions was identical to the original.

A.8.4 Results

Table A.12: Free Trade Original and Replication Results

	Support for Free Trade				
	Original	Wave 1	Wave 2	Wave 1	Wave 2
	(1)	(2)	(3)	(4)	(5)
Expert	0.110***	0.080***	0.002	0.126***	0.073***
	(0.028)	(0.022)	(0.024)	(0.015)	(0.018)
Positive	−0.078**	−0.015	−0.003	−0.035*	−0.037
	(0.038)	(0.030)	(0.033)	(0.019)	(0.024)
Negative	−0.114***	−0.148***	−0.020	−0.146***	−0.056**
	(0.039)	(0.031)	(0.033)	(0.021)	(0.025)
Pos + Neg	−0.186***	−0.106***	−0.014	−0.151***	−0.049**
	(0.039)	(0.031)	(0.033)	(0.021)	(0.024)
Constant (Control)	0.686	0.713	0.708	0.784	0.766
	(0.029)	(0.024)	(0.026)	(0.016)	(0.019)
Sample	Original	GfK	GfK	MTurk	MTurk
N	1,578	2,084	1,838	2,972	2,307
R^2	0.031	0.026	0.0003	0.046	0.010

*p < .1; **p < .05; ***p < .01
Robust standard errors are in parentheses.

The results of this experiment are remarkably consistent across all three versions. The Expert treatment causes about a 10 point increase in support for free trade; on GfK this effect dissipates after 10 days but declines to only 7 points on MTurk. The Positive treatment is oddly ineffective: null or even negative effects across the board. The Negative treatment

and the Positive + Negative treatments both exert strongly negative effects on free trade support, both of which persist in the MTurk sample but not the GfK sample.

A.9 Studies 10 and 11: Frame Breadth

Hopkins and Mummolo (2015) show that frames in one issue area by and large affect only "target" attitudes and not attitudes in other domains.

A.9.1 Original Study

Sample: 3,269 Adult Americans (GfK).

Dates of data collection: The 2008 Presidential campaign.

Treatments: Subjects were assigned at random to see two of these treatment texts. Subjects saw one at random, answered a random two of the outcome variables, then saw a second treatment, and answered the remainder of the treatment questions. "The argument below was recently made by a U.S. Senator. Please take a moment to read the argument carefully and then tell us what you think. [Treatment Text] Do you think the Senator is making a convincing argument? Please tell us why or why not. [Text entry]. "

- **Crime**: America is very vulnerable to violent crime, with forty-two Americans murdered every single day on average. Innocent people can be killed in their front yards. Across the country, we have to do everything we can to reduce the threat of violent crime. We have to stop violent criminals before they act. This means cracking down on the smaller offenses that all too often lead to violent crime, and making sure that convicted criminals always serve out their full sentences.

- **Health Care**: Health care is one of the most complicated issues we face. It involves 1 of every 6 dollars spent here in the United States. The health care system includes millions of doctors and nurses and thousands of hospitals and clinics. Together, they regularly make decisions that can mean life or death. The government in Washington can't even balance its own budget. How can we trust it to run something as complicated as the health care system?

- **Stimulus**: With a recession as deep as this one, there are more than 10 million unemployed Americans, and it's going to take years for our economy to recover. In February 2009, the government in Washington made things worse by passing an $800 billion stimulus package, which is more than $2,500 for every person living in this country. Now, it looks like a lot of that money didn't help the economy. Unemployment is still very high. The money went to pork-barrel projects and federal bureaucrats rather than creating jobs for unemployed Americans. The government in Washington can't even balance its own budget. How can we trust it to spend so much taxpayer money?

- **Terror**: The September 11th attacks and the news that al-Qaeda was planning new attacks on U.S. soil show how vulnerable America still is to terrorists. Innocent people

> can be killed while traveling to visit family or going to work. Across the country, we have to do everything we can to reduce the threat of terrorism. We have to stop terrorists before they act. This means conducting more frequent searches of suspicious people boarding planes, trains, subways, and buses.

Outcomes: Spending preferences were measured in all four areas, regardless of treatment assignment. The response options were 1: Decreased a lot, 2: Decreased a moderate amount, 3: Decreased a little; 4: Kept about the same; 5: Increased a little; 6: Increased a moderate amount; 7: Increased a great deal.

- **Crime Spending** Should federal spending on dealing with crime be increased, decreased, or kept the same?
- **Health Care Spending** Should federal spending on health care be increased, decreased, or kept the same?
- **Stimulus Spending** Should federal spending to stimulate the economy be increased, decreased, or kept the same?
- **Terrorism Spending** Should federal spending on the war on terrorism be increased, decreased, or kept the same?

The relatively complicated random assignment procedure requires the analyst to make some choices. I chose to define a subject as in treatment if they answered the "target" dependent variable after being exposed to the corresponding treatment (For example, if a person saw the health care treatment *after* answering the health care treatment, I would define that subject as being in control for the health care experiment).

A.9.2 GfK Replication Study

Sample: 3,189 GfK respondents in wave 1; 2,510 in wave 2.

Dates of data collection: August 2015.

Treatments: I assigned each subject to only see one of the treatments, which was shown before any of the dependent variables. This design simplifies the analysis while still allowing for the assessment of "direct" and "spillover" effects across issues.

Outcomes: The wording of the outcome questions was identical to the original.

A.9.3 MTurk Replication Study

Sample: 2,972 Mechanical Turk respondents in wave 1; 2,281 in wave 2.

Dates of data collection: July 2015.

Treatments: Treatments were defined in the same way as in the GfK replication.

Outcomes: The wording of the outcome questions was identical to the original.

A.9.4 Results

Table A.13: Crime Spending Original and Replication Results

	Support for Crime Spending				
	Original	Wave 1	Wave 2	Wave 1	Wave 2
	(1)	(2)	(3)	(4)	(5)
Crime Argument	0.079	0.113	−0.083	−0.016	0.068
	(0.068)	(0.069)	(0.070)	(0.069)	(0.076)
Constant (Control)	4.445	4.481	4.583	4.220	4.185
	(0.035)	(0.030)	(0.034)	(0.029)	(0.032)
Sample	Original	GfK	GfK	MTurk	MTurk
N	3,269	3,189	2,510	2,972	2,281
R^2	0.001	0.001	0.001	0.00002	0.0004

$^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$
Robust standard errors are in parentheses.

Table A.14: Health Care Spending Original and Replication Results

	Support for Health Care Spending				
	Original	Wave 1	Wave 2	Wave 1	Wave 2
	(1)	(2)	(3)	(4)	(5)
Health Care Argument	−0.044	−0.044	0.054	−0.080	−0.049
	(0.081)	(0.085)	(0.088)	(0.076)	(0.088)
Constant (Control)	4.606	4.289	4.256	4.928	4.920
	(0.043)	(0.037)	(0.041)	(0.034)	(0.039)
Sample	Original	GfK	GfK	MTurk	MTurk
N	3,271	3,188	2,513	2,972	2,281
R^2	0.0001	0.0001	0.0002	0.0004	0.0001

$^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$
Robust standard errors are in parentheses.

Table A.15: Stimulus Spending Original and Replication Results

	Support for Stimulus Spending				
	Original	Wave 1	Wave 2	Wave 1	Wave 2
	(1)	(2)	(3)	(4)	(5)
Stimulus Argument	−0.317***	−0.455***	−0.113	−0.415***	−0.252***
	(0.087)	(0.079)	(0.087)	(0.074)	(0.081)
Constant (Control)	4.200	4.233	4.099	4.655	4.600
	(0.048)	(0.035)	(0.038)	(0.031)	(0.035)
Sample	Original	GfK	GfK	MTurk	MTurk
N	3,266	3,179	2,512	2,972	2,281
R^2	0.006	0.012	0.001	0.012	0.004

$^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$
Robust standard errors are in parentheses.

Table A.16: Terrorism Spending Original and Replication Results

	Support for Terrorism Spending				
	Original	Wave 1	Wave 2	Wave 1	Wave 2
	(1)	(2)	(3)	(4)	(5)
Terrorism Argument	0.361***	0.265***	0.056	0.155**	0.031
	(0.076)	(0.080)	(0.088)	(0.077)	(0.087)
Constant (Control)	3.571	4.233	4.323	3.291	3.228
	(0.042)	(0.036)	(0.038)	(0.035)	(0.039)
Sample	Original	GfK	GfK	MTurk	MTurk
N	3,272	3,180	2,513	2,972	2,281
R^2	0.011	0.004	0.0002	0.001	0.0001

*p < .1; **p < .05; ***p < .01
Robust standard errors are in parentheses.

A.10 Studies 12 and 13: Polarization

Levendusky and Malhotra (2015) show that descriptions of a polarized electorate led subjects to view Republicans and Democrats as further apart on issues.

A.10.1 Original Study

Sample: 1587 American adults (GfK).

Dates of data collection: 11/29/12 - 12/12/12.

Treatments: Subjects were assigned at random to see one of three treatments: **Polarized**, **Moderate**, or a **Placebo**.

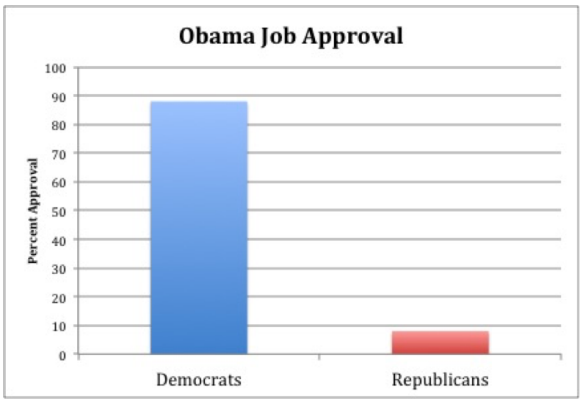
Polarized:

Electorate as Divided as Ever
Jefferson Graham (USA Today)

In the aftermath of the 2012 presidential election, interviews with voters at a diner in Smithfield, PA reveal an electorate as divided as ever. When asked about the importance of the election results, Republican Marlene Evers of nearby Fairchance said, "I can't believe Obama won. He is a radical socialist. He will destroy the Christian values set forth by the Founding Fathers that have made this country great. If he gets his way, he'll overturn 5,000 years of tradition and allow gay marriage, destroying the American family. We must stop him any way we can."

Later on that evening, Democratic voter and Obama supporter Dan Thompson of Masontown pointed to economic issues as influencing his vote in the election. "The Republican Party is for corporate greed and will do nothing but destroy the lives and hopes of regular working people in this country. They tried to use voter ID laws to steal this election, because they know the American people reject their ideas." He added, "Bush was a complete idiot who bankrupted this nation with the Iraq War, and Romney would have been just as bad, destroying the economy. Republicans want to roll back women's reproductive freedom by restricting access to contraception and labeling women who defend it sluts and prostitutes."

As we left Smithfield, it is clear that Republicans and Democrats in the area seem as divided as ever before. This same pattern also holds nationally: Democrats and Republicans across the country are deeply divided. For example, Gallup data released last week shows that while nearly 9 in 10 Democratic voters (88 percent) approve of President Obama's job as president, less than 1 in 10 Republicans (8 percent) approves. This 80 point gap between the parties in approval is among the largest ever recorded (see figure). "Differences in Obama's approval reflect fundamental divides between the parties," says Stanford political science professor Neil Malhotra. "Democrats and Republicans really do hold different beliefs."



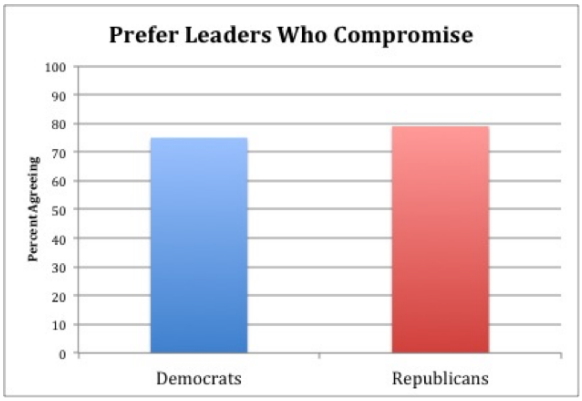
Moderate:

Electorate Remains Moderate
Jefferson Graham (USA Today)

In the aftermath of the 2012 presidential election, interviews with voters at a diner in Smithfield, PA reveal few real divisions in the electorate. When asked about Obama's victory, Republican Marlene Evers of nearby Fairchance said, "I don't agree with all of Obama's economic policies, but he seems to be trying hard to resolve America's economic problems. He's doing things that we all agree with, like trying to bring down the deficit. He's also trying to find a middle ground on social issues like his gay marriage decision. While he supports gay marriage, he did not push to change federal on policy on this issue, knowing that it might upset some voters. I am pro-life, but I agree with President Obama that women need access to safe and affordable family planning tools."

Later on that evening, Democratic voter and Obama supporter Dan Thompson of Masontown pointed to economic issues as influencing his vote in the election. "I'm not an ideologue. I find myself mostly in the middle, and really just want the country to get back on track and find common-sense solutions to get our economy fixed." Thompson also noted that he wanted a break from the culture wars, and wants politicians to stop focusing on controversial social issues like abortion. "Americans can all agree that, even if we support the right to abortion, it should be rare and avoided, and the President's policies are trying to reduce the need for abortion in this country."

As we left Smithfield, it is surprising to find that Republicans and Democrats in the electorate seem to want the same things, very different from the picture we get from Washington. This same pattern also holds nationally: Democrats and Republicans across the country are not really very divided. For example, recent data from the Pew Center for the People and the Press show that Democrats and Republicans alike overwhelmingly support leaders who compromise to get things done. 75 percent of Democrats feel this way, as do 79 percent of Republicans, a nearly identical level (see figure). "This shows that there is no divide between ordinary Democrats and Republicans," says Stanford political science professor Neil Malhotra. "Democrats and Republicans really do want the same things."



Placebo:

Blake Predicts Pop Singer Will Win Season 3 of the Voice

Blake Shelton knows what he wants to happen on The Voice this season. "There are some pop singers that are really good," said Shelton, who coached pop singer Dia Frampton to the final four during season 1. "We had a few last year but we didn't end up with a pop singer winning the show. Jermaine Paul, an R & B singer, won season 2."

The celebrity coaches – Blake, along with Christina Aguilera, Cee Lo Green and Adam Levine – will be back in the rotating chairs for season 3's blind auditions.

During the blind auditions, the celebrity coaches will hear contestants sing without seeing what they look like. If they like what they hear, the coaches can push the button, turn their chair around and see the face behind the voice, selecting the singer for their team. So, there's no guarantee that Blake will coach the aspiring pop stars.

"I know anything can happen," he says, "but I really believe in my heart that one of several pop singers that we have on our show this year have a shot at winning the whole thing."

Outcomes: I focus on two dependent variables. The first, *Perceived Polarization*, is built from subjects' responses to a series of policy questions. After giving their response to each question in the list below, subjects were asked how they think a "typical Democratic voter" and a "typical Republican voter" would respond to each question. The outcome variable is the average of the absolute values of the differences in subjects' Democratic and Republican Responses.

- The tax rates on the profits people make from selling stocks and bonds, called capital gains taxes, are currently lower than the income tax rates many people pay. Do you think that capital gains tax rates should be increased, decreased, or kept about the same? (7 point scale. Decreased a lot: -3; Increased a lot: 3)
- There is some debate about whether or not undocumented immigrants who were brought to this country illegally as children should be deported. Which of the following positions on the scale below best represents your position on this issue? (7 point scale. Very strongly oppose deportation: -3; Very strongly support deportation: 3)
- The United States is currently considering signing additional free trade agreements with Central American, South American, and Asian countries. The Democratic Party wants to make it more difficult for the U.S. to enter into such agreements. The Republican Party wants to make it easier to do so. What do you think? Do you support or oppose the United States signing more free trade agreements with Central American, South American, and Asian countries? (7 point scale. Very strongly oppose free trade: -3; Very strongly support free trade: 3)

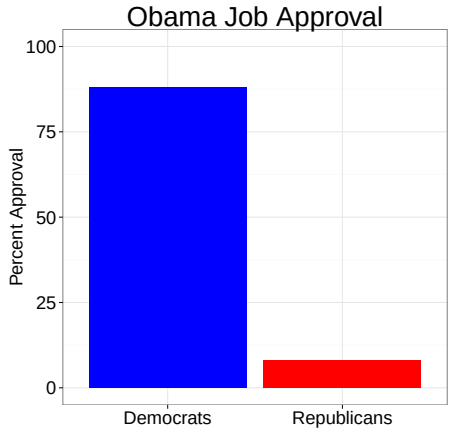
- Public financing of elections is when the government pays for the cost of campaigning for various offices, rather than the campaigns relying on donations from the general public, corporations, or unions. Democrats typically support public financing plans while Republicans have wanted to eliminate them. What do you think? Do you support or oppose the government paying for the public financing of elections? (7 point scale. Very strongly oppose public financing: -3; Very strongly support public financing: 3)

A.10.2 GfK Replication Study

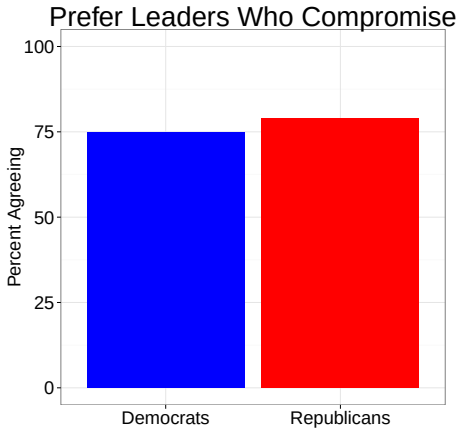
Sample: 2,115 GfK respondents in wave 1; 1,855 in wave 2.

Dates of data collection: August 2015.

Treatments: The wording of the **Polarized** and **Moderate** conditions is the same as above. I used slightly different figures:



(a) Polarized



(b) Moderate

I used a different placebo text, due to the time-bound nature of reality TV:
Placebo:

The Lasting Appeal of So You Think You Can Dance

After 12 seasons of dance, you've got to shake things up a bit and bring in something fresh – and the new format really has done that. From what I've seen, being in Vegas and watching the audition cities that I had not seen previously, we are getting some of the best of the best talent. And on my side, the Street side, we're getting some incredible people who previously would not have even tried for So You Think You Can Dance. We're dealing with people who have never left their cities, much less taken any dance classes or had any

formal training, and now they're starting to come out and wanting to show what they can do – because they have the chance to do what it is that they do.

There's something about a family show like So You Think You Can Dance offering a wide variety of talent in many different packages – whether it be color, creed, size, anything – because you get to see these people doing what it is they're strongest at, and you never know who that's going to inspire as they're watching. And I think that's been one of the strongest common threads through every season: that it's ongoing inspiration for the future generation, and there's always somebody that you can connect with. Out of the 20, there's at least one person, and even if they don't make it to the top 20, if you watch through the audition specials, you'll see someone that you connect with; they'll strike a chord. It moves you.

Outcomes: The wording of the outcome questions was identical to the original.

A.10.3 MTurk Replication Study

Sample: 2,972 Mechanical Turk respondents in wave 1; 2,288 in wave 2.

Dates of data collection: July 2015.

Treatments: Treatments were defined in the same way as in the GfK replication.

Outcomes: The wording of the outcome questions was identical to the original.

A.10.4 Results

As shown in Table A.17, the polarized treatment did cause subjects to view themselves as more moderate. By contrast, as shown in Table A.18, subjects viewed Republicans and Democrats as being further apart on issues. This effect appears to have persisted for 10 days.

Table A.17: Extremity Original and Replication Results

	Extremity of Policy Views				
	Original	Wave 1	Wave 2	Wave 1	Wave 2
	(1)	(2)	(3)	(4)	(5)
Placebo Treatment	−0.008	0.016	−0.037	−0.001	−0.030
	(0.064)	(0.044)	(0.049)	(0.030)	(0.033)
Polarizied Treatment	−0.042	0.043	0.031	−0.036	−0.039
	(0.061)	(0.043)	(0.048)	(0.030)	(0.033)
Constant (Moderate)	1.318	1.395	1.432	1.540	1.626
	(0.045)	(0.031)	(0.035)	(0.021)	(0.024)
Sample	Original	GfK	GfK	MTurk	MTurk
N	1,521	2,115	1,855	2,972	2,288
R^2	0.001	0.001	0.001	0.001	0.001

*p < .1; **p < .05; ***p < .01
Robust standard errors are in parentheses.

Table A.18: Perceived Polarization Original and Replication Results

	Perceived Polarization				
	Original	Wave 1	Wave 2	Wave 1	Wave 2
	(1)	(2)	(3)	(4)	(5)
Placebo Treatment	0.260**	0.315***	0.022	0.320***	0.041
	(0.123)	(0.096)	(0.107)	(0.059)	(0.064)
Polarizied Treatment	0.413***	0.548***	0.202**	0.413***	0.131**
	(0.121)	(0.093)	(0.097)	(0.058)	(0.064)
Constant (Moderate)	2.550	2.639	2.960	3.126	3.550
	(0.084)	(0.068)	(0.072)	(0.041)	(0.045)
Sample	Original	GfK	GfK	MTurk	MTurk
N	1,471	2,115	1,845	2,972	2,288
R^2	0.013	0.020	0.003	0.018	0.002

*p < .1; **p < .05; ***p < .01
Robust standard errors are in parentheses.

A.11 Study 14: Immigration

A.11.1 Original Study

Brader, Valentino and Suhay (2008) find that news articles focusing on the negative economic impacts of immigration decrease support for immigration.

Sample: 353 Adult Americans (GfK)

Dates of data collection: 10/21/03 - 11/5/03

Treatments: Subjects could be assigned to one of four treatment stories or a placebo news story. The treatment stories were made to look like articles in the New York Times.

Negative European/Latino


The New York Times

Immigration Concerns Governors

Questions Raised About Economic, Cultural Impact of Immigrants

NEW YORK (AP) — During the 1990s, more immigrants entered the United States than in any previous decade, and the growing number of immigrants in the U.S. clearly has some Americans worried. At a state governors' convention in June, many governors called for the Bush Administration and Congress to step in to restrict the flow of immigrants.

Several governors voiced concern that immigrants are driving down the wages of American workers, while taxpayers are forced to meet the rising costs of social services for the newcomers. Governors say these views are shared by many of their constituents.

John Baine, shift manager at a large auto parts factory in Cleveland, said he is angered that "a number of friends have been laid-off or forced to take a pay cut" because of the influx of cheap immigrant labor.

Nancy Petrey, a Boulder, Colo. nurse, has seen staff let go for similar reasons. "People give twenty years of their lives to this hospital and then, boom, they're out the door because some foreigner will do their job for half the pay," Petrey said. "It just isn't right."

Governors also say constituents are worried that the country is no longer a "melting pot," because new immigrants are not adopting American values or blending into their new social world.

Mary Stowe, an Omaha-based sales associate, says she is frustrated by the fact that recent immigrants to her area "do not learn English or make any effort to fit in."

Bob Callaway, a construction supervisor in Newark, says he sees similar problems with immigrants hired by his company. "These people are totally unwilling to adopt American values like hard work and responsibility," Callaway said. "I try not to complain, but sometimes they are so pushy and uncooperative – it's not acceptable."

When asked his opinion, Jose Sanchez, a recent immigrant from Mexico, says he welcomes the chance for a better life in America. "Many of my cousins find work here and now it's my turn. I want a good job and benefits."

"But," Sanchez added, "that doesn't mean I have to change who I am. We love our culture. I'm proud to be from Mexico."

While there was agreement at the convention that the federal government needs to do more to help states manage the rising tide of newcomers, few governors agree on exactly why immigration levels have increased.

Some blame the Immigration Act passed by Congress in 1990, which loosened federal restrictions on immigration. Others point to the fact that large companies are attracting immigrants to the U.S. with the promise of prosperity, a practice that has become widespread in recent years.

Still others maintain that, in a world full of turmoil, people are attracted here by the hope of a better way of life.

Whatever is bringing immigrants to these shores in record numbers, everyone seems certain that the numbers will continue to grow.

Associated Press

Jose Sanchez is one of thousands of new immigrants who arrived in the U.S. during the first half of this year.

(a) Negative Latino


The New York Times

Immigration Concerns Governors

Questions Raised About Economic, Cultural Impact of Immigrants

NEW YORK (AP) — During the 1990s, more immigrants entered the United States than in any previous decade, and the growing number of immigrants in the U.S. clearly has some Americans worried. At a state governors' convention in June, many governors called for the Bush Administration and Congress to step in to restrict the flow of immigrants.

Several governors voiced concern that immigrants are driving down the wages of American workers, while taxpayers are forced to meet the rising costs of social services for the newcomers. Governors say these views are shared by many of their constituents.

John Baine, shift manager at a large auto parts factory in Cleveland, said he is angered that "a number of friends have been laid-off or forced to take a pay cut" because of the influx of cheap immigrant labor.

Nancy Petrey, a Boulder, Colo. nurse, has seen staff let go for similar reasons. "People give twenty years of their lives to this hospital and then, boom, they're out the door because some foreigner will do their job for half the pay," Petrey said. "It just isn't right."

Governors also say constituents are worried that the country is no longer a "melting pot," because new immigrants are not adopting American values or blending into their new social world.

Mary Stowe, an Omaha-based sales associate, says she is frustrated by the fact that recent immigrants to her area "do not learn English or make any effort to fit in."

Bob Callaway, a construction supervisor in Newark, says he sees similar problems with immigrants hired by his company. "These people are totally unwilling to adopt American values like hard work and responsibility," Callaway said. "I try not to complain, but sometimes they are so pushy and uncooperative – it's not acceptable."

When asked his opinion, Nikolai Vandinsky, a recent immigrant from Russia, says he welcomes the chance for a better life in America. "Many of my cousins find work here and now it's my turn. I want a good job and benefits."

"But," Vandinsky added, "that doesn't mean I have to change who I am. We love our culture. I'm proud to be from Russia."

While there was agreement at the convention that the federal government needs to do more to help states manage the rising tide of newcomers, few governors agree on exactly why immigration levels have increased.

Some blame the Immigration Act passed by Congress in 1990, which loosened federal restrictions on immigration. Others point to the fact that large companies are attracting immigrants to the U.S. with the promise of prosperity, a practice that has become widespread in recent years.

Still others maintain that, in a world full of turmoil, people are attracted here by the hope of a better way of life.

Whatever is bringing immigrants to these shores in record numbers, everyone seems certain that the numbers will continue to grow.

Associated Press

Nikolai Vandinsky is one of thousands of new immigrants who arrived in the U.S. during the first half of this year.

(b) Negative European

Immigration Concerns Governors
Questions Raised About Economic, Cultural Impact of Immigrants

NEW YORK (AP) - During the 1990s, more immigrants entered the United States than in any previous decade, and the growing number of immigrants in the U.S. clearly has some Americans worried. At a state governors' convention in June, many governors called for the Bush Administration and Congress to step in to restrict the flow of immigrants.

Several governors voiced concern that immigrants are driving down the wages of American workers while taxpayers are forced to meet the rising costs of social services for the newcomers. Governors say these views are shared by many of their constituents.

John Baine, shift manager at a large auto parts factors in Cleveland, said he is angered that "a number of friends have been laid-off or forced to take a pay cut" because of the influx of cheap immigrant labor.

Nancy Petrey, a Boulder, Colo. nurse, has seen staff let go for similar reasons. "People give twenty years of their lives to this hospital and then, boom, they're out the door because some foreigner will do their job for half the pay," Petrey said. "It just isn't right."

Governors also say constituents are worried that the country is no longer a "melting pot," because new immigrants are not adopting American values or blending into their new social world.

Mary Stowe, an Omaha-based sales associate, says she is frustrated by the fact that recent immigrants to her area "do not learn English or make any effort to fit in."

Bob Callaway, a construction supervisor in Newark, says he sees similar problems with immigrants hired by his company. "These people are totally unwilling to adopt American values like hard work and responsibility," Callaway said. "I try not to complain, but sometimes they are so pushy and uncooperative – it's not acceptable."

When asked his opinion, [Nikolai Vandisnky]/[Jose Sanchez], a recent immigrant from [Russia]/[Mexico], says he welcomes the chance for a better life in America. "Many of my cousins find work here and now it's my turn. I want a good job and benefits."

"But," [Vandinsky]/[Sanchez] added, "that doesn't mean I have to change who I am. We love our culture. I'm proud to be from [Russia]/[Mexico]."

While there was agreement at the convention that the federal government needs to do more to help states manage the rising tide of newcomers, few governors agree on exactly why immigration levels have increased.

Some blame the Immigration Act passed by Congress in 1990, which loosened federal restrictions on immigration. Others point to the fact that large companies are attracting immigrants to the U.S. with the promise of prosperity, a practice that has become widespread in recent years.

Still others maintain that, in a world full of turmoil, people are attracted here by the hope of a better way of life.

Whatever is bringing immigrants to these shores in record numbers, everyone seems certain that the numbers will continue to grow.

Positive European/Latino

The New York Times

Immigration Heartens Governors

Promise Seen in Economic, Cultural Contribution of Immigrants

NEW YORK (AP) — During the 1990s, more immigrants entered the United States than in any previous decade, and the growing number of immigrants in the U.S. clearly has some Americans hopeful about the future. At a state governors' convention in June, many governors called for the Bush Administration and Congress to protect the flow of immigrants from further restrictions.

Several governors said they are encouraged by how immigrants are helping to strengthen the economy, while also providing a welcome boost to tax revenues. Governors say these views are shared by many of their constituents.

John Baine, shift manager at a large auto parts factory in Cleveland, says he is enthusiastic about how much the influx of immigrant labor has "helped the company keep a lid on costs and remain competitive."

Nancy Petrey, a Boulder, Colo. nurse, has seen similar benefits for the hospital where she works. "These people take jobs that are often hard for us to fill, and they're willing to work shifts that other people don't want," Petrey said. "It's a big help."

Governors also say many constituents take pride in the fact that the country is still a "melting pot," where immigrants continue to bring new experiences and ideas that enrich American culture.

Mary Stowe, an Omaha-based sales associate, says she admires what it must take to "leave home and come to a place that is so different, without knowing the language or anything about the way of life here."

Bob Callaway, a construction supervisor in Newark, says he sees similar qualities in the immigrants hired by his company. "These people are determined and persistent," Callaway said. "I've gotta give 'em credit, they'll do what it takes to get ahead. That's something I respect."

When asked his opinion, Jose Sanchez, a recent immigrant from Mexico, says he welcomes the chance for a better life in America. "Many of my cousins find work here and now it's my turn. I want a good job and benefits."

"But," Sanchez added, "that doesn't mean I have to change who I am. We love our culture. I'm proud to be from Mexico."

While there was agreement at the convention that the federal government needs to do more to help states manage the rising tide of newcomers, few governors agree on exactly why immigration levels have increased.

Some credit the Immigration Act passed by Congress in 1990, which loosened federal restrictions on immigration. Others point to the fact that large companies are attracting immigrants to the U.S. with the promise of prosperity, a practice that has become widespread in recent years.

Still others maintain that, in a world full of turmoil, people are attracted here by the hope of a better way of life.

Whatever is bringing immigrants to these shores in record numbers, everyone seems certain that the numbers will continue to grow.

Jose Sanchez is one of thousands of new immigrants who arrived in the U.S. during the first half of this year.

Associated Press

(a) Positive Latino

The New York Times

Immigration Heartens Governors

Promise Seen in Economic, Cultural Contribution of Immigrants

NEW YORK (AP) — During the 1990s, more immigrants entered the United States than in any previous decade, and the growing number of immigrants in the U.S. clearly has some Americans hopeful about the future. At a state governors' convention in June, many governors called for the Bush Administration and Congress to protect the flow of immigrants from further restrictions.

Several governors said they are encouraged by how immigrants are helping to strengthen the economy, while also providing a welcome boost to tax revenues. Governors say these views are shared by many of their constituents.

John Baine, shift manager at a large auto parts factory in Cleveland, says he is enthusiastic about how much the influx of immigrant labor has "helped the company keep a lid on costs and remain competitive."

Nancy Petrey, a Boulder, Colo. nurse, has seen similar benefits for the hospital where she works. "These people take jobs that are often hard for us to fill, and they're willing to work shifts that other people don't want," Petrey said. "It's a big help."

Governors also say many constituents take pride in the fact that the country is still a "melting pot," where immigrants continue to bring new experiences and ideas that enrich American culture.

Mary Stowe, an Omaha-based sales associate, says she admires what it must take to "leave home and come to a place that is so different, without knowing the language or anything about the way of life here."

Bob Callaway, a construction supervisor in Newark, says he sees similar qualities in the immigrants hired by his company. "These people are determined and persistent," Callaway said. "I've gotta give 'em credit, they'll do what it takes to get ahead. That's something I respect."

When asked his opinion, Nikolai Vandinsky, a recent immigrant from Russia, says he welcomes the chance for a better life in America. "Many of my cousins find work here and now it's my turn. I want a good job and benefits."

"But," Vandinsky added, "that doesn't mean I have to change who I am. We love our culture. I'm proud to be from Russia."

While there was agreement at the convention that the federal government needs to do more to help states manage the rising tide of newcomers, few governors agree on exactly why immigration levels have increased.

Some credit the Immigration Act passed by Congress in 1990, which loosened federal restrictions on immigration. Others point to the fact that large companies are attracting immigrants to the U.S. with the promise of prosperity, a practice that has become widespread in recent years.

Still others maintain that, in a world full of turmoil, people are attracted here by the hope of a better way of life.

Whatever is bringing immigrants to these shores in record numbers, everyone seems certain that the numbers will continue to grow.

Nikolai Vandinsky is one of thousands of new immigrants who arrived in the U.S. during the first half of this year.

Associated Press

(b) Positive European

Immigration Heartens Governors
Promise Seen in Economic, Cultural Contribution of Immigrants

NEW YORK (AP) – During the 1990s, more immigrants entered the United States than in any previous decade, and the growing number of immigrants in the U.S. clearly has some Americans hopeful about the future. At a state governors' convention in June, many governors called for the Bush Administration and Congress to protect the flow of immigrants from further restrictions.

Several governors said they are encouraged by how immigrants are helping to strengthen the economy, while also providing a welcome boost to tax revenues. Governors say these views are shared by many of their constituents.

John Baine, shift manager at a large auto parts factory in Cleveland, says he is enthusiastic about how much the influx of immigrant labor has "helped the company keep a lid on costs and remain competitive."

Nancy Petrey, a Boulder, Colo. nurse, has seen similar benefits for the hospital where she works. "These people take jobs that are often hard for us to fill, and they're willing to work shifts that other people don't want," Petrey said. "It's a big help."

Governors also say many constituents take pride in the fact that the country is still a "melting pot," where immigrants continue to bring new experiences and ideas that enrich American culture.

Mary Stowe, an Omaha-based sales associate, says she admires what it must take to "leave home and come to a place that is so different, without knowing the language or anything about the way of life here."

Bob Callaway, a construction supervisor in Newark, says he sees similar qualities in the immigrants hired by his company. "These people are determined and persistent, " Callaway said. "I've gotta give 'em credit, they'll do what it takes to get ahead. That's something I respect."

When asked his opinion, [Nikolai Vandisnky]/[Jose Sanchez], a recent immigrant from [Russia]/[Mexico], says he welcomes the chance for a better life in America. "Many of my cousins find work here and now it's my turn. I want a good job and benefits."

"But," [Vandinsky]/[Sanchez] added, "that doesn't mean I have to change who I am. We love our culture. I'm proud to be from [Russia]/[Mexico]."

While there was agreement at the convention that the federal government needs to do more to help states manage the rising tide of newcomers, few governors agree on exactly why immigration levels have increased.

Some blame the Immigration Act passed by Congress in 1990, which loosened federal restrictions on immigration. Others point to the fact that large companies are attracting immigrants to the U.S. with the promise of prosperity, a practice that has become widespread in recent years.

Still others maintain that, in a world full of turmoil, people are attracted here by the hope of a better way of life.

Whatever is bringing immigrants to these shores in record numbers, everyone seems certain that the numbers will continue to grow.

Outcomes:

- **Support for Immigration**: Do you think the number of immigrants from foreign countries who are permitted to come to the United States to live should be increased a lot, increased a little, left the same as it is now, decreased a little, or decreased a lot? (1: Decreased a lot, 5: Increased a lot)
- **Negative Impact**: In your opinion, how likely is it that immigration will have a negative financial impact on many Americans? (Very Likely, Somewhat Likely, Somewhat Unlikely, Very Unlikely) (1: Very Unlikely, 4: Very Likely)

A.11.2 Replication Study

Sample: 2,138 Mechanical Turk subjects in wave 1; 1,773 in wave 2.

Dates of data collection: July 2015.

Treatments: The treatments used the identical text as the original.

Outcomes: I used the same outcome measures with the same wording. In the second wave, I randomly varied whether subjects saw the questions as written above or in a "matrix" style format, in order to thwart subjects "learning" the right way to answer. This manipulation does not affect responses, so it is not presented below.

A.11.3 Results

In both the original and in the replication, the negative frame has a negative effect on *Support for Immigration* and a positive effect on the *Negative Impact* dependent variable. These effects appear to dissipate after 10 days.

Table A.19: Immigration Original and Replication Results

	Support for Immigration			Negative Impact		
	Original	Wave 1	Wave 2	Original	Wave 1	Wave 2
	(1)	(2)	(3)	(4)	(5)	(6)
Negative Frame	-0.561^{***}	-0.116^{*}	0.029	0.430^{***}	0.140^{**}	0.052
	(0.157)	(0.068)	(0.075)	(0.145)	(0.058)	(0.061)
Positive Frame	-0.140	0.178^{***}	0.116	0.124	-0.094^{*}	-0.080
	(0.163)	(0.066)	(0.074)	(0.147)	(0.057)	(0.061)
Constant (Control)	2.347	2.958	2.857	2.611	2.198	2.181
	(0.136)	(0.054)	(0.062)	(0.123)	(0.048)	(0.050)
Sample	Original	MTurk	MTurk	Original	MTurk	MTurk
N	352	2,138	1,773	353	2,138	1,771
R^2	0.051	0.013	0.002	0.034	0.012	0.004

$^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$
Robust standard errors are in parentheses.

A.12 Study 15: System Threat

A.12.1 Original Study

Craig and Richeson (2014) show small effects of a news story describing the US as becoming "majority-minority" on various measures of system threat.

Sample: 611 American Adults (GfK).

Dates of data collection: 5/23/2012 - 6/1/2012.

Treatments:

Majority Minority Treatment

In a Generation, Racial Minorities May Be the U.S. Majority

New U.S. Census Bureau data suggest that America will become a majority-minority nation much faster than once predicted. The nation's racial minority population is steadily rising, advancing an unmistakable trend that could make minorities the new American majority by midcentury.

The data show a declining number of White adults and growing under-18 populations of Hispanics, Asians, and other minorities. Demographers calculate that by 2042, Americans who identify themselves as Hispanic, Black, Asian, American Indian, Native Hawaiian, or Pacific Islander will together outnumber non-Hispanic Whites. The main reasons for the accelerating change are rapid immigration growth and significantly higher birthrates among racial and ethnic minorities. As White baby boomers age past their childbearing years, younger Hispanic parents are having children and driving U.S. population growth. For example, there are now roughly 9 births for every 1 death among Hispanics, compared to a roughly one-to-one ratio for Whites.

The latest figures are predicated on current and historical trends, which can be thrown awry by several variables, including prospective overhauls of public policy.

Placebo

U.S. Census Bureau Reports Residents Now Move at a Higher Rate.

New U.S. Census Bureau data suggest that the rate of geographical mobility, or the number of individuals who have moved within the past year, is increasing.The national mover rate increased from 11.9 percent in 2008 (the lowest rate since the U.S. Census Bureau began tracking the data) to 12.5 percent in 2009.

According to the new data, 37.1 million people changed residences in the U.S. within the past year. 84.5 percent of all movers stayed within the same state. Renters were more than five times more likely to move than homeowners. The estimates also reveal that many of the nation's fastest-growing cities are suburbs. Specifically, principal cities within metropolitan

areas experienced a net loss of 2.1 million movers, while the suburbs had a net gain of 2.4 million movers. For those who moved to a different county or state, the reasons for moving varied considerably by the length of their move.

The latest figures are predicated on current and historical trends, which can be thrown awry by several variables, including prospective overhauls of public policy.

Outcomes:

- **Way of Life**: Please indicate how much you agree or disagree with the following statement: The American way of life is seriously threatened. (1: Strongly Disagree, 7: Strongly Agree)
- **Support for Immigration**: Do you think the number of immigrants from foreign countries who are allowed to come to the U.S. to live should be increased a lot, increased a little, left the same as it is now, decreased a little, or decreased a lot? (1: Decreased a lot, 5: Increased a lot)

A.12.2 Replication Study

Sample: 1,276 Mechanical Turk subjects in wave 1; 1,065 in wave 2.

Dates of data collection: July 2015.

Treatments: The treatments used the identical text as the original. I also included a pure control (**No Frame**) condition.

Outcomes: I used a very slightly different wording for the immigration question: **Support for Immigration**: Do you think the number of immigrants from foreign countries who are permitted to come to the United States to live should be increased a lot, increased a little, left the same as it is now, decreased a little, or decreased a lot? For **Way of Life**, I used a 5-point scale (1: Strongly Disagree, 5: Strongly Agree). In order to make them comparable, I present effects on the **Way of Life** DV in standard units.

A.12.3 Results

The effects Majority Minority treatment could not be distinguished from zero in the original experiment, but did exert a negative effect on support for immigration in the replication.

Table A.20: System Threat Original and Replication Results

	Way of Life (Standardized)			Support for Immigration		
	Original	Wave 1	Wave 2	Original	Wave 1	Wave 2
	(1)	(2)	(3)	(4)	(5)	(6)
No Frame		0.094	0.096		−0.060	−0.089
		(0.067)	(0.075)		(0.077)	(0.084)
Majority Minority	−0.144	0.071	0.039	−0.003	−0.174**	−0.082
	(0.110)	(0.068)	(0.074)	(0.109)	(0.077)	(0.080)
Constant (Placebo)	3.156	1.723	1.679	2.398	3.018	2.946
	(0.075)	(0.046)	(0.052)	(0.076)	(0.054)	(0.057)
Sample	Original	MTurk	MTurk	Original	MTurk	MTurk
N	611	1,276	1,052	608	1,276	1,056
R^2	0.004	0.002	0.002	0.00000	0.004	0.001

*p < .1; **p < .05; ***p < .01
Robust standard errors are in parentheses.

A.13 Study 16: Expert Economists

(Johnston and Ballard 2014) found that informing the public of economists' expert opinion on various economic policies exerted a profound effect on subjects stated preferences. In five issue areas, subjects who learned of economists' views on policy were far more likely to support those positions than subjects in a control condition.

A.13.1 Original Study

Sample: 2,041 American Adults (Gfk).

Dates of data collection: 9/6/2013 - 9/23/2013.

Treatments: Subjects were assigned to respond to an economic opinion question on one of five topics. Subjects were assigned to see the "expert" or "control" versions of each question. Technically, this is a 2 x 5 factorial design, in which the first factor is whether subjects saw economists' opinions and the second factor is which economic opinion question seen. The estimand is the effect of learning economists' opinions on subjects' agreement with the economists' point of view.

Factor 1:

- **Expert**: A sample of professional economists with widely varying political preferences was asked whether they agreed or disagreed with the following statement: [Treatment Text] To what extent do you agree or disagree with this statement? (Strongly Agree, Agree, Disagree, Disagree Strongly, Uncertain)
- **Control**: To what extent do you agree or disagree with the following statement? (Strongly Agree, Agree, Disagree, Disagree Strongly, Uncertain)

Factor 2:

- **Immigration**: The average US citizen would be better off if a larger number of highly educated foreign workers were legally allowed to immigrate to the US each year. (Economists: Strongly Agree: 49, Agree: 46)
- **Health Care**: Long run fiscal sustainability in the U.S. will require cuts in currently promised Medicare and Medicaid benefits and/or tax increases that include higher taxes on households with incomes below $250,000. (Economists: Strongly Agree: 56, Agree: 35)
- **Trade with China**: Trade with China makes most Americans better off because, among other advantages, they can buy goods that are made or assembled more cheaply in China. (Economists: Strongly Agree: 59, Agree: 41)

- **Tax Cut**: A cut in federal income tax rates in the US right now would raise taxable income enough so that the annual total tax revenue would be higher within five years than without the tax cut. (Economists: Strongly Disagree: 57, Disagree: 39)
- **Gold Standard**: If the US replaced its discretionary monetary policy regime with a gold standard, defining a "dollar" as a specific number of ounces of gold, the price-stability and employment outcomes would be better for the average American. (Economists: Strongly Disagree: 66, Disagree: 34)

Outcomes:

- **Agree**: For each question, the agreement dependent variable was coded 1 if the subject chose either "Strongly Agree" or "Agree" when the economists did or "Strongly Disagree" or "Disagree" when the economists did.

A.13.2 Replication Study

Sample: 2,985 Mechanical Turk subjects in wave 1; 2,487 in wave 2.

Dates of data collection: July 2015.

Treatments: Instead of assigning subjects to see only one question, I had subjects see all five questions, randomizing within question whether a subject saw the "control" or "expert" version. The questions were otherwise identical.

Outcomes: I defined agreement in the same way as in the original.

A.13.3 Results

The results of the original and replication experiments are in close agreement. The Expert treatment increases agreement with the economists' point of view on every issue. This strong effect evaporates within 10 days.

Table A.21: Expert Economists: Immigration and Health Care

	Agree on Immigration			Agree on Health Care		
	Original	Wave 1	Wave 2	Original	Wave 1	Wave 2
	(1)	(2)	(3)	(4)	(5)	(6)
Expert Treatment	0.069	0.163***	0.020	0.075	0.172***	0.015
	(0.057)	(0.018)	(0.020)	(0.050)	(0.018)	(0.019)
Constant (Control)	0.249	0.387	0.454	0.220	0.345	0.323
	(0.038)	(0.013)	(0.014)	(0.034)	(0.012)	(0.013)
Sample	Original	MTurk	MTurk	Original	MTurk	MTurk
N	413	2,985	2,487	409	2,985	2,487
R^2	0.006	0.027	0.0004	0.007	0.030	0.0002

*p < .1; **p < .05; ***p < .01
Robust standard errors are in parentheses.

Table A.22: Expert Economists: China and Tax Cuts

	Agree on Trade with China			Agree on Tax Cut		
	Original	Wave 1	Wave 2	Original	Wave 1	Wave 2
	(1)	(2)	(3)	(4)	(5)	(6)
Expert Treatment	0.122**	0.209***	0.044**	0.176***	0.218***	−0.013
	(0.052)	(0.018)	(0.020)	(0.051)	(0.018)	(0.019)
Constant (Control)	0.207	0.440	0.420	0.148	0.376	0.360
	(0.033)	(0.013)	(0.014)	(0.032)	(0.012)	(0.014)
Sample	Original	MTurk	MTurk	Original	MTurk	MTurk
N	408	2,985	2,487	403	2,985	2,487
R^2	0.019	0.044	0.002	0.043	0.048	0.0002

*p < .1; **p < .05; ***p < .01
Robust standard errors are in parentheses.

Table A.23: Expert Economists: Gold Standard

	Agree on Gold Standard		
	Original	Wave 1	Wave 2
	(1)	(2)	(3)
Expert Treatment	0.122**	0.209***	0.044**
	(0.052)	(0.018)	(0.020)
Constant (Control)	0.207	0.440	0.420
	(0.033)	(0.013)	(0.014)
Sample	Original	MTurk	MTurk
N	408	2,985	2,487
R^2	0.019	0.044	0.002

*p < .1; **p < .05; ***p < .01
Robust standard errors are in parentheses.

A.14 Study 17: Mental Health

(McGinty et al. 2013) show that a news article about a mass shooting affects support for stricter gun control laws as well as perceptions of how dangerous people with serious mental illness are.

A.14.1 Original Study

Sample: 2,935 American Adults (GfK).

Dates of data collection: May 2012.

Treatments:

News The gunman who opened fire in an Indianapolis park yesterday morning has been identified as Indianapolis resident Jake Robinson, age 30. According to police, the shooter has a history of serious mental illness. Mr. Robinson's motivation for opening fire in Smith Park in central Indianapolis is unclear. Witnesses said Mr. Robinson arrived at the park around seven-thirty am and appeared agitated, pacing up and down and talking to himself.

At approximately 8:15 am, Mr. Robinson took a gun out of his bag and began to shoot. Three adults passing through the park on their way to work were shot and killed. Three more adults and two children were wounded. The police officer leading the investigation said that Jake Robinson used a semiautomatic weapon to shoot about 30 bullets in a row before he was tackled by a security guard from a nearby building. Little is known about Mr. Robinson, who lived alone and appears to have no immediate family. Mr. Robinson's cousin, who lives in South Carolina, said Mr. Robinson was hospitalized for mental illness last year.

LCM Ban Yesterday's shooting in downtown Indianapolis left residents looking for solutions to the problem of gun violence. According to the Indianapolis Coalition against Violence – a group whose membership includes city lawmakers, law enforcement officials, researchers, advocacy groups and citizens concerned about violence in Indianapolis – gun violence in the United States has reached epidemic proportions.

"With more than 65,000 Americans shot in an attack last year, we have to do something to keep dangerous guns off our streets," said Kim Jones, the spokesperson for the group. One proposal currently being considered by Congress is a good start, Jones said. Congress is considering legislation to ban large ammunition clips, which are military-style high capacity magazines that can shoot 30, 50, or 100 bullets without requiring the shooter to stop and reload. According to Kim Jones, "Getting this law in place is one way to protect the public from dangerous guns."

Mental Illness Yesterday's shooting in downtown Indianapolis left residents looking for solutions to the problem of gun violence. According to the Indianapolis Coalition against Violence – a group whose membership includes city lawmakers, law enforcement officials, researchers, advocacy groups and citizens concerned about violence in Indianapolis – gun violence in the United States has reached epidemic proportions.

"With more than 65,000 Americans shot in an attack last year, we have to do something to keep guns out of the hands of dangerous people," said Kim Jones, the spokesperson for the group. One proposal currently being considered by Congress is a good start, Jones said. Congress is considering legislation to require states to enter people with serious mental illness into a background check system used by gun dealers to identify people prohibited from buying guns, or face a penalty. According to Kim Jones, "Getting this law in place is one way to protect the public from dangerous people."

Outcomes:

- **Magazines**: As you may know, high-capacity gun magazines or clips can hold many rounds of ammunition, so a shooter can fire more rounds without manually reloading. Would you support or oppose a nationwide ban on the sale of high-capacity gun magazines that hold more than 10 rounds of ammunition? (Strongly Oppose: 1; Strongly Support: 5)
- **SMI Danger**: Do you agree with the following statement? "People with serious mental illness are, by far, more dangerous than the general population." (Strongly Disagree: 1; Strongly Agree: 5)

A.14.2 Replication Study

Sample: 2,985 Mechanical Turk subjects in wave 1; 2,474 in wave 2.

Dates of data collection: July 2015.

Treatments: The treatments used the identical text as the original.

Outcomes: The outcomes were the same as in the original.

A.14.3 Results

In the original study, the News treatment increased subjects support for a ban on high-capacity magazines as well as perceptions of the mentally ill as dangerous. These effects were not apparent in the replication. The effects of the LCM Ban treatment on support for a high capacity magazine ban were consistent across the original and the replication, and endured for at least 10 days.

Table A.24: Mental Illness Original and Replication Results

	Magazines			SMI Danger		
	Original	Wave 1	Wave 2	Original	Wave 1	Wave 2
	(1)	(2)	(3)	(4)	(5)	(6)
News	0.208***	0.054	0.004	0.147***	0.027	−0.018
	(0.055)	(0.054)	(0.059)	(0.040)	(0.042)	(0.044)
LCM Ban	0.315***	0.129**	0.170**	−0.035	0.001	−0.044
	(0.068)	(0.066)	(0.071)	(0.049)	(0.051)	(0.053)
Mental Illness	0.138**	−0.062	0.008	−0.061	−0.023	−0.025
	(0.068)	(0.068)	(0.074)	(0.049)	(0.051)	(0.054)
Constant (Control)	3.308	3.543	3.577	3.251	3.133	3.050
	(0.057)	(0.054)	(0.059)	(0.039)	(0.042)	(0.044)
Sample	Original	MTurk	MTurk	Original	MTurk	MTurk
N	2,935	2,985	2,474	2,933	2,985	2,474
R^2	0.012	0.003	0.003	0.005	0.0002	0.0004

*p < .1; **p < .05; ***p < .01
Robust standard errors are in parentheses.

A.15 Study 18: Contentious Global Warming

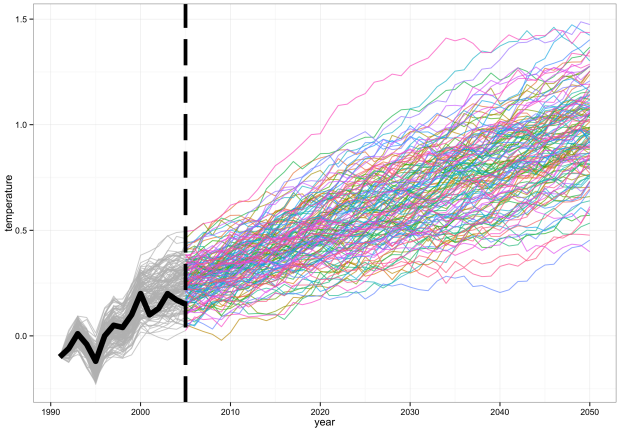
This study was conducted in collaboration with Andy Guess and was designed to measure subjects' ability to process scientific information about global warming under contentious conditions.

Sample: 2,156 Mechanical Turk Subjects in wave 1; 1,802 in wave 2.
Dates of data collection: June 2015.
This study used a pre-treatment measure of support: How strongly do you believe in the scientific evidence that global temperatures are rising due to human activity? (Very strongly, Not very strongly). Subjects who select "Very strongly" are coded as "proponents" while those who respond "Not very strongly" are coded as "opponents".
Treatments: Subjects could be assigned to a control or "hiatus" treatment. Additionally, subjects were assigned to a control processing condition or an "insults" processing condition.
Control Global warming refers to the increase in the Earth's temperature over the last century. Since 1900, temperatures have increased by 1.4 degrees Fahrenheit. Scientists, politicians, and citizens have been debating over what, if anything, to do to combat global warming.

In order to understand what to do about global warming, we need to know more than how much the Earth has warmed in the past. We need to know how much the Earth will warm in the future.

Climate scientists try to predict how much warming will occur by creating statistical models based on historical data, then using those models to predict future temperatures. Because any one model can be wrong, the UN's Intergovernmental Panel on Climate Change (IPCC) averages many forecasts, as shown in the figure below.

Figure A.4: Main Manipulation: Control



The colored lines show the models' best guesses for the changes in temperature. The

black line shows the actual observed temperature.

There is a great deal of agreement between the models and the observed temperature. This agreement means that we can have greater confidence in the models' projections into the future.

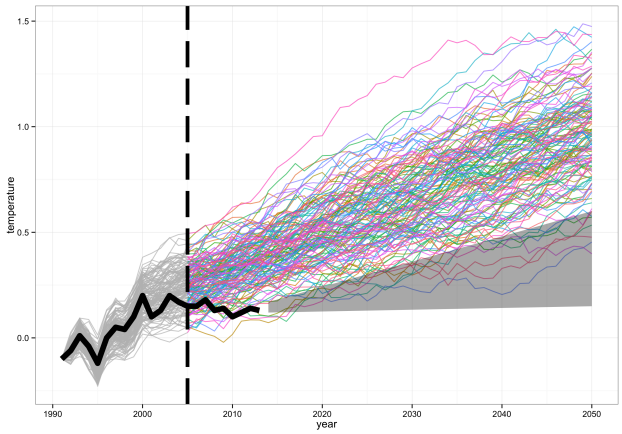
The Earth is warming, and the models predict that the warming trends will continue into the future.

Hiatus Global warming refers to the increase in the Earth's temperature over the last century. Since 1900, temperatures have increased by 1.4 degrees Fahrenheit. Scientists, politicians, and citizens have been debating over what, if anything, to do to combat global warming.

In order to understand what to do about global warming, we need to know more than how much the Earth has warmed in the past. We need to know how much the Earth will warm in the future.

Climate scientists try to predict how much warming will occur by creating statistical models based on historical data, then using those models to predict future temperatures. Because any one model can be wrong, the UN's Intergovernmental Panel on Climate Change (IPCC) averages many forecasts, as shown in the figure below.

Figure A.5: Main Manipulation: Hiatus Treatment



The colored lines show the models' best guesses for the changes in temperature. The solid black line shows the actual observed temperature.

The dotted line represents the moment in time when the predictions were made. To the left of the dotted line, the models and the observed temperature agree. But to the right of the dotted line, the actual levels of warming clearly disagree with the predictions of the models.

This disagreement means we cannot be confident of the models' projections into the future. The shaded region describes the updated projections based on more recent data.

Processing Conditions:

- Insults: (if "Very Strongly") Some people think that your position makes you an alarmist supporter of big government. How do you respond? [Text entry]. If ("Not very strongly") Some people think that your position makes you an unthinking opponent of science. How do you respond? [Text entry].
- Control. In this condition, subjects proceeded directly to their information treatment.

Outcomes:

- **Belief in Increase**: Since 1850, there is a well-documented increase in global temperatures. Some attribute this increase to human activity while others attribute it to natural causes. Which comes closest to your view? (1: I believe that the increase is entirely due to natural causes, 7: I believe that the increase is entirely due to human activity)
- **Degrees**: How many degrees do you believe the Earth will warm over the next 100 years? (Slider: -1 to 3)
- **Temp Cause**: Some people believe that increases in the Earth's temperature over the last century are due to the effects of pollution from human activities. Others believe that temperature increases are due to natural changes in the environment. Which comes closer to your view? (Temperature increases are mainly due to the effects of pollution from human activities, Temperature increases are mainly due to natural changes in the environment, Temperature increases are due to both human activities and natural changes, I do not believe the Earth's temperature has risen.) The outcome **Natural Cause** is coded 1 if the respondent chose "Temperature increases are mainly due to natural changes in the environment" or "Temperature increases are due to both human activities and natural changes" and 0 otherwise.

A.15.1 Results

The *Hiatus* treatment decreased subjects' belief that global warming was due to human causes and decreased subjects' predictions about future temperature increases.

Table A.25: Contentious Global Warming Results

	Belief in Increase		Degrees		Natural Cause	
	Wave 1	Wave 2	Wave 1	Wave 2	Wave 1	Wave 2
	(1)	(2)	(3)	(4)	(5)	(6)
Hiatus Treatment	−0.173*	−0.169	−0.114**	−0.097*	0.020	0.046
	(0.094)	(0.104)	(0.049)	(0.053)	(0.030)	(0.033)
Insult	−0.080	−0.019	0.016	0.004	0.018	0.022
	(0.094)	(0.103)	(0.049)	(0.053)	(0.030)	(0.033)
Constant (Control)	5.100	5.095	1.739	1.827	0.511	0.471
	(0.080)	(0.090)	(0.042)	(0.044)	(0.026)	(0.028)
N	1,082	911	1,082	911	1,082	912
R^2	0.004	0.003	0.005	0.004	0.001	0.003

*p < .1; **p < .05; ***p < .01
Robust standard errors are in parentheses.

A.16 Study 19: Newspapers

This study was conducted in collaboration with Emily Ekins and David Kirby. It shows that when subjects are shown newspaper opinion pieces, they update their attitudes and beliefs in the direction of the opinion piece.

Sample: 3,567 Mechanical Turk Subjects in wave 1; 2,989 in wave 2; 2,448 in wave 3.

Dates of data collection: July 2015.

A.16.1 Treatments

Subjects could be assigned to see either nothing or one of five op-ed pieces. The text of each treatment is below:

- Control
- Amtrak
- Veterans
- Climate
- Wall Street
- Flat Tax

Newsweek

The Amtrak Crash: Is More Spending the Answer?

By Randal O'Toole 5/13/15 At 4:46 PM

It is too soon to tell what caused the Amtrak train crash that killed seven people on May 12. But advocates of increased government spending are already beginning to use the crash to promote more spending on infrastructure and are criticizing Republicans who voted to reduce Amtrak's budget the day after the crash.

Yet there is a flaw in the assumption that spending more money would result in better infrastructure. In fact, in some cases, the problem is that too much money is being spent on infrastructure, but in the wrong places.

The reason for this is that politicians prefer to spend money building new infrastructure over maintaining the old. The result is that existing infrastructure that depends on tax dollars steadily declines while any new funds raised for infrastructure tend to go to new projects.

We can see this in the Boston, Washington, and other rail transit systems. Boston's system is $9 billion in debt, has a $3 billion maintenance backlog, and needs to spend nearly $700 million a year just to keep the backlog from growing. Yet has only budgeted $100 million for maintenance this year, and instead of repairing the existing system, Boston is spending $2 billion extending one of its light-rail lines.

Similarly, Washington's Metro rail system has a $10 billion maintenance backlog, and poor maintenance was the cause of the 2009 wreck that killed nine people. Yet, rather than rehabilitate their portions of the system, Northern Virginia is spending $6.8 billion building a new rail line to Dulles Airport; D.C. wants to spend $1 billion on new streetcar lines; and Maryland is considering building a $2.5 billion light-rail line in D.C. suburbs.

On the other hand, infrastructure that is funded out of user fees is generally in good shape. Despite tales of crumbling bridges, the 2007 Minnesota bridge collapse was due to a construction flaw and the 2013 Washington state bridge collapse was due to an oversized truck; lack of maintenance had nothing to do with either failure.

Department of Transportation numbers show that the number of bridges considered structurally deficient has fallen by more than 50 percent since 1990, while the average roughness of highway pavement has decreased. State highways and bridges, which are almost entirely funded out of user fees, tend to be in the best condition while local highways and bridges, which depend more on tax dollars, tend to be the ones with the most serious problems.

Before 1970, almost all of our transportation infrastructure was funded out of user fees and the United States had the best transportation system in the world. Since then, funding decisions have increasingly been made by politicians who are more interested in getting

their pictures taken cutting ribbons than in making sure our transportation systems run safely and smoothly.

Proponents of higher gas taxes and other increased funding on infrastructure may talk about crumbling bridges, but what they really want is to spend more money on new projects that are often of little value. For example, they want high-speed trains that cost more but go less than half the speed of flying and light-rail trains that cost more but can move fewer people than buses.

This country doesn't need more infrastructure that it can't afford to maintain. Instead, it needs a more reliable system of transport funding, and that means one based on user fees and not tax subsidies.

Randal O'Toole is a senior fellow with the Cato Institute and author of Gridlock: Why We're Stuck in Traffic and What to Do About It.

THE WALL STREET JOURNAL.

The Political Assault on Climate Skeptics

Members of Congress send inquisitorial letters to universities, energy companies, even think tanks. By Richard S. Lindzen March 4, 2015 6:50 P.M. ET

Research in recent years has encouraged those of us who question the popular alarm over allegedly man-made global warming. Actually, the move from "global warming" to "climate change" indicated the silliness of this issue. The climate has been changing since the Earth was formed. This normal course is now taken to be evidence of doom.

Individuals and organizations highly vested in disaster scenarios have relentlessly attacked scientists and others who do not share their beliefs. The attacks have taken a threatening turn.

As to the science itself, it's worth noting that all predictions of warming since the onset of the last warming episode of 1978-98 – which is the only period that the United Nations Intergovernmental Panel on Climate Change (IPCC) attempts to attribute to carbon-dioxide emissions – have greatly exceeded what has been observed. These observations support a much reduced and essentially harmless climate response to increased atmospheric carbon dioxide.

In addition, there is experimental support for the increased importance of variations in solar radiation on climate and a renewed awareness of the importance of natural unforced climate variability that is largely absent in current climate models. There also is observational evidence from several independent studies that the so-called "water vapor feedback," essential to amplifying the relatively weak impact of carbon dioxide alone on Earth temperatures, is canceled by cloud processes.

There are also claims that extreme weather – hurricanes, tornadoes, droughts, floods, you name it – may be due to global warming. The data show no increase in the number or intensity of such events. The IPCC itself acknowledges the lack of any evident relation between extreme weather and climate, though allowing that with sufficient effort some relation might be uncovered.

World leaders proclaim that climate change is our greatest problem, demonizing carbon dioxide. Yet atmospheric levels of carbon dioxide have been vastly higher through most of Earth's history. Climates both warmer and colder than the present have coexisted with these higher levels.

Currently elevated levels of carbon dioxide have contributed to increases in agricultural productivity. Indeed, climatologists before the recent global warming hysteria referred to warm periods as "climate optima." Yet world leaders are embarking on costly policies that have no capacity to replace fossil fuels but enrich crony capitalists at public expense, increasing costs for all, and restricting access to energy to the world's poorest populations that still lack access to electricity's immense benefits.

Billions of dollars have been poured into studies supporting climate alarm, and trillions of dollars have been involved in overthrowing the energy economy. So it is unsurprising that great efforts have been made to ramp up hysteria, even as the case for climate alarm is disintegrating.

The latest example began with an article published in the New York Times on Feb. 22 about Willie Soon, a scientist at the Harvard Smithsonian Center for Astrophysics. Mr. Soon has, for over 25 years, argued for a primary role of solar variability on climate. But as Greenpeace noted in 2011, Mr. Soon was, in small measure, supported by fossil-fuel companies over a period of 10 years.

The Times reintroduced this old material as news, arguing that Mr. Soon had failed to list this support in a recent paper in Science Bulletin of which he was one of four authors. Two days later Arizona Rep. Raul Grijalva, the ranking Democrat on the Natural Resources Committee, used the Times article as the basis for a hunting expedition into anything said, written and communicated by seven individuals – David Legates, John Christy, Judith Curry, Robert Balling, Roger Pielke Jr., Steven Hayward and me – about testimony we gave to Congress or other governmental bodies. We were selected solely on the basis of our objections to alarmist claims about the climate.

In letters he sent to the presidents of the universities employing us (although I have been retired from MIT since 2013), Mr. Grijalva wanted all details of all of our outside funding, and communications about this funding, including "consulting fees, promotional considerations, speaking fees, honoraria, travel expenses, salary, compensation and any other monies." Mr. Grijalva acknowledged the absence of any evidence but purportedly wanted to know if accusations made against Mr. Soon about alleged conflicts of interest or failure to disclose his funding sources in science journals might not also apply to us.

Perhaps the most bizarre letter concerned the University of Colorado's Mr. Pielke. His specialty is science policy, not science per se, and he supports reductions in carbon emissions but finds no basis for associating extreme weather with climate. Mr. Grijalva's complaint is that Mr. Pielke, in agreeing with the IPCC on extreme weather and climate, contradicts the assertions of John Holdren, President Obama's science czar.

Mr. Grijalva's letters convey an unstated but perfectly clear threat: Research disputing alarm over the climate should cease lest universities that employ such individuals incur massive inconvenience and expense – and scientists holding such views should not offer testimony to Congress. After the Times article, Sens. Edward Markey (D., Mass.), Sheldon Whitehouse (D., R.I.) and Barbara Boxer (D., Calif.) also sent letters to numerous energy companies, industrial organizations and, strangely, many right-of-center think tanks (including the Cato Institute, with which I have an association) to unearth their alleged influence peddling.

The American Meteorological Society responded with appropriate indignation at the singling out of scientists for their scientific positions, as did many individual scientists. On Monday, apparently reacting to criticism, Mr. Grijalva conceded to the National Journal that his requests for communications between the seven of us and our outside funders was "overreach."

Where all this will lead is still hard to tell. At least Mr. Grijalva's letters should help clarify for many the essentially political nature of the alarms over the climate, and the damage it is doing to science, the environment and the well-being of the world's poorest.

Mr. Lindzen is professor emeritus of atmospheric sciences at MIT and a distinguished senior fellow of the Cato Institute.

THE WALL STREET JOURNAL.

Blow Up the Tax Code and Start Over

Apply a 14.5% flat tax to personal income and to businesses. Cut deductions. Watch the economy roar.

By Rand Paul June 17, 2015 7:09 P.M. ET

Some of my fellow Republican candidates for the presidency have proposed plans to fix the tax system. These proposals are a step in the right direction, but the tax code has grown so corrupt, complicated, intrusive and antigrowth that I've concluded the system isn't fixable.

So on Thursday I am announcing an over $2 trillion tax cut that would repeal the entire IRS tax code – more than 70,000 pages – and replace it with a low, broad-based tax of 14.5% on individuals and businesses. I would eliminate nearly every special-interest loophole. The plan also eliminates the payroll tax on workers and several federal taxes outright, including gift and estate taxes, telephone taxes, and all duties and tariffs. I call this "The Fair and Flat Tax."

President Obama talks about "middle-class economics," but his redistribution policies have led to rising income inequality and negative income gains for families. Here's what I propose for the middle class: The Fair and Flat Tax eliminates payroll taxes, which are seized by the IRS from a worker's paychecks before a family ever sees the money. This will boost the incentive for employers to hire more workers, and raise after-tax income by at least 15% over 10 years.

Here's why we have to start over with the tax code. From 2001 until 2010, there were at least 4,430 changes to tax laws – an average of one "fix" a day – always promising more fairness, more simplicity or more growth stimulants. And every year the Internal Revenue Code grows absurdly more incomprehensible, as if it were designed as a jobs program for accountants, IRS agents and tax attorneys.

Polls show that "fairness" is a top goal for Americans in our tax system. I envision a traditionally All-American solution: Everyone plays by the same rules. This means no one of privilege, wealth or with an arsenal of lobbyists can game the system to pay a lower rate than working Americans.

Most important, a smart tax system must turbocharge the economy and pull America out of the slow-growth rut of the past decade. We are already at least $2 trillion behind where we should be with a normal recovery; the growth gap widens every month. Even Mr. Obama's economic advisers tell him that the U.S. corporate tax code, which has the highest rates in the world (35%), is an economic drag. When an iconic American company like Burger King wants to renounce its citizenship for Canada because that country's tax rates are so much lower, there's a fundamental problem.

Another increasingly obvious danger of our current tax code is the empowerment of a rogue agency, the IRS, to examine the most private financial and lifestyle information of every American citizen. We now know that the IRS, through political hacks like former IRS official Lois Lerner, routinely abused its auditing power to build an enemies list and

harass anyone who might be adversarial to President Obama's policies. A convoluted tax code enables these corrupt tactics.

My tax plan would blow up the tax code and start over. In consultation with some of the top tax experts in the country, including the Heritage Foundation's Stephen Moore, former presidential candidate Steve Forbes and Reagan economist Arthur Laffer, I devised a 21st-century tax code that would establish a 14.5% flat-rate tax applied equally to all personal income, including wages, salaries, dividends, capital gains, rents and interest. All deductions except for a mortgage and charities would be eliminated. The first $50,000 of income for a family of four would not be taxed. For low-income working families, the plan would retain the earned-income tax credit.

I would also apply this uniform 14.5% business-activity tax on all companies – down from as high as nearly 40% for small businesses and 35% for corporations. This tax would be levied on revenues minus allowable expenses, such as the purchase of parts, computers and office equipment. All capital purchases would be immediately expensed, ending complicated depreciation schedules.

The immediate question everyone asks is: Won't this 14.5% tax plan blow a massive hole in the budget deficit? As a senator, I have proposed balanced budgets and I pledge to balance the budget as president.

Here's why this plan would balance the budget: We asked the experts at the nonpartisan Tax Foundation to estimate what this plan would mean for jobs, and whether we are raising enough money to fund the government. The analysis is positive news: The plan is an economic steroid injection. Because the Fair and Flat Tax rewards work, saving, investment and small business creation, the Tax Foundation estimates that in 10 years it will increase gross domestic product by about 10%, and create at least 1.4 million new jobs.

And because the best way to balance the budget and pay down government debt is to put Americans back to work, my plan would actually reduce the national debt by trillions of dollars over time when combined with my package of spending cuts.

The left will argue that the plan is a tax cut for the wealthy. But most of the loopholes in the tax code were designed by the rich and politically connected. Though the rich will pay a lower rate along with everyone else, they won't have special provisions to avoid paying lower than 14.5%.

The challenge to this plan will be to overcome special-interest groups in Washington who will muster all of their political muscle to save corporate welfare. That's what happened to my friend Steve Forbes when he ran for president in 1996 on the idea of the flat tax. Though the flat tax was surprisingly popular with voters for its simplicity and its capacity to boost the economy, crony capitalists and lobbyists exploded his noble crusade.

Today, the American people see the rot in the system that is degrading our economy day after day and want it to end. That is exactly what the Fair and Flat Tax will do through a plan that's the boldest restoration of fairness to American taxpayers in over a century.

Sen. Paul, a Republican from Kentucky, is running for his party's presidential nomination.

The New York Times

The Other Veterans Scandal

By Michael F. Cannon And Christopher A. Preble June 15, 2014

Washington The Department of Veterans Affairs is mired in scandal. More than 57,000 veterans have been waiting at least three months for a doctor's appointment. Another 64,000 never even made it onto a waiting list. There are allegations that waits for care either caused or contributed to veterans' deaths.

But another, even larger problem with the Department of Veterans Affairs is being overlooked: Even when the department works exactly as intended, it helps inflict great harm on veterans, active-duty military personnel and civilians.

Here's how. Veterans' health and disability benefits are some of the largest costs involved in any military conflict, but they are delayed costs, typically reaching their peak 40 or 50 years after the conflict ends. Congress funds these commitments – through the Department of Veterans Affairs – only once they come due.

As a result, when Congress debates whether to authorize and fund military action, it can act as if those costs don't exist. But concealing those costs makes military conflicts appear less burdensome and therefore increases their likelihood. It's as if Congress deliberately structured veterans' benefits to make it easier to start wars.

The Department of Veterans Affairs is supposed to help wounded veterans, but its current design makes soldiers more likely to get killed or injured in the first place. The scandal isn't at the Department of Veterans Affairs. The scandal is the Department of Veterans Affairs.

Is there a better way? We propose a system of veterans' benefits that would be funded by Congress in advance. It would allow veterans to purchase life, disability and health insurance from private insurers. Those policies would cover losses related to their term of service, and would pay benefits when they left active duty through the remainder of their lives.

To cover the cost, military personnel would receive additional pay sufficient to purchase a statutorily defined package of benefits at actuarially fair rates. The precise amount would be determined with reference to premiums quoted by competing insurers, and would vary with the risks posed by particular military jobs.

Insurers and providers would be more responsive because veterans could fire them – something they cannot do to the Department of Veterans Affairs. Veterans' insurance premiums would also reveal, and enable recruits and active-duty personnel to compare, the risks posed by various military jobs and career paths.

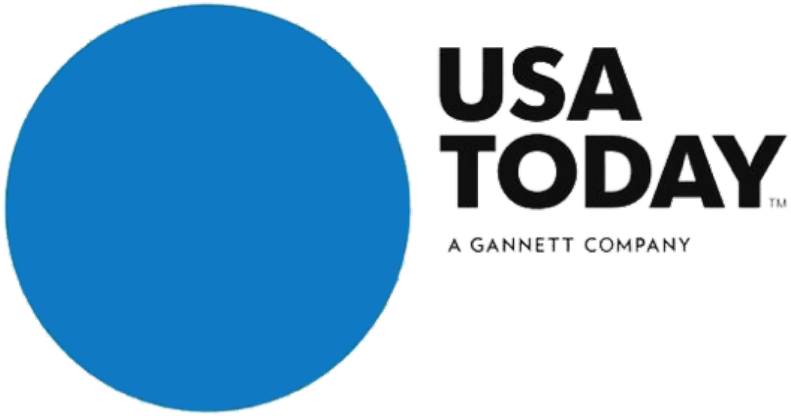
Most important, under this system, when a military conflict increases the risk to life and limb, insurers would adjust veterans' insurance premiums upward, and Congress would have to increase military pay immediately to enable military personnel to cover those added costs.

Consider how this system might have prevented Congress's misbegotten decision to authorize President George W. Bush to invade Iraq. In 2002, the Bush administration played down estimates that the war would cost as much as $200 billion, insisting the cost would be less than $50 billion. To give you a sense of how mistaken this was: The economists Linda J. Bilmes and Joseph E. Stiglitz put the cost of veterans' benefits alone, from the wars in Iraq and Afghanistan, at roughly $1 trillion.

Like others before her, Hillary Rodham Clinton has admitted that voting to authorize the Iraq invasion was a "mistake," though she made "the best decision I could with the information I had." How many members of Congress would have voted differently if confronted with the long-term health and disability needs of the troops they had already sent into Afghanistan and those they were sending into Iraq? How many would have pressed harder to end the wars sooner if they had to confront the mounting cost of veterans' benefits, in addition to the wars' other growing costs, every year the wars dragged on?

The alternative system we propose combines the universal goal of improving veterans' benefits with conservative Republicans' preference for market incentives and antiwar Democrats' desire to make it harder to wage war. Pre-funding veterans benefits could prevent unnecessary wars, or at least end them sooner. We can think of no greater tribute to the men and women serving in our armed forces.

Michael F. Cannon is the director of health policy studies, and Christopher Preble is the vice president for defense and foreign policy studies, at the Cato Institute.



Wall Street Offers Very Real Benefits: Opposing view

But headlines focus on the bad behavior.

By Thaya Knight 7:16 pm May 26, 2015

Not every person on Wall Street is a morally corrupt Gordon Gekko. Do Wall Street traders want to make money? Yes. Are they generally people who thrive in a fast-paced, competitive environment? You bet. And that is a good thing.

At its core, here's what Wall Street does: It makes sure that companies doing useful things get the money they need to keep doing those things. Do you like your smartphone? Does it make your life easier? The company that made that phone got the money to develop the product and get it into the store where you bought it with the help of Wall Street.

When a company wants to expand, or make a new product, or improve its old products, it needs money, and it often gets that money by selling stock or bonds. That helps those companies, the broader economy and consumers generally.

When we have flashing headlines about Wall Street traders acting badly, as we had last week with news of five major banks pleading guilty to criminal charges, it is very easy to hate Wall Street. But we only hear headlines about the worst behavior.

No one writes news stories about traders who go about their business every day, carefully complying with the many (and there are many) rules and regulations that govern their work. Also, the financial sector, which is usually what people mean when they say "Wall Street," isn't only or even mostly the big banks.

There are small firms, banks, funds and advisers that make up a large portion of our financial industry. While the news about corruption, corporate welfare and lawbreaking is very bad, it doesn't mean the entire industry is rotten. Or, more important, that we don't need it.

Wall Street could be better. We could eliminate regulations that crowd out competition for the big banks. We could reform the system to do away with "too big to fail," making it harder for bad traders to get away with bad behavior. Either way, we shouldn't lose sight of the very real economic and social benefits Wall Street provides.

Thaya Knight is associate director of financial regulation studies at the Cato Institute.

A.16.2 Outcomes

I asked a series a outcome questions on each of the five issue areas.

Amtrak Outcomes:

- Do you think the government should spend more, less, or about what it does now on transportation and infrastructure? [1: A lot more, A lot less]
- Would you prefer government pay for building and maintaining roads and infrastructure through raising taxes for transportation spending, or through charging user-fees, like paying tolls when you drive on the highways? [1: Fund entirely through tax increases, 4: Both Equally, 7; Fund entirely through user fees]
- If the government raised taxes to pay for more transportation spending, do you expect that money would primarily go toward building new infrastructure projects or maintaining and improving existing infrastructure? [1: Entirely toward NEW infrastructure projects, 4: Both Equally, 7: Entirely toward maintaining EXISTING infrastructure]
- For every dollar the government spends on transportation and infrastructure projects, about how many cents do you think are spent inefficiently? [Slider 0 - 100, How Many Cents Spent Inefficiently?]

Climate Outcomes:

- Would you say that climate change is best described as a ... (1: Crisis, 7: Not a problem at all)
- From what you've read and heard, do you believe increases in Earth's temperature are due... (1: Entirely due to the effects of pollution from human activity, 7: Entirely due to natural causes)
- Do you think the solution to the climate change problem will primarily come from government policies or technological innovation in the free market? (1: Entirely from the free market, 7: Entirely from government policies)
- Thinking about what's in the news, is the seriousness of global warming generally exaggerated, correct, or underestimated? (1: Generally exaggerated, 4: Generally Correct, 7: Generally underestimated)
- How many degrees (Fahrenheit) do you believe the Earth will warm over the next 100 years? (Select "0" if you think the temperature will stay about the same) [Slider -3 to 3]

Flat Tax Outcomes:

- Would you favor or oppose changing the federal tax system to a flat tax, where everyone making more than $50,000 a year pays the same percentage of his or her income in taxes? [1: Strongly Favor, 7: Strongly Oppose]
- What percentage of income, from 0 to 100, do you think Americans should pay in federal taxes on average? [Slider 0 - 100, Average Tax Rate]
- Do you favor or oppose reducing the business and corporate tax rate to 14.5% percent? [1: Strongly Favor, 7: Strongly Oppose]
- Do you think a flat tax on incomes over $50,000 without tax deductions or credits will do more to help all Americans or do more to help wealthy Americans? [1: Do more to help ALL Americans, 7: Do more to help WEALTHY Americans]

Veterans Outcomes:

- How would you rate your feelings toward the Department of Veterans Affairs (the VA) on a scale of 0 to 100, where a rating of 100 means you feel as warm and positive as possible and 0 means you feel as cold and negative as possible? How do you feel toward... [Department of Veterans Affairs]
- How much confidence do you have in the Department of Veterans Affairs' ability to care for veterans? [1: A Great Deal, 7: None At All]
- Would you favor or oppose changing the healthcare system for Veterans to a system where the government provides additional money sufficient for Veterans to purchase a government-approved health insurance plan from private health insurance companies? [1: Strongly Favor, 7: Strongly Oppose]
- For every dollar the government spends on Veterans Benefits, about how many cents do you think are spent inefficiently? [Slider 0 - 100, How Many Cents Spent Inefficiently?]

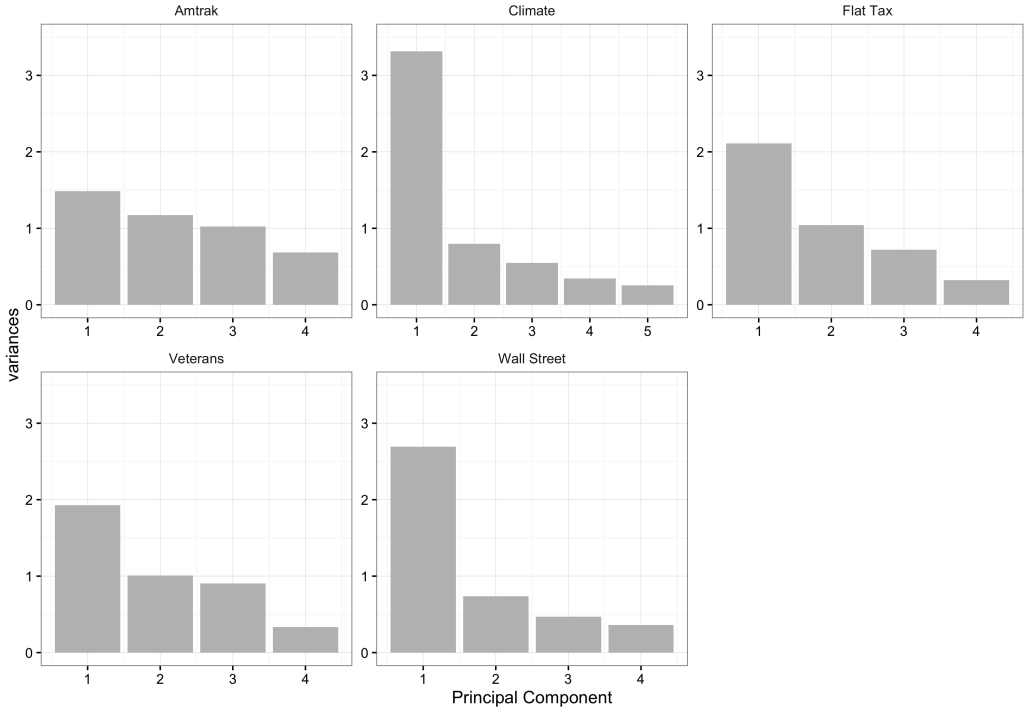
Wall Street Outcomes:

- How would you rate your feelings toward the following on a scale of 0 to 100, where a rating of 100 means you feel as warm and positive as possible and 0 means you feel as cold and negative as possible. How do you feel toward... [CEOs; Wall Street Bankers; Government Regulators]
- What percentage of Wall Street bankers, from zero to one hundred, do you think are corrupt? [Slider 0 - 100: % Wall Street Bankers Corrupt]
- How much confidence do you have in Wall Street bankers and brokers to do the right thing... [1: A Great Deal, 7: None at all]

- Compared to what it's doing now, do you think the federal government needs to regulate banks and financial institutions [1: A lot more, A lot less]

For each outcome area, I created a scale using principal components analysis, taking the first factor. The scree plots shown in Figure A.6 show that with the exception of the "Amtrak" set of dependent variables, a large portion of the variance in each issue area is accounted for by the first factor.

Figure A.6: Scree Plots for Five Issue Areas



A.16.3 Results

The tables below show the effects of each treatment on each outcome scale, in each wave. The effects are all on the diagonal – op-eds affect their target issue areas and (usually) not the off-issue areas. This pattern holds in all three waves.

Table A.26: Newspapers Wave 1

	Amtrak	Climate	Flat Tax	Veterans	Wallstreet
	(1)	(2)	(3)	(4)	(5)
Amtrak Op-Ed	0.719***	0.053	−0.059	0.084	−0.047
	(0.072)	(0.102)	(0.080)	(0.077)	(0.090)
Climate Op-Ed	0.090	0.561***	0.043	0.021	0.003
	(0.067)	(0.104)	(0.079)	(0.078)	(0.091)
Flat Tax Op-Ed	0.128*	0.186*	0.739***	0.100	−0.034
	(0.066)	(0.103)	(0.086)	(0.078)	(0.092)
Veterans Op-Ed	0.092	0.0002	−0.055	0.887***	−0.149*
	(0.068)	(0.102)	(0.081)	(0.077)	(0.089)
Wallstreet Op-Ed	0.043	0.091	0.133*	−0.146*	1.026***
	(0.065)	(0.102)	(0.079)	(0.078)	(0.091)
Constant (Control)	−0.178	−0.144	−0.132	−0.156	−0.136
	(0.046)	(0.070)	(0.056)	(0.054)	(0.062)
Wave	Wave 1	Wave 1	Wave 1	Wave 1	Wave 1
N	3,567	3,567	3,567	3,567	3,567
R^2	0.041	0.011	0.036	0.058	0.061

*p < .1; **p < .05; ***p < .01
Robust standard errors are in parentheses.

Table A.27: Newspapers Wave 2

	Amtrak	Climate	Flat Tax	Veterans	Wallstreet
	(1)	(2)	(3)	(4)	(5)
Amtrak Op-Ed	0.367***	−0.004	0.050	0.100	−0.087
	(0.074)	(0.115)	(0.088)	(0.087)	(0.103)
Climate Op-Ed	0.010	0.291**	0.165*	0.170**	−0.064
	(0.072)	(0.114)	(0.087)	(0.085)	(0.102)
Flat Tax Op-Ed	0.004	0.082	0.403***	0.113	−0.055
	(0.071)	(0.115)	(0.094)	(0.087)	(0.103)
Veterans Op-Ed	0.040	−0.042	−0.004	0.511***	−0.148
	(0.072)	(0.113)	(0.090)	(0.084)	(0.102)
Wallstreet Op-Ed	−0.075	−0.057	0.055	−0.065	0.451***
	(0.069)	(0.112)	(0.089)	(0.086)	(0.102)
Constant (Control)	−0.162	−0.120	−0.179	−0.251	−0.017
	(0.050)	(0.078)	(0.062)	(0.059)	(0.073)
Wave	Wave 2	Wave 2	Wave 2	Wave 2	Wave 2
N	2,989	2,989	2,989	2,989	2,989
R^2	0.015	0.004	0.009	0.018	0.015

*p < .1; **p < .05; ***p < .01
Robust standard errors are in parentheses.

Table A.28: Newspapers Wave 3

	Amtrak	Climate	Flat Tax	Veterans	Wallstreet
	(1)	(2)	(3)	(4)	(5)
Amtrak Op-Ed	0.299***	0.028	0.007	0.146	−0.066
	(0.082)	(0.128)	(0.095)	(0.097)	(0.109)
Climate Op-Ed	0.017	0.175	0.117	0.240**	−0.067
	(0.082)	(0.128)	(0.098)	(0.102)	(0.112)
Flat Tax Op-Ed	0.039	−0.007	0.322***	0.215**	−0.016
	(0.080)	(0.132)	(0.103)	(0.099)	(0.112)
Veterans Op-Ed	0.045	−0.084	−0.069	0.561***	−0.065
	(0.083)	(0.127)	(0.099)	(0.097)	(0.111)
Wallstreet Op-Ed	−0.026	−0.046	−0.043	−0.020	0.481***
	(0.077)	(0.124)	(0.099)	(0.096)	(0.110)
Constant (Control)	−0.137	−0.089	−0.175	−0.269	−0.066
	(0.056)	(0.086)	(0.066)	(0.069)	(0.076)
Wave	Wave 1	Wave 1	Wave 1	Wave 1	Wave 1
N	2,448	2,448	2,448	2,448	2,448
R^2	0.009	0.002	0.008	0.019	0.015

*p < .1; **p < .05; ***p < .01
Robust standard errors are in parentheses.

A.17 Study 20: Death Penalty

Peffley and Hurwitz (2007) present the results of a question wording experiment on support for capital punishment. The original study considered subgroup differences by respondent racial category, but I will present the unconditional results of both the original and replication experiments below.

A.17.1 Original Study

Sample: 1,142 White and Black Adult Americans.

Dates of data collection: October 19, 2000 – March 1, 2001.

Treatments:

- **Control**: Do you favor or oppose the death penalty for persons convicted of murder?
- **African Americans**: Some people say [FBI statistics show] that the death penalty is unfair because most of the people who are executed are African Americans. Do you favor or oppose the death penalty for persons convicted of murder?
- **Innocent**: Some people say [FBI statistics show] that the death penalty is unfair because too many innocent people are being executed. Do you favor or oppose the death penalty for persons convicted of murder?

Outcomes

- **Favor Death Penalty**: The response options were Strongly Oppose, Somewhat Oppose, Somewhat favor, Strongly Favor. The dependent variable was recoded 1 if the respondent chose Somewhat favor or Strongly favor. The original dataset was not available, but a tabulation by treatment condition and race was printed in the original article. Separate tabulations by the [Some people say]/[FBI statistics show] manipulations were not included, so I collapse over that margin in the analysis of the original experiment and in the replication, although the replication did include that manipulation.

A.17.2 Replication Study

Sample: 1,575 Mechanical Turk respondents.

Dates of data collection: March 2015.

Treatments: I used the identical treatments as in the original.

Outcomes: I used the identically worded outcome questions.

A.17.3 Results

In both the original and the replication, the Innocent treatment decreased support for the death penalty.

Table A.29: Death Penalty Original and Replication Results

	Favor Death Penalty	
	(1)	(2)
Innocent	-0.218^{**}	-0.402^{***}
	(0.095)	(0.098)
African Americans	0.069	-0.303^{***}
	(0.096)	(0.096)
Constant (Control)	2.629	2.734
	(0.077)	(0.089)
Sample	Original	MTurk
N	1,142	1,575
R^2	0.012	0.011

$^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$
Robust standard errors are in parentheses.

B Reanalysis and Replication of Lord, Ross, and Lepper (1979)

I embedded a replication of Lord, Ross, and Lepper's original 1979 design in the Capital Punishment study (Study 1) presented in Chapter 2. In this appendix, I present the original results and replication to show how remarkably well the original patterns of evidence are reproduced, nearly 40 years later and with a qualitatively different subject pool.

The central critique of the original study lies in the interpretation of results. In my view, the observed correlation between subjects' initial views and responses was inappropriately taken as evidence of polarization.

Other scholars have conducted replications of Lord, Ross, and Lepper as well (Pyszczynski, Greenberg and Holt 1985; Miller et al. 1993; Kuhn and Lao 1996; Munro and Ditto 1997; Corner, Whitmarsh and Xenias 2012). Those replications focused on a separate issue with the original study, the measurement strategy. Rather than asking subjects to report their own opinion change, subsequent efforts have taken the difference in measured levels of support before and after the experiment (e.g., Miller et al. 1993). No replication using this direct measure of change has found evidence of attitude polarization.

B.0.1 Hypotheses

Lord et al. (1979) was designed to evaluate two hypotheses using a within-subjects comparison of attitudes before and after encountering mixed evidence.

According to the authors, the biased assimilation hypothesis predicts that "individuals will dismiss and discount empirical evidence that contradicts their initial views but will derive support from evidence, of no greater probativeness, that seems consistent with their views" (p. 2099). The measure of biased assimilation is the observed correlation between

initial attitudes and evaluations of evidence. Strictly speaking, however, this operationalization of biased assimilation does not measure the extent to which subjects incorporate new and possibly discordant facts into their base of knowledge (as would be implied by the term "assimilation"). This distinction is important because, in principle, individuals can update their beliefs even if they rate the evidence (or the source of the evidence) as being of low quality.

The attitude polarization hypothesis predicts that mixed evidence makes attitudes and beliefs more extreme: "Our thesis is that belief polarization will increase, rather than decrease or remain unchanged, when mixed or inconclusive findings are assimilated by proponents of opposite viewpoints" (Lord et al. 1979, p. 2099). This prediction concerns treatment effect heterogeneity: mixed evidence should have a positive effect for proponents and a negative effect for opponents. While not stated explicitly in counterfactual language, initial beliefs are supposed to have a causal impact on subjects' responses to new information, that is, initial beliefs are hypothesized to moderate the effect of information on attitudes. This proposition is not adequately tested by this design, as subjects' initial beliefs are not randomly assigned.

B.0.2 Design of Lord et al. (1979)

1. An in-class survey was administered to 151 undergraduates that included three items on capital punishment: attitudes toward the death penalty, beliefs about its deterrent effect, and whether the relevant research supported their views.

2. A subset of 48 of the 151 was selected to participate in further research. Half of the 48 were in favor of capital punishment and the other half opposed; these subjects were selected because their answers to the three capital punishment survey items were internally consistent and showed a pattern of strong belief in the initial attitude.

3. Subjects were pseudo-randomly[1] assigned to one of four conditions using a 2x2 factorial design. The first factor was the order in which the fabricated evidence was presented: either pro- or anti-capital punishment came first. The second factor was the research method of the fabricated evidence: either time series or cross-sectional analysis was presented first.

[1]It appears that groups of four or more subjects were (cluster-) assigned to conditions by session, and that session conditions followed an alternating pattern. This procedure is pseudo-random insofar as subjects are not randomly assigned to sessions, nor are sessions randomly assigned to conditions.

4. Subjects were given a "results card" with a research finding and then asked to rate the changes in their beliefs about capital punishment and its deterrent effect on scales from -8 to 8.
5. Subjects were then given a "criticism card" in which the research method was described and methodological critiques were made. Subjects then were asked to evaluate how well the study had been done and how convincing they found the evidence on scales from -8 to 8.
6. Subjects wrote essays explaining their answers.
7. Subjects rated the changes in their beliefs about capital punishment and its deterrent effects once again.
8. Subjects then repeated steps 4-7, this time with evidence from the other point of view and using the other research design.

Some features of the design are worth emphasizing. First, the "running record" measurement strategy may have introduced systematic biases: by the end of the entire protocol, subjects had given their attitudes toward capital punishment five times, plus written an essay explaining their views. Repeated measurement itself may cause subjects to evaluate, re-evaluate, and possibly change their positions (Houston and Fazio 1989). Second, changes in attitudes toward capital punishment were obtained by asking subjects if and how their opinions changed in response to the two sets of evidence, rather than asking what their attitudes were post-evidence. The distinction between the self-reported measure of change and the "directly assessed" measure is a potential methodological weakness discussed at length in the original article and in subsequent replications (Lord et al. 1979, fn. 1; Miller et al. 1993, p. 562).

Finally, this study employed a within-subjects, rather than between-subjects, design for estimating the causal effect of mixed evidence on attitudes. Although the study did employ random assignment (the presentation order and research paradigm of the fabricated evidence), the treatment itself (mixed evidence) was not randomly assigned to some subjects but not others. Treatment effects are defined as the difference between two potential outcomes (Rubin 1974): subjects' attitudes in the absence of new information (control) and attitudes after encountering new evidence (treatment). This design implicitly assumes

that subjects' potential outcomes do not change over time and are unaffected by the measurement strategy – specifically, that untreated potential outcomes can be measured using a pre-test and that treated potential outcomes are observed during the post-test. Rather than make these assumptions, a between-subjects design compares treatment groups that are randomly assigned to reveal one potential outcome or the other.

B.0.3 Results of Lord et al. (1979)

In the Lord et al. study, the order in which the anti- and pro-deterrence evidence was presented had no discernible impact on the difference between evaluations of the two studies. Likewise, the evaluations of the studies were insensitive to the design of the fabricated evidence (either time-series or cross-sectional analysis) – subjects did not display a preference for one mode of inquiry over the other.

The results supported the biased assimilation hypothesis. Table B.1 shows subjects' mean evaluations of the studies' quality and persuasiveness. Proponents found the pro-deterrence study to be better conducted (difference = 3.1) and more convincing (difference = 3.2) than the anti-deterrence study. Opponents held the opposite opinion, finding the anti-deterrence study to be of higher quality (difference = -1.8) and more persuasive (difference = -2.2). Using t-tests, the authors found all four of these differences to be statistically significant at $p < 0.01$ or better. These results show a clear correlation between initial attitudes and evaluations of evidence.

Table B.2 presents the evidence the authors offer in support of attitude polarization: After seeing both sets of evidence, proponents reported moving an average of 1.5 scale points in the pro-capital-punishment direction and opponents reported having moved an average of 1.7 scale points in the anti-capital-punishment direction. A similar pattern holds for beliefs about the deterrent efficacy of capital punishment, which move in opposite directions for proponents (1.4) and opponents (-1.8). T-tests reveal that these differences are all significant at $p < 0.001$. These results show a clear correlation between initial positions and self-reported changes in beliefs and attitudes, which the original authors interpret as evidence of the attitude polarization hypothesis. Note, however, that the "Results only"

Table B.1: Biased Assimilation Results of Lord, Ross, and Lepper (1979)

Study	Proponents (N = 24)	Opponents (N = 24)
Mean ratings of how well the two studies had been conducted		
Prodeterrence	1.5	-2.1
Antideterrence	-1.6	-0.3
Difference	3.1	-1.8
Mean ratings of how convincing the two studies were as evidence on the deterrent efficacy of capital punishment		
Prodeterrence	1.4	-2.1
Antideterrence	-1.8	0.1
Difference	3.2	-2.2

panel of Table B.2 also includes evidence that does not support this hypothesis – proponents and opponents appeared to move in parallel after just reading the "results card" of each study.

Table B.2: Attitude Polarization Results of Lord, Ross, and Lepper (1979)

	Initial Attitudes	
Issue and Study	Proponents (N = 24)	Opponents (N = 24)
	Results only	
	Capital punishment	
Prodeterrence	1.3	0.4
Antideterrence	-0.7	-0.9
Combined	0.6	-0.5
	Deterrent efficacy	
Prodeterrence	1.9	0.7
Antideterrence	-0.9	-1.6
Combined	1.0	-0.9
	Details, data, critiques, rebuttals	
	Capital punishment	
Prodeterrence	0.8	-0.9
Antideterrence	0.7	-0.8
Combined	1.5	-1.7
	Deterrent efficacy	
Prodeterrence	0.7	-1.0
Antideterrence	0.7	-0.8
Combined	1.4	-1.8

B.0.4 Previous Replications

To date, there have been at least seven published replications of the Lord et al. (1979) design. A summary of these replications and their findings with respect to biased assimilation and attitude polarization is presented in Table B.3. All replications find evidence of biased assimilation, i.e., a positive correlation between initial attitudes and subsequent evaluations of evidence. Five of the replications gathered a self-reported measure of change; all five found the same pattern of attitude polarization as the original article (although one, Miller et al. 1993, found no significant polarization when examining attitudes on affirmative action). Five of the replications included the so-called "direct measure" of attitude change by asking participants for their post-evidence attitudes and comparing them to the pre-evidence attitude. None of these replications found evidence of attitude polarization using the direct measure.

Table B.3: Studies of Biased Assimilation and Attitude Polarization

					Attitude Polarization	
Authors	Year	Issue	Max N	Biased Assimilation	Self-reported	Directly Assessed
Lord, Ross, and Lepper	1979	Capital Punishment	48	Yes	Yes	–
Pyszczynski, Greenberg, and Holt	1985	Test of social sensitivity	40	Yes	–	No
Houston and Fazio	1989	Capital Punishment	102	Yes	–	–
Plous	1991	Technological Breakdowns	215	Yes	Yes	–
Miller, McHoskey, Bane, and Dowd	1993	Capital Punishment,	337	Yes	Yes	No
		Affirmative Action	109	Yes	No	–
Kuhn and Lao	1996	Capital Punishment	228	Yes	Yes	No
Munro and Ditto	1997	Homosexuality	174	Yes	Yes	No
Corner, Whitmarsh, and Xenias	2012	Climate Change	173	Yes	Yes	No
The present study	2016	Capital Punishment	670	Yes	Yes	No

In sum, the empirical evidence on biased assimilation and attitude polarization leads to a split decision. As a descriptive fact, biased assimilation seems to take place. There is clear evidence of a correlation between prior beliefs and evaluations of evidence. Attitude polarization, however, receives limited support in these replications. Leaving aside the difficulty of drawing causal inferences from the pre-post design, the conclusions are highly dependent on the measurement strategies. Which measurement is appropriate is a matter of dispute: Kuhn and Lao (1996) prefer the direct measure, as it may be free of a demand effect in which subjects feel compelled to report some change in beliefs even when there

is none. Ross (2012) prefers the self-reported measure, as it may refer to an individual's subjective assessment of how the information affected his or her beliefs.

I sympathize with those who criticize the self-reported change measure. Other scholars have suggested that individuals may not possess the necessary insight into their own cognitive mechanisms to be reliable guides to changes in their attitudes (Goethals and Reckman 1973; Nisbett and Ross 1980). That being said, the "direct measure" suffers from other biases, most notably regression to the mean – those with extreme attitudes may appear to moderate in the face of new evidence when no such change has occurred. Fortunately, experimental design employed in my replication obviates the need to take a strong stand on which is the better measure of "true" opinion change: the choice of measurement does *not* lead to substantively different conclusions when the experiment is analyzed according to the randomly assigned treatment groups.

B.0.5 Replication Study

I first present the results of a within-subjects analysis in the *Pro Con* condition of Study 1 only, corresponding to the original Lord et al. design. Table B.4 is presented in the same format as Table B.1 for ease of comparison. As in the original study, proponents ranked the Pro study as better conducted (difference = 3.24) and more convincing (difference = 4.10) than the Con study. Among opponents, the opposite pattern holds. The difference in mean quality rankings was -2.81 and the difference in mean persuasiveness ratings was -2.22. The pattern of biased assimilation observed among these online participants in 2014 is substantively identical to the pattern observed among Stanford undergraduates in the late 1970s. The point estimates only differ by an average of about 0.5 scale points, which is well within sampling variability.[2]

Table B.5 presents the average self-reported changes in attitudes and beliefs for both proponents and opponents of capital punishment. After reading only the results page, proponents reported average attitude change of 3.08 scale points in the pro direction, while

[2]I do not present standard errors in these tables for continuity of presentation, but they range between 0.40 and 0.55.

Table B.4: Replication: Biased Assimilation

Study	Proponents (N = 50)	Opponents (N = 68)
Mean ratings of how well the two studies had been conducted		
Prodeterrence	4.02	-0.56
Antideterrence	0.78	2.25
Difference	3.24	-2.81
Mean ratings of how convincing the two studies were as evidence on the deterrent efficacy of capital punishment		
Prodeterrence	4.00	-1.54
Antideterrence	-0.10	0.68
Difference	4.10	-2.22

opponents reported moving 3.09 scale points in the con direction. The same pattern holds for beliefs about deterrent efficacy (proponents 2.76; opponents -3.17). The second panel of Table B.5 shows the mean attitudes and beliefs after reading the study details and criticism. Again, proponents reported higher support for capital punishment after encountering evidence, and vice versa for opponents. The signs on the "combined" rows all match those in the original study (see Table B.2), though the magnitudes of the changes are two to three times as large. This difference may have as much to do with the measurement technology (my subjects reported their answers using online "sliders" whereas the Stanford subjects used pen and paper) as with differences across subjects and contexts.

Finally, in keeping with earlier replications (Kuhn and Lao 1996; Miller et al. 1993), I also measured attitude change using the so-called direct measure. In the pre-survey, subjects responded to 7-point attitude and belief questions (T1). At the very end of the main survey, after subjects had been exposed to both sets of evidence, they responded to the same attitude and belief items again (T2). The difference between them (T2 − T1) forms the direct measure of attitude change. Table B.6 displays mean responses for both periods and their difference for proponents and opponents. Mixed evidence does not appear to produce polarized attitudes either for proponents (difference = 0.04, s.e. = 0.13) or for opponents (difference = 0.06, s.e. = 0.07). Contrary to earlier replications, the direct

Table B.5: Replication: Attitude Polarization (self-reported measure)

	Initial Attitudes	
Issue and Study	Proponents (N = 50)	Opponents (N = 68)
	Results only	
	Capital punishment	
Prodeterrence	2.86	-0.44
Antideterrence	0.22	-2.65
Combined	3.08	-3.09
	Deterrent efficacy	
Prodeterrence	3.04	0.19
Antideterrence	-0.28	-3.37
Combined	2.76	-3.17
	Details, data, critiques, rebuttals	
	Capital punishment	
Prodeterrence	2.24	-1.13
Antideterrence	0.34	-1.93
Combined	2.58	-3.06
	Deterrent efficacy	
Prodeterrence	2.64	-1.26
Antideterrence	0.12	-2.19
Combined	2.76	-3.46

measure actually provides suggestive evidence of attitude *moderation* on the beliefs about deterrent efficacy. Proponents believe in deterrence 0.22 scale points less, and opponents 0.62 scale points more. Note, of course, that this apparent difference could easily be due to regression to the mean and be unrelated to exposure to mixed evidence.

Table B.6: Replication: Attitude Polarization (direct measure)

	Proponents (N = 50)			Opponents (N = 68)		
	T1	T2	T2−T1	T1	T2	T2−T1
Attitudes toward CP	5.90	5.94	0.04	1.81	1.87	0.06
	(0.12)	(0.15)	(0.13)	(0.10)	(0.12)	(0.07)
Beliefs about Efficacy	5.44	5.22	-0.22	1.94	2.56	0.62
	(0.11)	(0.17)	(0.13)	(0.12)	(0.13)	(0.13)

Subjects in *Pro Con* condition only.
Standard errors in parentheses.

The within-subjects analysis of the *Pro Con* condition again leads to an impasse. The attitude polarization hypothesis is strongly supported by the self-reported change measure and is roundly rejected by the direct measure.

The replication was successful insofar as it was able to reproduce the results of the original authors using the self-reported measure and also the results of subsequent scholars who preferred the direct measure. However, the original study and the replications constitute weak evidence for (or against) attitude polarization. This weakness stems from inadequate experimental design, as is illustrated by the contrast between the results of Study 1 and this appendix.

C Reanalysis of two Vaccine Correction Experiments

In the main text, I referred to a pair of survey experiments that showed evidence of a backlash effect of pro-vaccine messages among people who are predisposed to mistrust vaccines. The finding that "anti-vaxxers" report lower intention-to-vaccinate when treated with pro-vaccine information (compared with those in a control condition) was documented in Nyhan et al. (2014) and replicated in Nyhan and Reifler (2015). To my knowledge, this represents the only example of a backlash effect measured in a randomized survey experiment.

These backlash effects notwithstanding, the treatments also had correctly-signed effects on other vaccine-related attitudes among the group that is pre-disposed to "resist" pro-vaccine information. In Nyhan et al. (2014), the correction decreases the belief that vaccines cause autism among this group, and in Nyhan and Reifler (2015) it decreased the belief that vaccines are unsafe or cause the flu.

C.1 Reanalysis of Nyhan et al. 2014

Nyhan et al. (2014) tested the effects of four different interventions (relative to control) separately for subgroups of the experimental sample defined by terciles of a vaccine attitudes scale measured pre-treatment. In Tables C.1 and C.2 I present my reanalysis of the experiment focusing on only the "correction" intervention. Whereas the original analysis estimated treatment effects using an ordered logit specification, I use OLS with a control for the vaccine attitudes scale to improve precision. Table C.1 shows the "backlash" effect: the correction reduces intention-to-vaccinate among those in the bottom tercile of the vaccine attitudes scale but does not appear to alter the intention-to-vaccinate among others.

Table C.2, however, we see that the correction has the intended negative effect on the belief that vaccines cause autism.

Table C.1: Nyhan and Reifler (2014) Reanalysis: Intention to Vaccinate

	Intention to vaccinate		
	Bottom Tercile	Middle Tercile	Top Tercile
Autism Correction	-0.645^{***}	-0.093	0.068
	(0.179)	(0.122)	(0.070)
Pre-treatment Vaccine Attitude Scale	1.846^{***}	0.041	0.028
	(0.185)	(0.310)	(0.079)
Constant	-0.298	5.708	5.780
	(0.600)	(1.180)	(0.348)
N	270	202	235
R^2	0.388	0.004	0.008

$^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$
Robust standard errors are in parentheses.

Table C.2: Nyhan and Reifler (2014) Reanalysis: Autism

	Vaccines Cause Autism		
	Bottom Tercile	Middle Tercile	Top Tercile
Autism Correction	-0.237^{*}	-0.260	-0.288^{**}
	(0.122)	(0.192)	(0.138)
Pre-treatment Vaccine Attitude Scale	-0.599^{***}	-1.235^{**}	-1.038^{***}
	(0.129)	(0.616)	(0.233)
Constant	5.078	7.287	6.598
	(0.403)	(2.354)	(1.038)
N	267	200	231
R^2	0.123	0.066	0.140

$^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$
Robust standard errors are in parentheses.

C.2 Reanalysis of Nyhan and Reifler 2015

In Nyhan and Reifler (2015), a correction of vaccine misperceptions decreased intention-to-vaccinate among a subgroup who expressed "high" concern about vaccines. The original article uses an ordered probit specification, but I will analyze the data with OLS. An estimate of this unintended "backfire" effect is presented in the second column of Table C.3.

The estimates presented in the fourth and sixth columns of the table, however, show that this group updated in the direction of the correction on beliefs that vaccines are unsafe and that vaccines cause the flu.

Table C.3: Nyhan and Reifler (2015) Reanalysis

	Vac. Intention		Vac. Unsafe		Vac. Gives Flu	
Correction	0.176	-0.841^{*}	-0.215^{**}	-0.093	-0.330^{***}	-0.468^{*}
	(0.297)	(0.441)	(0.086)	(0.208)	(0.115)	(0.254)
Pre-treatment Concern	-0.172	-0.049	0.364^{***}	0.398^{*}	0.520^{***}	0.119
	(0.210)	(0.452)	(0.056)	(0.219)	(0.071)	(0.249)
Constant	4.070	3.440	0.847	0.709	1.007	2.313
	(0.522)	(1.973)	(0.111)	(0.971)	(0.174)	(1.199)
Concern	low	high	low	high	low	high
N	511	147	512	146	511	147
R^2	0.005	0.045	0.171	0.056	0.162	0.057

$^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$
Robust standard errors are in parentheses.

C.2.1 Discussion of Vaccine Experiments

The negative effect of corrections on anti-vaccine respondents' self-reported intention to vaccinate is robust: it holds up under ordered logit/probit and OLS specifications and replicates across two samples. The p-values on these negative effects are very small and would survive most multiple comparisons corrections. This backlash effect is not in dispute.

A puzzle, however, arises from the fact that in addition to this backlash on intention-to-vaccinate, anti-vaccine respondents also update their attitudes in the direction of evidence on other dimensions of vaccine attitudes. It could be that survey questions that ask respondents to predict their own future behavior are especially sensitive to motivated reasoning processes. It could be that subjects' attitudes towards vaccines are at least two-dimensional, and that decreasing the misperception that vaccines cause autism does not, in turn, increase the probability that subjects will vaccinate their children.