CrossMark

ORIGINAL PAPER

The Exaggerated Life of Death Panels? The Limited but Real Influence of Elite Rhetoric in the 2009–2010 Health Care Debate

Daniel J. Hopkins[1]

Abstract Experiments demonstrate that elites can influence public opinion through framing. Yet outside laboratories or surveys, real-world constraints are likely to limit elites' ability to reshape public opinion. Additionally, it is difficult to distinguish framing from related processes empirically. This paper uses the 2009–2010 health care debate, coupled with automated content analyses of elite- and mass-level language, to study real-world framing effects. Multiple empirical tests uncover limited but real evidence of elite influence. The language Americans use to explain their opinions proves generally stable, although there is also evidence that the public adopts the language of both parties' elites symmetrically. Elite rhetoric does not appear to have strong effects on Americans' overall evaluations of health care reform, but it can influence the reasons they provide for their evaluations. Methodologically, the automated analysis of elite rhetoric and open-ended questions shows promise in distinguishing framing from other communication effects and illuminating elite-mass interactions.

✉ Daniel J. Hopkins
danhop@sas.upenn.edu

1   Department of Political Science and Leonard Davis Institute of Health Economics, University of Pennsylvania, Philadelphia, USA

🌀 Springer

Introduction

For politicians, the ability to frame complex issues in advantageous ways is a central
tool. In 2000, a state legislator named Barack Obama premised his candidacy for
Congress on his capacity to frame issues, arguing that he was able to "best articulate
and frame the issues that are most important to voters" (Frontline 2009). Political
scientists devote substantial attention to framing as well. From 2010 to 2015, there
were 46 separate articles on the topic in three leading political science journals.[1] As
Druckman et al. (2012) summarize, a "generation of research shows that elites can
use frames...to affect public opinion."

Despite the wealth of recent studies on framing, there are outstanding questions
about how and to what extent framing durably shapes public opinion. For one thing,
research on framing has been primarily experimental. While there is some
observational evidence from shifts in salient frames (e.g. De Vreese 2003; Kellstedt
2003; Jacobs et al. 2003; Entman 2004; Smith 2007; Baumgartner et al. 2008;
Edwards 2009), scholars know less about how framing operates in real-world
conditions (Chong and Druckman 2011). In fact, there are a number of reasons to
suspect that framing's real-world influence might be quite limited (Barabas and Jerit
2010), from public inattention (Prior 2007; Arceneaux et al. 2012) and selective
information-seeking (Druckman et al. 2012) to conflicting media incentives
(Entman 2004; Hayes 2008) and investments in existing frames (Baumgartner
et al. 2009, Chap. 9). It is also plausible that elites craft their language using polling
and focus groups (Jacobs and Shapiro 2000; Jacobs and Burns 2004), a practice
which might make salient frames more of a reflection of public opinion than an
independent influence on it.

Studying framing in real-world contexts also foregrounds a significant measure-
ment problem. Framing is thought to influence public opinion by changing the
relative accessibility and applicability of different considerations in citizens' minds.
Yet when politicians emphasize a certain aspect of an issue, they also provide
different information and employ different persuasive strategies (Scheufele and
Iyengar 2012; Leeper and Slothuus 2015; see also Lenz 2013). Framing effects take
place when certain aspects of a broader issue become cognitively accessible, but not
when citizens learn new facts or are persuaded to evaluate existing facts differently.
In most real-world studies, it is accordingly difficult to distinguish framing from
learning or other types of persuasion.

To address these questions, we report a case study of attitudes toward the
Affordable Care Act (ACA), a major health policy reform signed by President
Obama in 2010. The initial debates surrounding the ACA in 2009–2010 provide a

---

[1] The three journals in question are the *American Journal of Political Science*, *American Political
Science Review*, and *Journal of Politics*.

critical opportunity to study real-world framing effects in a high-salience setting. Health care reform is a complex, multifaceted issue, giving supporters and opponents alike framing opportunities (Jacobs and Shapiro 2000). And frame they did. In September 2009, President Obama gave a prime-time Congressional address devoted entirely to health care. Between January 2009 and December 2010, U.S. Senators sent out 1,488 press releases related to the issue.

To many political commentators, the strength of public opposition to the ACA in the period after the law's passage was evidence of the effectiveness of some of these efforts. For instance, the *New York Times* stated the conventional wisdom when it wrote that "the Obama administration and Democrats ...largely lost the health care message war in the raucous legislative process" (Steinhauer and Pear 2011). Scholars have joined commentators in contending that politicians' rhetorical choices influenced public views of health care reform, with former Alaska Governor Sarah Palin's use of the phrase "death panels" cited as effective anti-ACA rhetoric (Scherer 2010; Shapiro and Jacobs 2010; see also Nyhan 2010; Kriner and Reeves 2014). Nine years after campaigning on his ability to frame issues, Barack Obama was perceived to have lost the battle to frame his signature domestic policy initiative. But the empirical evidence underpinning this presumption is limited.

There are also practical reasons to study real-world communication effects during the ACA debate. The Kaiser Family Foundation's (KFF) monthly surveys provide researchers with more than 30,000 fully observed respondents who were asked about health care reform between February 2009 and January 2012. The Pew Research Center and Kaiser together asked open-ended questions about health care reform in seven surveys; to our knowledge, there is no comparable set of open-ended questions spanning another major American policy debate.[2] By coupling these data with tools for automated content analysis such as Latent Dirichlet Allocation (Blei et al. 2003), we are able to observe the specific frames that both elected officials and American citizens use in describing their views of health care reform—and we can do so before and after the issue reached peak salience. This focus on word choice allows us to measure elite frames and public opinion on the same scale, and to identify the extent to which elite frames induce shifts in public arguments. Critically, it also provides a way of identifying whether elite frames do in fact change the aspects of a political issue which are cognitively accessible for citizens. Automated content analysis thus helps close the gap between theories and tests of framing effects.

The results suggest the impacts and limits of real-world political communications during a single, salient political debate. In contrast to the variability of elite frames over this period, public opinion moves only gradually, and the arguments Americans use prove stable. The punctuated nature of elite frames on health care makes them a poor fit to explain attitudes which change only gradually. There is evidence that the public incorporates the language used by political elites during the most intense phase of the debate, and that it does so roughly symmetrically (Berinsky 2007). However, the broad contours of the public's arguments for and

---

[2] The KFF data was obtained from the Roper Center's iPoll database. The Pew Research Center's surveys were downloaded from http://www.pewresearch.org/data/download-datasets/. All other code and data to replicate the research reported herein is available at https://dataverse.harvard.edu/dataverse/DJHopkins.

against health care reform were visible as early as July 2009—and as various measures outlined below indicate, that was prior to health care reform's spike in public salience, with more than twice as many cable television stories in the subsequent months. Elite messages do induce detectable shifts in public language, especially in the period after November 2009. But at least on the highly salient issue of health care attitudes, we do not observe more widespread shifts in opinion which are likely to reflect framing effects. *In short, elite rhetoric appears more likely to change the rationale underpinning evaluations of the ACA than to change the evaluations themselves.*

Certainly, these results are compatible with the existence of framing effects that unfold over longer periods of time (Smith 2007) or that are more pronounced for less salient issues, points developed in the Conclusion. But these results do reinforce the limits on the effects of elite rhetoric in a given political debate, and they underscore the complex, interactive nature of real-world communication effects (see also Jacobs and Shapiro 2000). This manuscript also aims to contribute methodologically, as it shows that the analysis of respondents' language and the application of automated content analysis to public opinion research allow for richer and more subtle measures of framing.

Prior Research and Hypotheses on Real-World Framing Effects

Scholars have defined framing effects in a variety of ways (e.g. Nelson 2011). Our goal is not to resolve such debates, but to identify a workable definition of framing before discussing its study in real-world settings. This section then grounds the paper's core hypothesis: for various reasons, real-world framing effects are likely to be quite limited. Indeed, even when political communications and public attitudes do move in tandem, those changes could be a product of persuasion, motivated reasoning, or other processes distinct from framing.

We follow Chong and Druckman (2010) by focusing on issue framing, which occurs "when a communication changes people's attitudes toward an object by changing the relative weights they give to competing considerations about an object" (p. 665). So defined, framing effects operate by influencing the cognitive accessibility of considerations relevant to an issue. Issue frames "embody a complex semantic structure" (Nelson 2011, p. 192) which facilitates certain mental associations at the expense of many others.

We formalize this conception by using the expectancy-value model of attitude formation (Nelson et al. 1997; Leeper and Slothuus 2015), in which attitude A toward a given object is the sum of considerations c_i and corresponding weights w_i. Overall attitude A is $\sum_{\forall i} c_i w_i$, or the sum of the weighted considerations. Framing effects are changes in w_i, not in c_i. Frames do not change people's considerations, but instead change the weights given to existing considerations. As a result, framing effects are expected to be stronger among those with more knowledge of or experience with a given subject, as they will have a wider variety of considerations to connect to that issue (Mutz 1994; Chong and Druckman 2007a; but see Lecheler

et al. 2009). Framing effects are distinctive from learning or persuasion, as the underlying mechanism of framing is cognitive accessibility (Huber and Paris 2012). In observational settings, however, we often lack the tools to differentiate framing from these related processes (Scheufele and Iyengar 2012; Leeper and Slothuus 2015).

The study of framing has developed rapidly in recent years, driven by experimental survey research (e.g. Iyengar and Kinder 1987; Chong and Druckman 2007b). This research generates expectations about real-world framing, including the hypothesis that framing effects will be especially pronounced when the frames come from trusted elites within one's party (Druckman 2003; Slothuus and de Vreese 2010; but see Bullock 2011). Attentive to concerns about external validity, scholars of framing have increasingly adopted research designs that more closely approximate real-world settings, whether by allowing for competition among frames (Sniderman and Theriault 2004; Chong and Druckman 2007b) or choices about exposure to frames (Arceneaux et al. 2012; Druckman et al. 2012).

The Limits of Real-World Framing Effects

Still, challenges remain in generalizing these experimental findings to real-world settings. In the typical framing experiment, the researcher chooses the frames with few limitations. In real-world settings, however, politicians and political elites are likely to face significant constraints in choosing frames. One stems from intra-party coordination problems, as lone officials are unlikely to advance frames that are at odds with their co-partisans or contemporary discourse generally (Baumgartner et al. 2009). Even if parties successfully coordinate, the news media might prove uninterested in transmitting a frame to the public (e.g. Entman 2004; Hayes 2008).

An additional limitation stems from the use of polling and focus groups in shaping contemporary elite rhetoric (Jacobs and Burns 2004). To the extent that political elites craft messages based on public opinion research, the causal effect in practice might be the opposite of that identified in experimental studies, with elites adopting the language and frames already used by the public. These various pathways suggest the importance of measuring not simply baseline opinion but also baseline word choice in public discussions of a given issue.

Still other limitations stem from facets of political psychology. For instance, citizens with partisan loyalties are likely to discount frames offered by the other party (Taber and Lodge 2006; Slothuus and de Vreese 2010). A related limitation stems from public inattention, as few prospective voters closely follow political discourse. The effects of political rhetoric are also likely to decay rapidly after exposure (Hill et al. 2013), further limiting framing effects. Given the limitations identified here, it is not surprising that studies of real-world framing effects report mixed results (e.g. Kellstedt 2003; Smith 2007; but see Jacobs et al. 2003; Edwards 2009). In light of these theoretical considerations and prior results, we expect that real-world framing effects are likely to be more limited than those observed in most survey experiments.

Absent a way to measure the relative accessibility of different considerations on a given political issue, prior research on real-world framing has been unable to

distinguish framing from other, related processes. Even when political communications do reshape citizens' ultimate evaluations, such effects are not necessarily evidence of framing (see especially Scheufele and Iyengar 2012; Leeper and Slothuus 2015). Framing is distinctive from other communication effects in that it involves changes in the accessibility and applicability of pre-existing cognitive considerations. If attitudes are indeed shifting through framing, we may well be able to detect a corresponding change in the weights which citizens place on different aspects of an issue.[3] A framing effect requires that both the weights placed on different considerations and the ultimate evaluation change.

Alternately, it is quite plausible that citizens will adopt their overall issue positions based on their partisan loyalties or other heuristics, and then turn to salient political rhetoric to justify those evaluations. Such claims are consistent with evidence from political psychology about the extent to which voters engage in motivated reasoning (Lenz 2013; Lodge and Taber 2013). As Lodge and Taber (2013) note, "conscious deliberation and rumination is from this perspective the *rationalization* of multiple unconscious processes that recruit reasons to justify and explain beliefs, attitudes, and actions" (p. 22). On a salient, partisan issue such as the ACA, citizens may first adopt an overall evaluation of the law based on partisan cues, affect, or other heuristics. Only subsequently will they adopt justifications for those views, often using the same reasons given by prominent party leaders. If so, we should expect not framing but what we might call "parroting" or motivated reasoning—the use of co-partisans' rhetoric to justify the party's position. Communications may change the relative weights given to different considerations without changing their overall evaluations.

Measuring Real-World Frames via LDA

Measurement problems have limited our capacity to study real-world framing. At the elite level, it is resource-intensive to measure the frames that elites employ, let alone observe those frames' changes over time or their causes. Outside of experimental settings, surveys do not typically include the questions needed to observe the subtle changes in mass cognition that framing posits, such as the increased accessibility of certain considerations. Also, prior scholarship has not typically measured frames and public opinion on the same scale, making it difficult to test the alternative explanation that politicians adopt frames already in use among citizens.

This section contends that analyses of word choice at both the elite and mass levels offer substantial leverage to address these limitations—and thus to understand framing and mass-elite dynamics. It first outlines why a variant of cluster analysis—LDA (Blei et al. 2003)—is a promising technique for measuring elite frames in real-

---

[3] To be sure, those changes in accessibility may be fleeting—and elite framing could influence public opinion even without influencing public word choice, especially if its effects do not rely on memory-based processing (see also Druckman and Leeper 2012). Put differently, in an online processing model of cognition, frames have the potential to influence attitudes even if the individual does not remember or employ the words used in the frame.

world settings. Departing from prior applications of LDA in political science (e.g. Lauderdale and Clark 2014), this section proposes using LDA to analyze even one-sentence documents such as open-ended survey responses. The claim is not that LDA will always return recognizable frames, but that when applied to similarly structured documents on a single issue, it has the potential to do so. By analyzing open-ended survey responses, we can also identify a closely related communication effect: changes in the rationale underlying a given attitude.

As the theoretical discussion makes clear, issue frames are rhetorical structures which call attention to a subset of the considerations relevant to an issue. Put differently, frames can be thought of as subtopics of a larger issue. But at a time of fragmented and polarized sources of political information, how can we measure the prominent frames in a real-world political debate?

Frames are closely associated with a speaker's choices of words. Given that, within a given topic, we might think of frames as being probability distributions over the relevant vocabulary. The health care frame emphasizing an expansion of governmental authority is more likely to use words including "government," "takeover," and perhaps "death" or "panel." In such a frame, terms like "affordable" or "pre-existing condition" would have far lower probabilities of occurring.

In recent years, computer scientists have developed models that closely match this conception of frames as being probability distributions over a vocabulary within a given topic. The most prominent of such models is LDA, a multilevel Bayesian model (Blei et al. 2003). LDA represents each document in a collection or corpus of texts as a mixture from a pre-defined number of clusters or topics. This means that a single document can draw from different clusters—or frames, in the applications below. Each cluster is in turn represented by a distinctive probability distribution over the corpus's vocabulary. By using an unsupervised technique such as LDA, researchers can be open to clusters of language they may not have anticipated ex ante (Grimmer and King 2011).

To understand LDA, it is valuable to state the model using the notation of Blei et al. (2003). Let K be a pre-defined number of clusters in a set of documents, with $\mathbf{w}$ a vector representing a document and V representing the number of unique words in the vocabulary. LDA models a collection of documents as emerging through the following process. First, the length of the document N is chosen from a Poisson distribution with prior parameter ξ ($N \sim Poisson(\xi)$). Then, the distribution of topics in that document θ is chosen from a Dirichlet distribution with prior parameter α ($\theta \sim Dirichlet(\alpha)$). From there, for each of the N words w_n, we first draw a word-specific topic z_n from a Multinomial distribution with parameter vector θ ($z_n \sim Multinomial(\theta)$), and then choose a word w_n from $p(w_n|z_n, \beta)$ ($w_n \sim p(w_n|z_n, \beta)$), where β is a $V \times K$ matrix indicating the probability that each of the K topics associates with each of the V words.[4]

Of the estimated parameters, a few are of particular interest. One is the $\hat{\boldsymbol{\beta}}$ matrix which allows us to identify the words that are associated with each topic. The second is the document-specific $\hat{\theta}$ vector, which provides the share of each

---

[4] The fact that the Dirichlet distribution is conjugate to the Multinomial enables researchers to fit the model using either a Gibbs sampler or variational inference (Blei et al. 2003).

document which falls into each topic. As with similar models in political science (Quinn et al. 2010; Roberts et al. 2014), this approach can be highly valuable in reducing the dimension of a textual data set and partitioning a set of documents into meaningful subtopics (Grimmer and Stewart 2013).[5]

Open-Ended Survey Responses

The impacts of real-world frames are almost always measured using survey questions with fixed responses. Closed-ended survey questions allow for easy comparisons across respondents. Yet they have limitations as well. Open-ended survey questions are known to provide a window into citizens' cognitive processes. Summarizing this trade-off, Payne (1951) writes that an open-ended question's "virtues and its faults all stem from this open feature. Its results are as full of variety as a country store, and just as hard to divide into departments" (p. 54). Open-ended questions elicit additional information, but information which is more difficult to summarize or compare across respondents. Automated techniques like LDA have the potential to reduce the disadvantages associated with open-ended questions, as they enable us to cluster responses based on word usage. We can identify that a respondent complaining about "too much control" and one talking about a "government takeover" are voicing related concerns. Yet to date, within political science, LDA and similar techniques have been applied primarily to lengthier documents of at least a few hundred words (e.g. Quinn et al. 2010; Grimmer 2010; Lauderdale and Clark 2014). Empirically, it is unclear whether brief open-ended responses of several words will provide sufficient density for such techniques to return meaningful clusters (see also Eisenstein et al. 2010; Roberts et al. 2014).

Even without clustering, the ability to analyze public word choice provides leverage in understanding framing. Words are the way that we identify elite frames, but to date, they have not been part of scholars' strategy for measuring framing effects among the public. By comparing elites' and citizens' word use, we will be able to measure whether the public adopts elite-level word choices as the health care debate unfolds. This approach allows for a more subtle test of elite influence than is commonly found in real-world studies, as we can observe whether citizens adopt elites' messages in explaining their views. It also has the critical advantage that it allows us to rule out endogeneity in word choice by identifying whether elites are adopting the frames already prominent among the public.

---

[5] LDA allows words to be characteristic of multiple topics. Yet in theory, LDA has a potential limitation in this application. Political rhetoric commonly involves a combination of frames. Yet LDA's use of the Dirichlet distribution to model topic generation means that the topics are assumed to be independent of one another. We thus confirm our analyses of elite rhetoric with the Correlated Topic Model (Blei and Lafferty 2006), which replaces the Dirichlet distribution giving rise to the topic probabilities with a logistic normal distribution.

Elite Framing: Data and Results

In the course of a campaign or debate, elite frames shift as circumstances change. Claims that health care reform was being passed through corrupt, back-room deals became meaningful only as the legislative process unfolded. To observe shifting frames, we need sources of political rhetoric that are available at many points in time. Since this paper's primary goal is to assess politicians' framing capacities, we also seek at least some rhetoric that is not yet filtered by the media. Speeches on the House or Senate floor might prove useful, but they are available mostly when a bill is formally being debated, truncating the over-time variation. For those reasons, press releases are an unparalleled source of information, as Grimmer (2010) shows. From January 2009 to July 2010, U.S. Senators alone sent out 1488 press releases using a health care-related term.[6] Often written in a form that mimics newspaper articles, press releases enable politicians to frame the issues in their preferred way. At the same time, we also analyze 218 television appearances to ensure that our results hold in content more likely to reach the public.

Figure 1 depicts the distribution of press releases by month, and shows that they peak at times of key legislative events. The figure illustrates the arc of the public debate, with the number of press releases growing in the summer of 2009, spiking in the fall and winter of 2009, and then spiking again with the bill's passage in March 2010. Other measures of salience depict a similar trajectory: while there was growing public and elite attention to health care reform as early as the summer of 2009, there was far more attention in the fall and winter of 2009/2010.[7] The figure's gray lines show select surveys, and reinforce that the bulk of the public debate took place after our baseline surveys in February and July 2009. Still, even in early 2009, the issue was not wholly unfamiliar.

To prepare the press releases and television appearances for model-fitting, they were first pre-processed using common techniques (e.g. Hopkins and King 2010). All words were reduced to common stems such as "senat" for "Senate," "Senator," etc. (Porter 1980), and those word stems that appeared in fewer than 1% of the documents were removed.[8] In addition, we remove proper nouns which are likely to be specific to certain documents, such as Senators' names, states' names, and party designations (such as "R-WY").[9] In all, this pre-processing leaves us with 2309 word stems for the press releases and 2057 for the television appearances.

---

[6] The search terms were "healthcare," "health care," "obamacare," "health reform," and "health insurance reform."

[7] Lexis–Nexis keyword searches of CNN and Fox News transcripts indicate that attention to health care reform rose to a peak of 603 stories in August 2009, remained high through the fall and winter, and reached their maximum of 667 stories in March 2010. According to the Gallup Most Important Problem series (Baumgartner and Jones 2013), the share of the public citing health care as the nation's most important problem rose from 7% in the second quarter of 2009 to 16% in the third quarter, and remained above 13% in the first quarter of 2010.

[8] Like the significant majority of automated analyses in political science to date, our analyses focus only on single words or unigrams, although incorporating common bigrams (such as "public option") would be a useful extension.

[9] For the press releases, we remove 170 such text strings. Given the differing format of the press appearances, we need remove only a small set of 11 names, typically those of the shows' hosts.
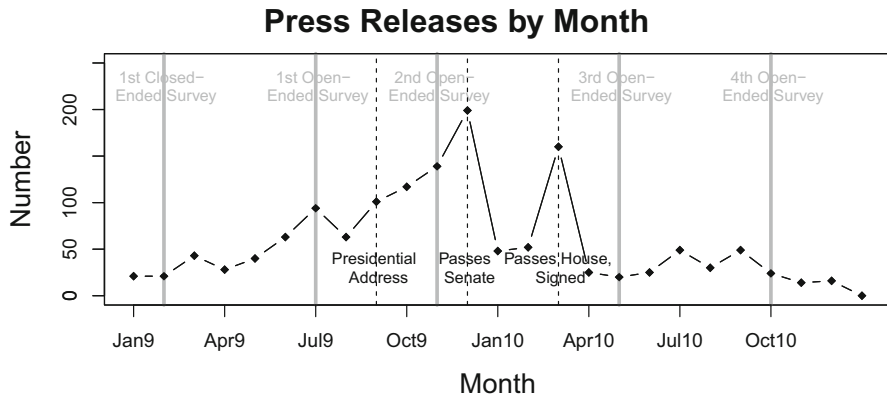
Press Releases by Month

Fig. 1 This figure depicts the number of press releases by month. It also identifies key events in the legislative timeline (*black lines*) as well as the dates of select surveys (*gray lines*)

To assess framing effects, we must first identify the elite frames employed in a given policy area. This section does so by applying LDA to 1,488 press releases from U.S. Senators between January 2009 and December 2010.[10] LDA returns a θ vector of length $K - 1$ indicating the topic probabilities as well as a β matrix. The β matrix has dimension $K * V$, with each cell indicating the probability that each word would be used conditional on each topic.[11]

There is no consensus on the definition of framing, let alone the lines that might demarcate one frame from another in a particular empirical application. Depending on the context, it might prove helpful to distinguish claims about "death panels" from claims about "government intrusion," or it might be hair-splitting. Similarly, there is no single choice for K, the number of clusters, that will be valid across applications. Still, the principles that operate in choosing a value of K are the same as those that govern the development of a classification scheme for hand-coding: we aim to identify a parsimonious representation that returns clusters which are theoretically meaningful, internally coherent, and externally differentiated (see also Boyd-Graber et al. 2009). Following the discussion of coding schemes in Hopkins and King (2010), we used an iterative process to choose K. Specifically, we allowed K to vary over the range from 5 to 15, and then chose the value of K that maximized the share of clusters that were clearly interpretable as coherent frames.[12] After

---

[10] Specifically, with the `topicmodels` package for R, we fit LDA using a collapsed Gibbs sampler. The Gibbs sampler was run for 200,000 iterations after discarding 30,000 burn-in iterations, and then thinned by 1/1000. The prior for α is set to 50/K, or 4.1667. Estimation using Variational Expectation-Maximization yields highly similar results. We fit the CTM using Variational Expectation-Maximization, with the two tolerance parameters set to 10^{-5}. The convergence of the Gibbs sampler was checked informally by running it three times with varying starting points and observing essentially identical clustering patterns.

[11] CTM also returns a $(K - 1) \times (K - 1)$ matrix Σ indicating the covariances between the transformed topic probabilities.

[12] The results from models fit with K set to 10 and 14 are available in Figs. 7 and 8 in Online Appendix, respectively.
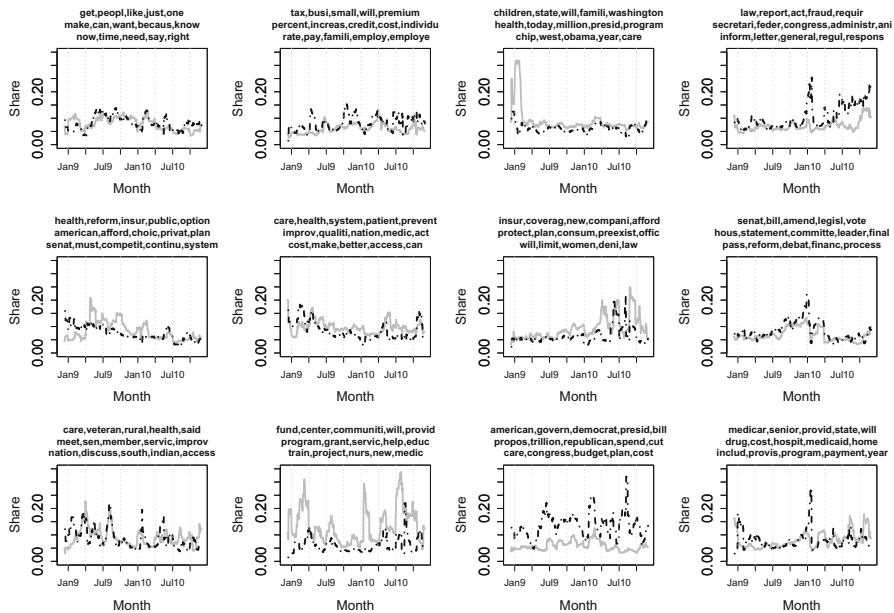
Fig. 2 Variation in topics over time. Model: LDA, fit to 1488 press releases. Each of the 12 clusters depicted here are separate topic returned by an LDA model fit to Senators' press releases. The *lines* depict the share of all press release content which falls into each topic, with *dashed black lines* representing Republican press releases and *solid gray lines* representing Democratic press releases. These shares have been smoothed over the prior 30 days

experimenting with various options and soliciting outside advice, we chose 12 clusters. The core conclusions are not sensitive to this choice, and several results below attest to the utility of the particular model employed (Grimmer and Stewart 2013). Indeed, the chosen solution includes the major frames we had expected ex ante.

Having fit the LDA model, one can then compute various quantities of interest. Here, we focus on one: the share of each document which falls into each of the K topics and how it changes over time. Figure 2 illustrates this quantity of interest for each of the 12 clusters.[13] LDA returns a θ vector for each document indicating the share of that document which falls into each of the K topics. In the figure, we average those shares for all press releases which fall in the prior 30 days. Solid gray lines indicate trends in cluster usage among Democrats, while the dashed black lines do the same for Republicans. To allow readers to assess the coherence of the topics, each cluster is labeled with the 15 word stems with the largest differences in probabilities between that cluster and the corpus overall.[14]

The LDA results in Fig. 2 have several notable features. Certainly, the clusters returned correspond in sensible ways to our expectations. For instance, one

---

[13] For the robustness check using CTM, see Fig. 9 in Online Appendix.

[14] For a more detailed discussion of interpreting the output from topic models, see Boyd-Graber et al. (2009).

prominent Republican cluster throughout the debate draws on word stems including "American," "govern," "Democrat," "presid," "bill," and "trillion" (bottom row, second from right). That frame emphasizes the cost of the legislation and the increasing role of government it entails. It reaches a maximum of 32% of all Republican content. The GOP consistently discussed health care reform as a costly expansion of government. It is also worth noting that the words "death" and "panel" are not a constitutive element of this (or any) cluster.

Another heavily Republican cluster is more oriented towards business costs and taxes, with prominent word stems including "tax," "busi," "small," "will," and "premium," (second column, top row). This frame emerges in the fall of 2009, but is common throughout the debate. It accounts for 17% of Republican rhetoric at its zenith. Notice as well the Medicare-oriented frame, defined by stems including "medicar," "senior," "provid," and "state" (bottom right). This frame is used by both parties, but with a spike in Republican usage in January 2010 that reaches 27%. Just after health care reform passed the U.S. Senate, Republican press releases portrayed health care reform as undermining Medicare. Republican Senators framed health care reform in different ways as the debate unfolded. Republican press releases also emphasized fraud and oversight, as the cluster defined by "law," "report," and "fraud" (top right) illustrates.

The LDA-based partitions also allow us to observe the heavily Democratic clusters, including a "public option" frame—"health," "reform," "insur," "public," and "option" (left column, middle row)—which was more salient in the fall of 2009 than it was later, and which accounts for up to 21% of Democratic rhetoric. The suggestion of Lynch and Gollust (2010) that Democratic rhetoric emphasized affordable insurance coverage and middle-class economic security proves accurate. See, for example, the cluster defined by the stems "insur," "coverag," "new," "compani," and "afford" (third column, middle row), which reaches a maximum of 26%. Another cluster seems to focus on the Children's Health Insurance program, and it gets attention primarily from Democrats before the ACA debate started in earnest (top row, third column). In periods when the ACA is less salient, Democrats' health-care related press releases focus more on funding for community health programs (bottom row, second column).

Not all clusters are dominated by one party. For example, Senators from both parties commonly talk about rural and veterans' health care (bottom row, left column). There is another cluster on legislative procedure that spikes in late 2009, just when the bill came up in Senate (middle row, right column). A third cluster provides generic health care language that was used by both parties (middle row, second column).[15] When considering the difference between the 916 Democratic press releases and the 552 Republican press releases, the average absolute difference across topics is 0.04. When viewed over time, the average standard error within a

---

[15] Given the prior for α, it is not surprising that the LDA model distributes the probability relatively evenly across the twelve categories. The lowest-probability frame still had 0.074 of the total probability mass, while the highest-probability frame had only 0.095.

topic is 0.06 for Republicans and 0.08 for Democrats, indicating substantial punctuation over time.[16]

To be sure, the engaged public does not typically read press releases—its members instead watch and read journalists, who themselves serve as an important filter on what elite rhetoric reaches the public. Accordingly, we also fit an LDA model using the transcripts from 218 appearances on Sunday talk shows by Members of Congress and Obama administration officials between July 2009 and March 2010, when the legislation was signed into law.[17] In this data set, we use 1999 separate word stems that appear in more than 1% of the television appearances.[18]

In television appearances, officials are speaking to a much broader audience, and we observe exactly what attentive citizens observe. The results are presented in Fig. 3—and despite the markedly smaller sample size, they corroborate the core claims made above using Senators' press releases. Republicans are consistently more likely to talk about "trillions," "taxes," and "government." and their discussion of "Medicare" peaks in late 2009 and early 2010. The Democrats are likely to draw from a cluster of words including "public" and "option" early in the debate as well as highlighting issues related to costs and coverage. Later in the debate, the Democrats' rhetoric shifts to focus on insurance companies and affordability. For representatives of both parties, discussions of Senate voting and procedure ramp up as the Senate's formal consideration of the ACA in December 2009 approaches.[19] While our conclusions about framing are drawn primarily from Senators' press releases, they are not specific to that mode of communication. Elite frames on health care show considerable over-time variability.

Measuring Public Opinion

When scholars measure the effects of framing in real-world settings, they almost exclusively rely on closed-ended survey questions (but see Gamson 1992). To increase this study's comparability with prior research, and to ensure that its subsequent results are not the product of its novel method, the Appendix follows that approach. Specifically, it outlines analyses of 30,370 responses to a survey question about whether health care reform is good or bad for the country as a whole asked in 32 surveys between February 2009 and January 2012. Americans' attitudes

---

[16] While most of these clusters are legitimate frames, some are procedural clusters without significant framing value. The cluster labeled with the stems "law," "report," "act," and "fraud" (top right) appears to be one such cluster.

[17] These analyses include television appearances by figures of both parties—including Senators, Representatives, and members of the Obama administration—that dealt with health care reform. Such press appearances are widely seen as attempts at agenda-setting, and present prime opportunities to offer health care-related frames. We collected such press appearances from five media outlets: Fox News, CNN, ABC, NBC, and CBS.

[18] We exclude a small number of names—typically those of the shows' hosts—as well as common English-language stop words.

[19] As with the press releases, we also observe a few catch-all clusters without substantive interpretations.
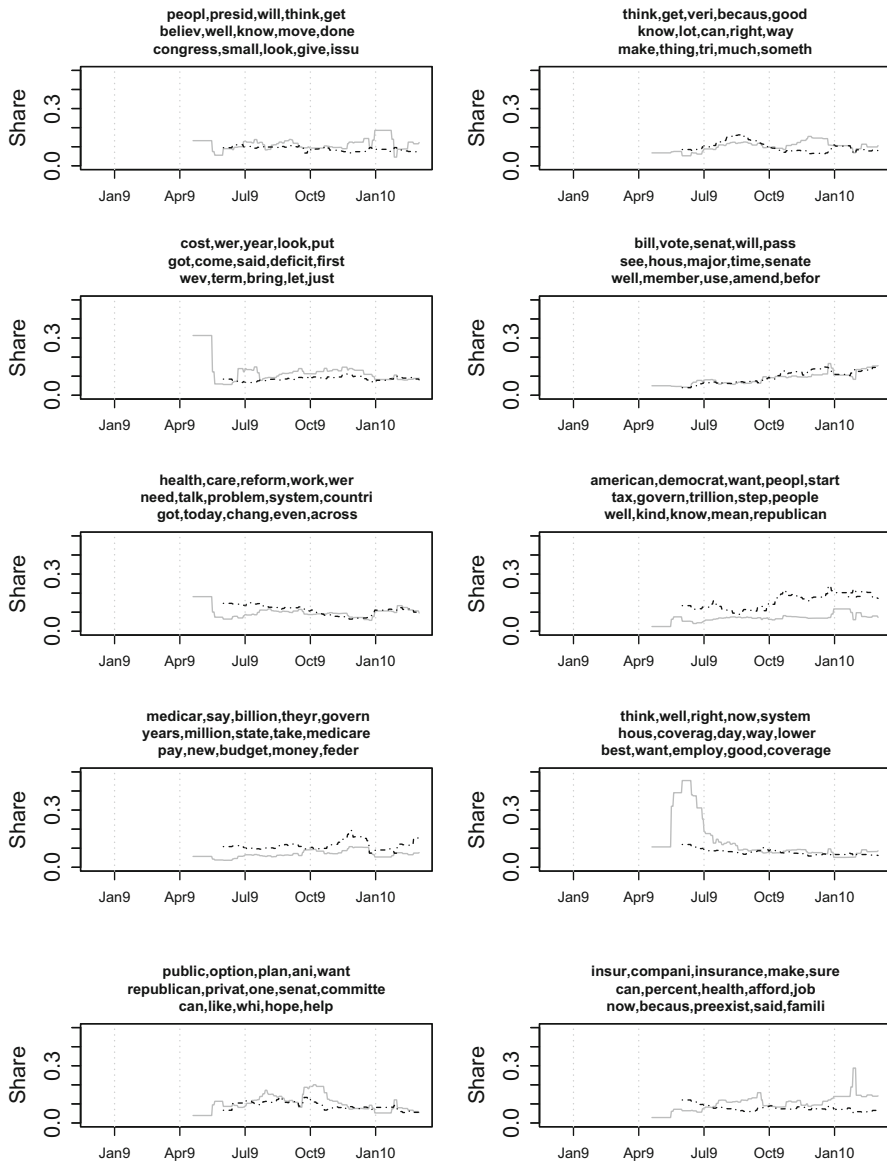
Fig. 3 Variation in topics over time. Topic probabilities smoothed over the prior 28 days. Model: LDA, fit to 218 transcripts from television appearances by administration officials and Congressional leaders. The *dashed black lines* indicate Republican press releases while the *straight gray lines* indicate Democratic press releases

on health care reform correlate primarily with their partisanship and their ethnic and racial identities, correlations which grow and then remain stable over the course of the debate. It is instructive to contrast the stability in the correlates of health care attitudes with the far more punctuated changes in prominent elite frames. This

stability in the correlates of health care attitudes is at odds with the variability of elite frames identified above, and thus with claims that the public is strongly responsive to the frames salient at a particular moment (see also Druckman et al. 2012). By using some of the sudden shifts in elite frames identified above, we can also test framing's influence on sub-groups. Yet Granger tests detailed in the Appendix indicate no such influence: the Republican emphasis on Medicare in the winter of 2010 did not affect those on Medicare, the Democratic emphasis on expanding coverage did not influence those without it, and the Republican emphasis on costs and taxes did not move those with higher incomes.

Still, it is plausible that frames could influence public opinion without shifting overall policy attitudes. Frames might instead operate by shifting the reasons people give for a policy attitude, for example. In this section, we apply LDA to open-ended responses from seven surveys during the health care debate to examine the reasons respondents give for their views. Even given sparse responses that are rarely more than several words, LDA effectively partitions the responses into meaningful clusters and enables scholars to observe the stability of public opinion.

Analyzing Open-Ended Responses

We identified seven telephone surveys conducted from July 2009 through November 2011 which asked open-ended questions about health care reform. The first two were conducted by the Pew Research Center in July 2009 (n = 1506) and November 2009 (n = 1003),[20] while the remaining five were conducted by the Kaiser Family Foundation in May 2010 (n = 1210), October 2010 (n = 1202), March 2011 (n = 1202), June 2011 (n = 1201), and November 2011 (n = 1209).[21] In all, we observe 8,533 responses over a period of 29 months spanning from before the salient public debate until a year after the 2010 mid-term elections. To illustrate the research design, Fig. 1 uses gray lines to depict the first closed-ended survey and the first four open-ended surveys alongside the distribution of press releases by month. As that figure shows, while attempts at framing were already underway in July 2009, the most intense efforts at framing were still to come. What's more, given the punctuated elite frames uncovered in Fig. 2, it is clear that our open-ended baseline took place before key shifts in the content of elites' frames.

While reviewing the open-ended responses, we identified spelling or formatting errors in between 9 and 14% of the Pew open-ended responses and corrected them.[22] We then conducted standard pre-processing, including stemming words (Porter 1980), removing stop words, and removing words of one and two letters. There are 3715 unique word stems in the corpus as a whole, but we then remove all but the 225 word stems which appear in at least 0.25% of the documents. Our analyses focus on the 6363 respondents who used at least one of those 225 words in

---

[20] Pew asked respondents, "what would you say is the main reason you favor/oppose the health care proposals being discussed in Congress?"

[21] The Kaiser question asked, "Can you tell me in your own words what is the main reason you have a favorable/an unfavorable opinion of the health care reform law?"

[22] For the Kaiser surveys, the comparable error rates were 1–2%.

their responses. We initially examine these open-ended responses separately from the press releases described above, as a joint model of two such different modes of text would primarily leverage those words that separate the two formats. We again experiment with various numbers of clusters, from four to twelve, before choosing six as a maximally coherent representation which distinguishes the major frames. Still, as with the press releases, it is worth noting that the information about which respondents were surveyed at which points in time does not formally enter the model. Thus any emerging temporal patterns reflect the data, not the modeling assumptions. Despite being an unsupervised method, can LDA meaningfully structure text as sparse as open-ended responses?

As shown in Fig. 4, the answer is an unambiguous "yes." The figure presents the results using six panels, each of which illustrates the share of responses in a particular category over the first 36 months of the Obama administration. As a reference point, health care reform became salient in August and September of 2009. (For a measure of the salience of health care reform, see the distribution of press releases by month in Fig. 1 in the Appendix. The volume of press releases on health care rises steadily from August 2009 to a peak in December 2009, and then peaks again in March 2010 with the law's passage.) The response shares among supporters and opponents are shown separately, with opponents depicted using dotted red lines and supporters depicted using solid black lines.

The cluster of words at the top left shows one coherent set of responses dominated by concerns about the expansion of government control. Unsurprisingly, this cluster is far more likely to be used by health care reform's opponents: it never accounts for more than 9% of supporters' words and always accounts for at least 18% of opponents' words. While there is some evidence of a decline in its use in October 2010, the general trend is one of consistent usage. Opponents were concerned about health care reform's impact on the scope of government as early as July 2009, before the issue became salient nationally, and remained so well after the polarized debate and the passage of the law. The same is generally true for a separate cluster emphasizing the legislation's cost and tax impacts (top right), a cluster which declines in use but always accounts for at least 19% of opponents' explanations.

These two clusters correspond to clusters of words identified above in the Republicans' press releases. But it is critical to note that these clusters were prominent in public opinion even before health care reform became a central issue in the late summer and early fall of 2009. The timing strongly suggests that the the effect of the elite-level frames was not causal: the core sources of opposition to the ACA were visible long before the legislation itself took shape. While there is some evidence of a shift in rhetoric among health care reform opponents, there is also considerable stability. Consistently, opponents are concerned that the ACA represents an increase in governmental authority that is likely to entail significant costs to the government. Even the catch-all sixth cluster, shown on the bottom right, reinforces this point, as "socialized" is among its prominent words. In no case do we see the words "death" or "panel," an observation which further limits the direct impact of Governor Palin's formulation. In fact, in the first survey after Governor Palin's use of the term, not a single respondent employed the phrase.
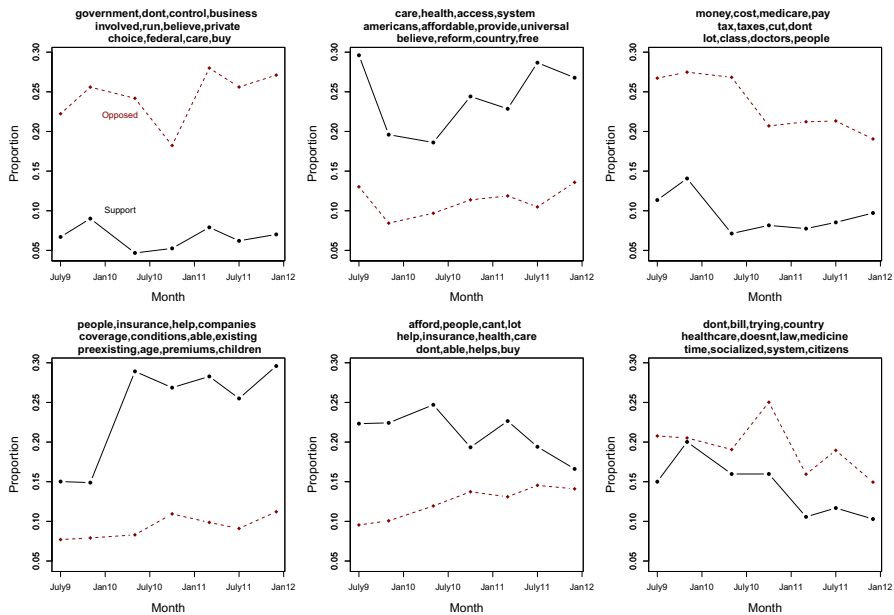
Fig. 4 Variation in open-ended responses for health care reform supporters (*solid black lines*) and opponents (*dashed red lines*) over time. The x-axis indicates time in months. Model: LDA, fit to 6355 open-ended survey responses in seven surveys (Color figure online)

Stability is clear among supporters as well. We do see health care reform supporters shifting into the category defined by words like "people," "insurance," "help," and "companies" (bottom left), although the concomitant decline is from a similar cluster emphasizing the affordability of health insurance (bottom middle). What's more, even after the decline, 17% of all rhetoric by supporters draws on the affordability-related frame. Supporters also make heavy use of a frame emphasizing universal access and affordability (top middle), one defined by words such as "access," "affordable," "provide," and "universal." This frame always accounts for at least 19% of the explanations given by supporters for their views. It is worth noting that this frame is used somewhat more before the salient political debate (30% in July 2009) and afterwards (27% in November 2011), meaning that it might be a default frame among supporters.

Elite-level framing might explain the observed subtle shifts from focusing on universal coverage and on affordability to focusing more specifically on insurance companies. Above, we saw that Democratic Senators focused on words like "insurance," "coverage," and "afford," a focus which grew more pronounced after the passage of the ACA in March 2010. Still, the overall finding is stability among mass-level supporters and opponents alike. There are shifts in the prominence of frames, but all of the core frames that citizens use during and after the salient moments of the health care debate were used to a significant extent beforehand.

LDA is an unsupervised technique, so there is no guarantee that the clusters it returns will be useful for any given analysis. But in this case, they surely are.

Identifying Mass-Elite Convergence in Language

Both in experimental and real-world settings, studies of framing analyze the effect of variation in frame exposure on public opinion, which is measured through closed-ended survey questions. Yet this approach has multiple limitations. One we began to address above: framing might influence the rationale behind a policy attitude without changing the attitude itself, making it plausible that the typical approach underestimates elite influence. Second, to the extent that political elites use polls, focus groups, or other feedback from their constituents in developing frames, frames' purported effects could in fact be reflections of endogeneity. Put differently, politicians' might choose precisely those frames that resonate with pre-existing public opinion, making their causal impact unclear. Above, we saw the utility of models like LDA in making sense of the core arguments at each level. Yet if we apply LDA to a joint data set of press releases and open-ended responses, the resulting clusters are defined to an unhelpful extent by the differences in format across the two types of text. For that reason, this section details an alternative approach to examining the relationship between elite and mass-level word choice over time, despite differences in these two modes of expression. At its base, this technique examines the changing similarity in word choice between elite and mass speech.

We begin by identifying the word stems that appear in more than 1% of the 1,488 press releases and more than 0.25% of the open-ended survey responses. There are exactly 100 such word stems; neither "death" nor "panel" is among them.[23] For parsimony, we focus on the open-ended responses from the three surveys which span the most salient parts of the debate. The first is July 2009, before health care reform became highly salient with the August town hall meetings and the September Presidential address. The second is November 2009, in the thick of the legislative debate. And the third is May 2010, more than 1 month after the passage of the ACA. We then identify the press releases that occur between each of these surveys, giving us three measurements of opinion and two measurements of elite framing. With baseline measures of public language on the issue, we are able to measure how it changes as the issue becomes salient and as political leaders deploy their frames. We can thus identify framing effects as separate from elites' co-optation of public language.

We aim to measure the relationship between elite rhetoric and mass rhetoric, and to chart changes in those relationships over time. We begin by representing a given set of surveys or press releases as a probability distribution across the V words in the overall vocabulary. The distribution of interest is $p(w_v) = \frac{w_v}{\sum_{i=1}^{V} w_v}$, calculated simply as the share of the total words in a given corpus accounted for by each particular word w_v. We can then calculate the distance d between the distribution of words in a

---

[23] Although "death panel" is technically a bigram, the absence of these unigrams implies its absence as a two-word phrase.

group of press releases $p_{pr}(w_v)$ and the distribution of the same words among the public $p_{sur}(w_n)$ using any of several distance metrics. We focus here on *changes* in the distance metrics when comparing opinions prior to a set of press releases with those subsequent to the press releases. Declines in distance indicate that when elites adopt language in their press releases, the members of the public subsequently shift their vocabulary by adopting the same words.

To give some sense of these distributions, we plot the difference in the distributions of words from four open-ended survey questions in Fig. 5. On the x-axis, we plot each word's overall frequency on the logged scale, while on the y-axis we plot the difference in word usage between supporters (positive numbers) and opponents (negative numbers). We divide this difference by each word's standard deviation across documents, reducing the magnitude for words that are more variable on average. Here again, the stability of the popular vocabulary over time is on display, with opponents consistently using words like "pay," "cost," and "money" and supporters using words like "help," "health," "care," and "afford."

Distance in Speech

How can we use these probability distributions to measure framing effects? To take a specific example, we measure the effect of Democratic frames on American citizens in the early stage of the debate by creating three vectors of word
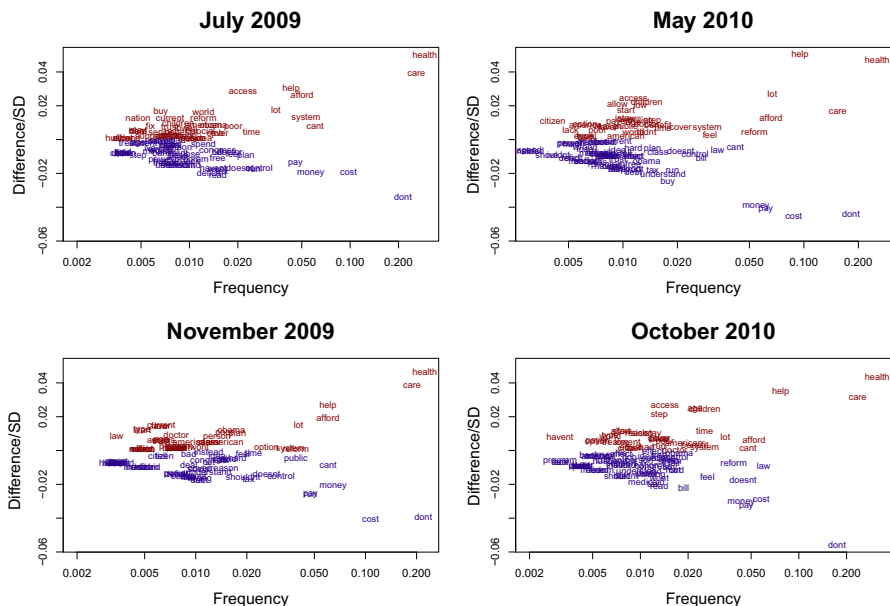

Fig. 5 This figure presents the difference in word usage for supporters and opponents (y-axis; rescaled by dividing by the word's standard deviation across responses) as a function of each word's overall frequency on the log scale (x-axis). Words that appear more frequently in supporter's comments are in *red* and above zero on the y-axis, while those used more frequently by opponents are in *blue* and below zero on the y-axis (Color figure online)

distributions. The first and second denote the distribution of words in the July 2009 Pew open-ended responses and the November 2009 Pew open-ended responses, while the third denotes the distribution of words in Democratic press releases made public between those surveys. In July 2009, for example, the stems "health" and "care" accounted for 29% of the common words used by citizens, with "don't" and "cost" being prominent as well. We then measure the distance between this distribution and the distribution of words in subsequent press releases using the symmetrized KL divergence metric, a common distance metric in content analyses.[24] The result: 1.48, a number that is not meaningful on its own but which provides a baseline estimate of the relationship between elite language and mass language. By replicating this process for the same press releases and the open-ended responses measured *after* the press releases were made public, we can measure whether supporters grew more likely to use specific words after elites employed those words. The distance between the word distribution in the press releases and that in the subsequent open-ended responses measured in November is 1.65. In the early stages of the debate, the public did not grow more likely to use the words in Democratic press releases.

To measure whether such differences are meaningful, we use bootstrapping, randomly drawing 10,000 new data sets of press releases and open-ended responses and repeating the measurement procedure outlined above. The increase in the symmetrized KL divergence between Democratic press releases and citizens' language is in the unexpected direction but insignificant for the early period between July and November ($p = 0.95$, one-sided), as illustrated on the left side of Fig. 10 in the Online Appendix. We see the same, insignificant result on the Republican side ($p = 0.925$). However, when we conduct the same analyses for the second period, between November 2009 and May 2010, we see a substantial and significant decline in distance ($p < 0.001$), one that holds for Republican Senators and citizens' language as well ($p = 0.003$). We reach the same conclusion when estimating a more intuitive metric, the average absolute difference between the word frequencies in the open-ended responses and those in the press releases. For Democrats, that figure drops from 0.0102 to 0.0098 between November 2009 and May 2010, while for Republicans it drops from 0.0094 to 0.0088. The relationship between Senators' press releases and the electorate's language increases notably during the most salient phase of the health care reform debate. It is not simply that political elites and the public use similar words: the similarity actually grew after the elites' public statements.

We examine how robust this conclusion is to the use of alternate measures such as Euclidean distance or the Pearson's correlation (with Fig. 10 in the Online Appendix illustrating the results).[25] We again observe that between November and May, mass-level word choice became more closely related to the word choice in

---

[24] The Kullback-Liebler divergence between two distributions P_{pr} and P_{sur} is defined as $\sum_i P_{pr}(i) Log\left(\frac{P_{pr}(i)}{P_{sur}(i)}\right)$. To symmetrize this metric, we add the divergence from P_{pr} to P_{sur} to the divergence from P_{sur} to P_{pr}.

[25] Keep in mind that the Pearson's correlation is negatively correlated with distance, so upward arrows in the right panel of Fig. 10 indicate an increasingly close relationship.

Senators' press releases. There are some differences across these metrics which are noteworthy.[26] Also, using the Pearson's correlation, the result for Republicans in the second period is not statistically significant ($p = 0.143$). Still, the broad pattern holds, and was confirmed using still other distance metrics such as the Canberra and Manhattan metrics. The results for the first period are inconclusive—they are typically insignificant and vary across metrics. For the second period, however, they are consistent. From November 2009 to May 2010, citizens increasingly used the words common among Democratic elites, and the same appears true for Republican Senators' word choices as well. Elite rhetoric might not influence aggregate opinion in demonstrable ways, but it does shape how citizens talk about an issue. It does so primarily after November 2009, a fact which further motivates this study's emphasis on the period from July 2009 to May 2010. Also, in contrast to the arguments of several commentators, there is no evidence that Republican words were more likely to be adopted than Democratic words.

Discussion and Conclusion

For mass- and elite-level opponents of the ACA, the legislation's cost and the increased governmental role it authorized were central reasons for opposition. Consider the following argument: "This may be your last chance to weigh the consequences of taking the first step toward establishment of socialized medicine in the United States...When costs get out of line...there are three possible courses of action. The first is to reduce the benefits; the second is to increase taxes; the third is to impose government controls of the services." Certainly, that argument would have been an unsurprising addition to the press releases and TV apperances analyzed above. But that particular quotation comes from the 1965 debates over Medicare, when it was written by the President of the American Medical Association. Even across decades, the rhetoric used to support and oppose health care reform has been notable in its consistency, a point our mass-level results reinforce. Such consistency places a low ceiling on the influence that politicians' frames at any one moment can have: a variable cannot explain a constant. This observation in turn suggests the value of studying not just specific political debates but the contours of public debate over broad swaths of time (see also Smith 2007; Noel 2014). Framing is not simply a source of elite influence over public opinion, but a complex, interactive process that unfolds over years.

This manuscript applied automated content analyses to the health care rhetoric of public officials and American citizens from 2009 to 2010. These tools enable us to summarize the central arguments for and against health care reform and to analyze shifts over the course of the debate. We see, for example, that supporters at the elite and mass levels alike emphasized the expansion of insurance and increased access

---

[26] The Euclidean distance detects statistically significant declines in the first time period from late July to November 2009, with p-values of 0.004 for Democratic press releases and 0.003 for Republican press releases.

to health care. Yet supporting citizens used these arguments to explain their views even before the health care reform debate came to dominate headlines, a fact which suggests the limited influence of elite rhetoric on overall evaluations. As detailed in the Appendix, more traditional tests of real-world framing effects using Granger tests do not indicate any substantive effect on public attitudes among the sub-groups targeted by specific rhetoric.

Analyzing the fit between elite and mass word choice enables us to measure communication effects—whether framing or not—with increased precision and subtlety. When doing so, we see that citizens grow more likely to use the same words as their elected officials over the course of a public debate. In this case, such effects appear to be roughly symmetric: there is an increasing correlation between the word choices of both Democratic and Republican Senators and the citizenry at large. Even when it does not move overall evaluations, elite rhetoric appears to shape the rationale underpinning those evaluations. This pattern of findings is broadly consistent with the results of Lodge and Taber (2013), which emphasizes unconscious, often affective processes in shaping public opinion. For Druckman (2004), "framing effects constitute one of the most stunning and influential demonstrations of irrationality" (p. 671). But as this paper illustrates, the absence of framing effects does not imply rational decision-making, either.

The changing distance between elite-level word choice and mass-level word choice is a metric that could be productively employed and developed in future analyses, whether experimental or observational. One might study which subsets within the electorate are especially likely or unlikely to adopt elite rhetoric, for example. Alternatively, one might analyze the connection between individual legislators and the language and opinions of their constituents. Scholars should certainly examine whether similar findings emerge in other policy realms, including those that are lower in salience and where citizens lack extensive personal experience. The fact that health care reform was the subject of a similar debate in 1993–1994 (Jacobs and Shapiro 2000; Kriner and Reeves 2014) might make this issue atypical as well—and suggests the value of studying policy debates on less salient issues. Methodologically, future research might also develop unified statistical models that allow us to track shifts in vocabulary across different types of speech.

Scholars of public opinion have long recognized the trade-offs inherent in using open-ended or closed-ended survey questions (Payne 1951). Still, research on framing and on public opinion has relied almost entirely on closed-ended questions. The growing set of tools for automated content analysis has the potential to shift this balance, and to improve the measurement of key concepts in public opinion research. For framing to operate, the cognitive accessibility and applicability of specific considerations must change—and automated content-analytic tools provide a novel way to measure whether they do. By using press releases, television appearances, and open-ended survey questions to test hypotheses about framing, this paper provides an example of how text and automated content analysis can provide new vantagepoints on old questions.

Appendix: Framing and Sub-group Effects

Frames and Sub-group Opinion

Here, we proceed as analyses of real-world framing typically do and examine closed-ended survey data. Doing so allows us to assess the stability of public opinion on health care reform, which might place an upper bound on framing effects: variable elite frames cannot explain stable opinions. Analyzing citizens' closed-ended assessments of health care reform also enables us to measure sub-group trends that provide tests of targeted framing effects. The results advance the claim that real-world framing effects are limited, as we find no evidence that sub-groups targeted by frames subsequently shift their attitudes. At the same time, the limitations of this common approach help motivate the use of open-ended responses to measure public opinion and mass-elite interactions.

To analyze public health care attitudes, we turn to 32 telephone surveys of American adults conducted by the Kaiser Family Foundation between February 2009 and January 2012. With at least 1,200 respondents, the Kaiser surveys jointly provide us with information on 30,370 Americans' attitudes toward health care reform.[27] For this analysis, we focus on a single question: "Do you think the country as a whole would be better off or worse off if the president and Congress passed health care reform, or don't you think it would make much difference?"[28] This question was asked consistently throughout the debate, providing us with a common metric of respondents' health care reform attitudes. The responses are coded on a scale from 1 to 3, with 1 indicating that the country will be worse off, 2 indicating no difference, and 3 indicating that the country will be better off.

Drawing on Gelman and Hill (2006), the analysis proceeds by estimating separate linear models of attitudes toward health care reform in each month. The models include basic demographics, such as a respondent's gender, race, ethnicity, age and education in years, and income in thousands of dollars. They also include measures for self-reported Republicans and Independents, although in most administrations the survey did not push Independents to "lean" to a party. In addition, the models employ a few health care-related measures that gauge respondents' self-interest in the debate. These include indicator variables for receiving Medicare or Medicaid as well as an indicator for lacking health insurance and a five-category measure of self-reported health.[29] To analyze any shifts in

---

[27] Prior to listwise deletion, there are 43,887 respondents. The results remain quite similar when omitting the measure of respondents' income, which itself accounts for 47% of the missingness.

[28] After the bill's passage, the language was modified to ask about conditions "under the new health reform law."

[29] Not all respondents over 65 report insurance through Medicare, although a significant majority do.
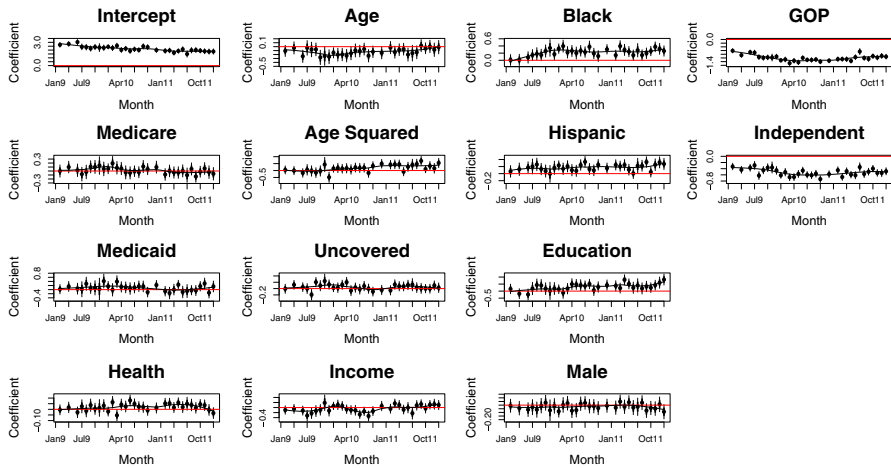
Fig. 6 Data: 30,370 American adults interviewed in 32 monthly Kaiser Family Foundation surveys between February 2009 and January 2012. The x-axis indicates the month. The dependent variable is each respondent's assessment that health care reform is good for the country as a whole, measured on a scale from 1 ("worse off") to 3 ("better off"). *Each dot* indicates the mean coefficient estimate for an OLS model of that month's data, with *lines* reflecting 95% confidence intervals

public opinion over time, the analysis ran separate linear regression models for each month's survey. It then aggregated each variable's coefficients into a single plot, presented in Fig. 6. The x-axis indicates time while the y-axis indicates the size of each coefficient. Note that the scale of the y-axis differs by variable. Each dot indicates the corresponding month's mean coefficient, while the surrounding line depicts the 95% confidence interval.

The first thing to notice: partisanship is a uniquely powerful predictor of attitudes toward health care reform, and increasingly so as the debate went on (see also Kriner and Reeves 2014). From the coefficient magnitudes, we see that identifying as a Republican or as an Independent leads to markedly reduced assessments of health care reform's benefits. There are small but discernible changes over time in these coefficients, with increasing evidence of partisan polarization as the debate unfolds. We see, for example, that Republican identifiers are less negative on health care reform during the spring and summer of 2009, when President Obama was still enjoying a honeymoon and when health care had yet to take center stage. GOP assessments of health care reform drop notably in August of 2009, as angry town hall meetings captured national media attention. They then remain at a plateau during the fall of 2009 before dropping again in January 2010, just after the ACA passed the U.S. Senate. Both the GOP and Independent coefficients reach their nadir around the spring of 2010, when the ACA was passed and signed into law. As with several other covariates, including the indicators for race and ethnicity, the pattern for partisanship shows initial polarization in the early months of the debate and then impressive stability thereafter. Notice as well that the objective measures of self-interest, such as self-reported health or receipt of Medicare or Medicaid, have very

little predictive power at any point during the debate. Those without health care coverage are mildly more positive in their assessments, but there is little evidence those assessments are bolstered by Democrats' use of related frames just after the law's passage. This contrasts with findings for subjective measures of self-interest reported in Henderson and Hillygus (2011). If anything, Medicare recipients appear slightly more positive about health care reform in early 2010, when Republican Senators were emphasizing its reductions in Medicare spending (but see Campbell 2011).

Figure 6 also demonstrates the stability of opinions, especially after the issue of health care reform became salient in August and September of 2009. The changes afterward tend to be small, such as the potential uptick in positive views among the well-educated in late 2011, the gradual increase in support among the elderly, or the gradual decline in the baseline respondent. This stability of sub-group opinion contrasts with the punctuation in frames identified above: the sudden shifts in elite framing do not match up with the stability of sub-group opinion. In that way, these results amplify the conclusion of Druckman et al. (2012) that choosing to be exposed to certain initial health care frames induces opinion stability thereafter.

Granger Tests of Framing Effects

Expecting a single frame to produce a homogeneous effect across the population might set the bar for framing too high. Prior research shows that partisans respond to frames or arguments in quite different ways, and Americans' responses to health care frames might also hinge on the frames' personal relevance. We thus consider whether three of the pronounced shifts in framing identified in the article disproportionately influenced the sub-groups targeted by those frames. First, if the sudden spike in Medicare-related rhetoric by Republicans in January 2010 was influential, those on Medicare might become less sanguine about health care reform just afterward. Second, if the Democratic emphasis on extending coverage and increasing its affordability after the law's passage was influential, we might expect the prominence of that frame to be especially powerful among those without health insurance. Alternately, the Republicans drew from a frame about rising taxes and increased business costs at multiple times, a frame that might have been more influential among well-to-do respondents.

To test these possibilities, we begin with the 20 months for which we have Senators' press releases and public opinion data, which covers the period from February 2009 to December 2010. We extract the relevant coefficient from the regressions conducted above which predict assessments about health care reform's impact on the country as a whole. We then use Granger tests to examine the sequencing of the frames and any shifts in sub-group opinion during the subsequent month. To measure frames, we consider both the difference in partisan usage of the frame and the share of usage by the party deploying the frame strategically, as Table 1 shows. On the Medicare frame, for example, we consider the difference between Republican and Democratic usage as well as Republican usage alone. Frames are measured based on the maximum share of a party's discourse that drew

Table 1 This table reports the results of Granger tests when examining the relationship between the "Medicare," "Coverage," and "Taxes" frames presented in Fig. 2 and the opinions of relevant sub-groups as presented in Fig. 6

	F statistic	*p* value
ΔMedicare	0.134	0.719
Republicans—Medicare	0.436	0.519
Democrats—Medicare	0.057	0.814
ΔCoverage	0.184	0.674
Republicans—Coverage	3.440	0.082
Democrats—Coverage	1.051	0.321
ΔTaxes	1.706	0.210
Republicans—Taxes	0.261	0.617
Democrats—Taxes	2.114	0.165

These tests examine how frames predict subsequent public opinion

Table 2 This table reports the results of Granger tests when examining the relationship between the "Medicare," "Coverage," and "Taxes" frames presented in Fig. 2 and the opinions of relevant sub-groups as presented in Fig. 6

	F statistic	*p* value
ΔMedicare	0.004	0.952
Republicans—Medicare	0.084	0.776
Democrats—Medicare	0.067	0.799
ΔCoverage	2.561	0.129
Republicans—Coverage	0.053	0.821
Democrats—Coverage	1.957	0.181
ΔTaxes	0.560	0.465
Republicans—Taxes	0.002	0.966
Democrats—Taxes	1.692	0.212

These tests examine how sub-group opinion predicts subsequent frames

on related language in the prior month, although the results are robust to the use of means instead.

As the consistently high p-values indicate, there is no evidence that increased use of a frame shifts the opinion of the relevant sub-group in the subsequent month. This same pattern of null results holds if we increase the lag to 2 months. It also holds when we multiply the frame shares by the total number of press releases in each month, which allows us to measure the frames' salience. Groups targeted by frames are not differentially responsive to those frames.

By contrast, if political rhetoric is responsive to changes in public opinion, it is less clear precisely whose opinions rhetoric should track—should it track overall opinions or those of key sub-groups? Still, it is noteworthy that when we repeat these analyses allowing sub-group opinion to lead political rhetoric on related issues, we do not find significant relationships either, as illustrated by Table 2.

References

Arceneaux, K., Johnson, M., & Murphy, C. (2012). Polarized political communication, oppositional media hostility, and selective exposure. *The Journal of Politics, 74*(01), 174–186.

Barabas, J., & Jerit, J. (2010). Are survey experiments externally valid? *American Political Science Review, 104*(2), 226–242.

Baumgartner, F. R., & Bryan, D. J. (2013). Policy Agendas Project. http://www.policyagendas.org/.

Baumgartner, F. R., Berry, J. M., Hojnacki, M., Leech, B. L., & David, C. K. (2009). *Lobbying and policy change: Who wins, who loses, and why.* Chicago: University of Chicago Press.

Baumgartner, F. R., De Boef, S., & Boydstun, A. E. (2008). *The decline of the death penalty and the discovery of innocence.* New York: Cambridge University Press.

Berinsky, A. J. (2007). Assuming the costs of war: Events, elites, and American public support for military conflict. *Journal of Politics, 69*(4), 975–997.

Blei, D. M., & Lafferty, J. D. (2006). Correlated topic models. *Advances in Neural Information Processing Systems, 18*, 147.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research, 3*, 993–1022.

Boyd-Graber, J., Chang, J., Gerrish, S., Wang C., & Blei, D. (2009). Reading tea leaves: How humans interpret topic models. In *Proceedings of the 23rd annual conference on neural information processing systems.*

Bullock, J. G. (2011). Elite influence on public opinion in an informed electorate. *American Political Science Review, 105*(3), 496–515.

Campbell, A. L. (2011). Policy feedbacks and the impact of policy designs on public opinion. *Journal of Health Politcs, Policy, and Law, 36*(6), 961–73.

Chong, D., & James, N. D. (2007a). A theory of framing and opinion formation. *Journal of Communication, 57*, 99–118.

Chong, D., & James, N. D. (2007b). Framing public opinion in competitive democracies. *American Political Science Review, 101*(04), 637–655.

Chong, D., & Druckman, J. N. (2010). Dynamic public opinion: Communication effects over time. *American Political Science Review, 104*(4), 663–680.

Chong, D., & Druckman, J. N. (2011). Public-elite interactions: Puzzles in search of researchers. In Y. S. Robert & L. Jacobs (Eds.), *The Oxford Handbook of American Public Opinion and the Media.* New York: Oxford University Press.

De Vreese, C. H. (2003). *Framing Europe: Television news and European integration.* Amsterdam: Het Spinhuis.

Druckman, J. N. (2003). On the limits of framing effects: Who can frame? *Journal of Politics, 63*(4), 1041–1066.

Druckman, J. N. (2004). Political preference formation: Competition, deliberation, and the (ir) relevance of framing effects. *American Political Science Review, 98*(04), 671–686.

Druckman, J. N., & Leeper, T. J. (2012). Learning more from political communication experiments: Pretreatment and its effects. *American Journal of Political Science, 56*(4), 875–896.

Druckman, J. N., Fein, J., & Leeper, T. J. (2012). A source of bias in public opinion stability. *American Political Science Review, 106*(2), 430–454.

Edwards, G. C. (2009). *The strategic president: Persuasion and opportunity in presidential leadership.* Princeton, NJ: Princeton University Press.

Eisenstein, J., O'Connor, B., Smith, N. A., & Xing, E.P. (2010). A latent variable model for geographic lexical variation. In *Proceedings of the 2010 conference on empirical methods in natural language processing.* Association for Computational Linguistics, pp. 1277–1287.

Entman, R. M. (2004). *Projections of power: Framing news, public opinion, and U.S. Foreign Policy.* Chicago: University of Chicago Press.

Gamson, W. A. (1992). *Talking politics.* New York: Cambridge University Press.

Gelman, A., & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models.* New York: Cambridge University Press.

Grimmer, J. (2010). A Bayesian hierarchical topic model for political texts: Measuring expressed agendas in senate press releases. *Political Analysis, 18*(1), 1–35.

Grimmer, J., & King, G. (2011). General purpose computer-assisted clustering and conceptualization. *Proceedings of the National Academy of Sciences, 108*(7), 2643–2650.

Grimmer, J., & Stewart, B. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis* (Forthcoming).

Hayes, D. (2008). Does the messenger matter? Candidate-media agenda convergence and its effects on voter issue salience. *Political Research Quarterly, 61*(1), 134–146.

Henderson, M., & Sunshine Hillygus, D. (2011). The dynamics of health care opinion, 2008–2010: Partisanship, self-interest, and racial resentment. *Journal of Health Politics, Policy, and Law, 36*(6), 945–960.

Hill, S. J., Lo, J., Vavreck, L., & Zaller, J. (2013). How quickly we forget: The duration of persuasion effects from mass communication. *Political Communication, 30*, 521–547.

Hopkins, D. J., & King, G. (2010). A method of automated nonparametric content analysis for social science. *American Journal of Political Science, 54*(1), 229–247.

Huber, G. A., & Paris, C. (2012). Assessing the programmatic equivalence assumption in question wording experiments: Understanding why americans like assistance to the poor more than welfare. *Public Opinion Quarterly, 77*, 385–397.

Iyengar, S., & Kinder, D. R. (1987). *News that matters*. Chicago: University of Chicago Press.

Jacobs, L. R., & Burns, M. (2004). The second face of the public presidency: Presidential polling and the shift from policy to personality polling. *Presidential Studies Quarterly, 34*(3), 536–556.

Jacobs, L. R., & Shapiro, R. Y. (2000). *Politicians don't pander: Political manipulation and the loss of democratic responsiveness*. Chicago: University of Chicago Press.

Jacobs, L. R., Page, B. I., Burns, M., McAvoy, G., & Ostermeier, Eric. (2003). What presidents talk about: The nixon case. *Presidential Studies Quarterly, 33*(4), 751–771.

Kellstedt, P. M. (2003). *The mass media and the dynamics of American racial attitudes*. New York: Cambridge University Press.

Kriner, D. L., & Reeves, A. (2014). Responsive partisanship: Public support for the Clinton and Obama health care plans. *Journal of Health Politics, Policy and Law, 39*(4), 717–749.

Lauderdale, B. E., & Clark, T. S. (2014). Scaling politically meaningful dimensions using texts and votes. *American Journal of Political Science, 58*(3), 754–771.

Lecheler, S., de Vreese, C., & Slothuus, R. (2009). Issue importance as a moderator of framing effects. *Communication Research, 36*(3), 400–425.

Leeper, T., & Slothuus, R. (2015). *Can citizens be framed? How information, not emphasis, changes opinions*. Aarhus C: Aarhus University.

Lenz, G. S. (2013). *Follow the leader?: How voters respond to politicians' policies and performance*. Chicago: University of Chicago Press.

Lodge, M., & Taber, C. S. (2013). *The rationalizing voter*. Cambridge: Cambridge University Press.

Lynch, J., & Gollust, S. E. (2010). Playing fair: Fairness beliefs and health policy preferences in the United States. *Journal of Health Politics, Policy and Law, 35*(6), 849–887.

Mutz, D. C. (1994). Contextualizing personal experience: The role of mass media. *The Journal of Politics, 56*(3), 689–714.

Nelson, T. (2011). Issue framing. In R. Y. Shapiro & L. Jacobs (Eds.), *The Oxford Handbook of American Public Opinion and the Media* (pp. 189–203). New York: Oxford University Press.

Nelson, T. E., Clawson, R. A., & Oxley, Z. M. (1997). Media framing of a civil liberties conflict and its effect on tolerance. *American Political Science Review, 91*(3), 567–583.

Noel, H. (2014). *Political ideologies and political parties in America*. New York: Cambridge University Press.

Nyhan, B. (2010). Why the "Death Panel" myth wouldn't die: Misinformation in the health care reform debate. *The Forum, 8*(1), 1–24.

Payne, S. L. (1951). *The art of asking questions*. Princeton, NJ: Princeton University Press.

Porter, M. F. (1980). An algorithm for suffix stripping. *Program: Electronic Library and Information Systems, 14*(3), 130–137.

Prior, M. (2007). *Post-broadcast democracy: How media choice increases inequality in political involvement and polarizes elections*. New York: Cambridge University Press.

Quinn, K. M., Monroe, B. L., Colaresi, M., Crespin, M. H., & Radev, D. R. (2010). How to analyze political attention with minimal assumptions and costs. *American Journal of Political Science, 54*(1), 209–228.

Roberts, M. E., Stewart, B. M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S. K., et al. (2014). Structural topic models for open-ended survey responses. *American Journal of Political Science, 58*(4), 1064–1082.

Scherer, M. (2010). The White House scrambles to Tame the news cyclone. *Time* March 4.

Scheufele, D. A., & Iyengar, S. (2012). The state of framing research: A call for new directions. In K. Kenski & K. H. Jamieson (Eds.), *The Oxford handbook of political communication theories* (pp. 1–26). New York: Oxford University Press.

Shapiro, R. Y., & Jacobs, L. (2010). Simulating representation: Elite mobilization and political power in health care reform. *The Forum*, *8*(1), 1–15.

Slothuus, R., & de Vreese, C. H. (2010). Political parties, motivated reasoning, and issue framing effects. *Journal of Politics*, *72*(3), 630–645.

Smith, M. A. (2007). *The right talk: How conservatives transformed the great society into the economic society*. Princeton, NJ: Princeton University Press.

Sniderman, P. M., & Theriault, S. M. (2004). The structure of political argument and the logic of issue framing. *Studies in Public Opinion: Attitudes, Nonattitudes, Measurement Error, and Change* (pp. 133–65). Princeton, NJ: Princeton University Press.

Steinhauer, J., & Pear, R. (2011). G.O.P. Newcomers set out to undo Obama victories. *New York Times*.

Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science*, *50*(3), 755–769.