The Clues in the News: Unpacking Thermostatic Responsiveness to Policy [1]

Fabian Neuner
University of Michigan
fgneuner@umich.edu

Stuart Soroka
University of Michigan
ssoroka@umich.edu

Christopher Wlezien
University of Texas at Austin
Wlezien@austin.texas.edu

Abstract: There is a sizeable literature finding evidence of thermostatic responsiveness across a range of salient policy domains and countries. We thus far have only a partial sense for what drives responsiveness to policy change at the individual level, however. One possibility is that individuals learn what they need to know from the mass media, but there is very little work exploring either the prevalence of relevant policy cues in media content, or citizens' abilities to pick up on those cues. This paper represents a first attempt to examine both, first through an automated content analysis of 35 years of reporting on both defense and welfare in the *New York Times* and the *Washington Post*; then through an exploratory coding exercise fielded using MTurk. Results point to the surprising frequency, and impact, of policy cues in media content, and the ease with which citizens correctly identify policy content in news; the one remaining (and we suspect difficult) task is the design of an experiment that allows for an exploration of the conditions under which there is thermostatic responsiveness at the individual level.

"Few people read the New York Times or the Washington Post, and even these papers really only give us a glimmer of insight into what government is doing." (Soroka and Wlezien 2010: 30)

The reciprocal relationship between public policy and public preferences is central to representative democratic governance. That said, there have since the dawn of democracy been real concerns about both governments' willingness to respond to citizens and whether this even makes sense given questions about citizens' ability to provide useful input to governments. Converse's (1964) work has been particularly influential; and there is now a vast literature chronicling the political ignorance of the average citizen. At the same time, there is a growing body of work on "thermostatic responsiveness" suggesting that citizens do in fact adjust their preferences for policy change based in part on recent policy.[2] This is not true for all policy domains, of course. But in salient policy domains,

[1] Prepared for delivery at the Toronto Political Behaviour Workshop, Ryerson University, November 2015. This work was funded in part by the *Philip Converse and Warren Miller Fellowship Fund in American Political Behavior*, awarded by the Institute for Social Research at the University of Michigan to Fabian Neuner and Stuart Soroka. The paper also relies on on Lexicoder software development by Mark Daku and Lori Young, and a good deal of database work from Daniel Hiaeshutter-Rice.

[2] As regards thermostatic responsiveness, see, e.g., Wlezien (1995; 1996), Erikson, et al (2002), Soroka and Wlezien (2005), Jennings (2009), Soroka and Wlezien (2010), Wlezien and Soroka (2012). As regards the public's inattentiveness and information, see, e.g., Berelson, et al (1954), Converse (1964; 1970), Bennett (1988), Page and Shapiro (1992), Delli Carpini and Keeter (1996), Popkin and Dimock (1999).

there is an accumulating body of evidence suggesting that publics can and do respond to policy change in the US and other countries.

How can we reconcile evidence of thermostatic responsiveness with evidence of uninformed publics? Past work (i.e., Page and Shapiro 1992) has focused on the advantages of aggregation.[3] Aggregation cannot entirely account for thermostatic responsiveness, however. What accounts for the underlying trend? Regular citizens do not directly observe policy change or read the Federal Register for that matter. Why does opinion react to policy change? Our account is very simple, and based on the following propositions: (a) thermostatic responsiveness requires only very basic levels of knowledge about policy and policy change, (b) this very basic information is readily available in media content, and (c) citizens are readily able to pick up on these informational media cues.

This paper offers a next step in testing these hypotheses. Recent work (Soroka and Wlezien 2015) uses automated content analysis to examine over 100,000 news stories on welfare and defense policy in US. That research finds a surprisingly large number of informational cues – both direct, and indirect – about the direction of policy in news content. We first revisit some of those findings here. We then draw on the same body of media content in a coding exercise, validating automated coding of policy direction cues prevalent in news content with human coding using Amazon Mechanical Turk (MTurk). We see this not just as a test of the automation, but a test of whether regular citizens are able to extract information about policy from news content. We are buoyed by what appear to be very strong results in both our content-analytic data and coding exercises, which suggest that some basic conditions for mass-media-informed thermostatic responsiveness are met. In a concluding section, then, we consider the possibility of experiments designed to elicit thermostatic responsiveness, in order to further explore who is responding and what exactly they are responsive to.

How easy is it for citizens to learn about policy from news articles with seemingly complex, and sometimes quite subtle, informational cues about public policy change? An answer to this question is, we believe, central to understanding how and why representative democracy works. The work below thus takes a first step towards identifying some of the micro-level dynamics behind thermostatic responsiveness to policy change.

Mass Media & Thermostatic Responsiveness

If there is thermostatic public responsiveness, people will adjust their preferences for "more" or "less" policy in response to policy change, favoring less (more) policy in the wake of policy increases (decreases), other things being equal. Formally, the public's relative preference (R) represents the difference between the publics preferred level of policy (P^*) and policy (P) itself,

$$R_t = P_t^* - P_t \, , \qquad (1)$$

[3] Though see Althaus (2003) for a more complicated and less complimentary view of aggregation.

where the subscripted t indicates time.[4] R can change because either P or P^* changes. Because we often do not have measures of P^* and also because the three variables are not measured using the same metric, it is necessary to rewrite the equation as follows:

$$R_t = B_1\, U_t + B_2\, P_t + e_t\ , \qquad (2)$$

where U is a set of additional exogenous predictors of P^*.

If the coefficient for policy (B_2) in equation 2 is less than 0, we have evidence of thermostatic responsiveness. This often is true but it sometimes is not (Wlezien 1995). Moreover, even when the coefficient is less than 0, its size varies, owing partly to characteristics of issues and partly to political institutions (Soroka and Wlezien 2010; Wlezien and Soroka 2012). First, the public is more responsive in highly salient domains. This is highly intuitive. Second, political context matters. In particular, federalism dampens public responsiveness, seemingly because it lowers the clarity of information about what policymakers are doing at different levels of government.

Thermostatic public responsiveness clearly is important. It provides the basis for holding policymakers accountable. It also provides informed signals that policymakers could effectively represent. Establishing that there is thermostatic public responsiveness, and understanding how it varies across issues and political context, is thus of some significance. There is, as we have noted, a growing body of work that does exactly this. (See citations in note 2.) But there is relatively little work that seeks to understand the underlying mechanism(s) driving the relationship. We know that most people pay little attention to politics and have little knowledge about it. How do people get the information? How is it that the patterns are so strikingly similar across various subgroups, including those with low levels of education and political interest?

What do we know? To begin, we know that thermostatic responsiveness does not require a high level of information. This must be true – otherwise we cannot account for the patterns that we (and others) observe. We have posited elsewhere (Soroka and Wlezien 2010) that people only need to have a sense for the direction of policy change – whether policy has gone up or down – and perhaps also the magnitude – whether it has gone up by a little or a lot. These are very basic informational cues. But even these cues need to come from somewhere.

We know that the information is mediated in some way. Again, this must be true. People do not directly observe what policymakers do in most domains. They also don't have copies of the federal budget (or other policies) on their coffee tables. Information about what policymakers do is thus conveyed to the public by some other means. But how exactly? Answering this question obviously is important unto itself, as it also helps us understand how the public learns about policy, which is critical to making representative democracies work.

The most likely suspect is mass media. Although its structure and form has changed substantially over time, the public still relies primarily on large-scale media organizations

[4] The equation can apply across space and time, though the latter is more common and also the focus of the analysis in this paper.

for its information about politics.[5] That said, the prospect that mass media provide the information that drives the thermostatic model may seem at odds with common critiques of media content. There are considerable bodies of work detailing a range of biases in media content,[6] and a good deal of work lamenting a lack of policy content specifically (e.g., Lawrence, 2000). There is a considerable body of work, in short, suggesting that media content offers a barely perceptible and systematically biased view of public policy.

Just as we argue that citizens may not need much information to form general opinions, we argue that media can be inefficient and biased in many different ways but still provide the basic information citizens need to assess the direction of policy change, at least in very salient policy areas that attract a lot of public, political and media attention. Skeptics abound, to be sure. But there is already work suggesting that people can and do learn about policy when there is sufficient media coverage (see Barabas and Jerit 2009 for both a review of the field, and a new study); and existing work also identifies some areas in which the public actually has relatively high policy-specific knowledge (e.g., Delli Carpini and Keeter 1996).

How can we best explore this possibility? This paper offers a first attempt, based on the following expectations. If mass media actually are driving thermostatic public responsiveness, we should be able to identify the following patterns:

1. Mass media content will contain some sufficient number of cues about the direction of policy change.
2. Mass media cues about policy change will reflect – to some limited degree, at least – what actually happens to policy.
3. Citizens will be able to identify these cues about policy change.
4. Citizens' policy preferences will respond (thermostatically) to the cues about policy they receive through mass media.

The sections that follow examine the first three of these patterns in turn. We return to the fourth in a concluding section.

Are There Policy Cues in Mass Media?

We begin with an exploration of the number of policy cues available in mass media. We focus for the time being on news content in the *New York Times* and *Washington Post*, from 1980 to the present. Media data are drawn from Factiva full-text indices. For now, we focus on two policy domains: defense and welfare. We select these because they are highly salient domains in which there is ample evidence of thermostatic public responsiveness (e.g., Wlezien 1995, 2004; Soroka and Wlezien 2010).

We discuss the extraction of relevant media articles in more detail in the Appendix. Note here that our aim was to extract policy-relevant articles in the two newspapers, relying primarily on subject codes in the Factiva database, checked by human coders. Our search

---

[5] See, for instance, http://www.people-press.org/2012/02/07/cable-leads-the-pack-as-campaign-news-source/. Of course, citizens may also rely on "new" media for policy information; and these too could facilitate thermostatic responsiveness, particularly if people's various new media sources also track policy over time.

[6] For instance, consider work on sensationalism and/or negativity in news content, e.g., Altheide 1997; Davie and Lee 1995; Lichter and Noyes 1995; Meyrowitz 1985; Patterson 1994; Sabato 1991; Soroka 2014.

produced a database of roughly 70,000 stories, which we analyze using two packages designed for automated content analysis in R. One is relatively common, the tm package. The other, Lexicoder, was released several years ago as a Java-based utility for basic large-scale content analyses (see, e.g., Young and Soroka 2012; Soroka 2006; 2012). We rely here on the newest implementation (3.0), still in Java, but built to be run from the command line and thus also readily accessible within R.[7]

We begin with a simple text-cleaning function in Lexicoder that removes punctuation and changes all words to lower case. Subsequent analyses rely on both word-count and dictionary-count functions. We begin with a simple dictionary search for terms related to spending. (The dictionary is included in the Appendix.) The first rows of Table 1 show the number of articles that include and do not include at least one of these spending keywords. The defense search clearly casts a wider net than does the welfare search. Even so, we have roughly 20,000 defense articles that include one spending keyword, and 11,000 similar articles in welfare. These form the body of data on which we will focus our analyses below.

Table 1. Frequency of Spending Mentions

	Defense	Welfare
Articles with no spending mention	31534	5216
Articles with at least one spending mention	21271	11388
Number of sentences mentioning spending	*68873*	*55003*

The final row of Table 1 shows a different quantity: the number of sentences, rather than articles, mentioning spending. These sentences are identified using the Lexicoder *kwic* (keyword in context) function, which both identifies and then extracts (into a separate database) all instances in which a given set of keywords is used. Note that the *kwic* function captures not just the keyword, but the surrounding *N* words, up to the beginning or end of the sentence. We set a wide, 30-word window here, and in so doing capture the entire sentence in which a set of spending-related keywords are used. This results in a separate database of 123,876 such instances.

Note that our data reflect a potentially large understatement of the frequency of policy information in media content. There will be many articles that deal with policy but do not mention spending explicitly, for instance; and (relatedly) there are many aspects of policy that are not about spending. That said, our corpus focuses on a subset of policy-relevant information that passes a relatively high bar information-wise. For those who suspect that media do not provide any information about the direction of spending, the size alone of the *kwic* dataset should be striking. Even ignoring decades of missing data, our corpus amounts to just under 10 spending-related sentences on these two topics every day for the past 35 years.

It thus appears that – in these domains at least – there are a good number of cues related to the direction of policy. Do those cues actually reflect policy change? We explore this possibility by producing a measure of the "media policy signal" – the direction of policy, as suggested by media content. We begin by identifying all instances in which a spending

[7] The new version of Lexicoder was publicly released in late-August, 2015, and is readily available to academic researchers at www.lexicoder.com.)

keyword co-occurs with a direction keyword – for instance, spend* occurring alongside "more" or "less" offers very clear information about the direction of fiscal policy. (Again, dictionaries are included in the Appendix.) Focusing on co-occurrences of spending and direction keywords leaves us with roughly 63,000 sentences in our database that suggest (albeit ignoring varying degrees of relevance) the direction of fiscal policy.[8]

Converting these dictionary counts into an accurate media policy signal is not straightforward, but we rely on what is probably the simplest route here: we use co-occurring spending and directional keywords to attribute one of two codes to every retrieval, (a) increasing spending (+1), or (b) decreasing spending (-1); and then calculate the sum of all mentions, aggregated by fiscal year, over the 35-year time period.[9]

Is there *any* correspondence? Are there any hints in these data that media content captures over-time trends in policy? Figure 1 offers some preliminary tests, relying on a comparison of the media policy signal and actual budgetary policy, drawn from the Policy Agendas' database of appropriations (available at policyagendas.org).[10] We distinguish spending on each of welfare and defense using the definitions in Wlezien (2004) and Soroka and Wlezien (2010); in each case, we rely on constant dollars, i.e., adjusting for inflation.

Figure 1 illustrates these data in two different ways. The left panels show a simple scatterplot comparing the direction of spending change to the media signal, where the x-axis shows the upward or downward change in appropriations while the y-axis shows upward or downward change suggested by the aggregation of media content. Each dot shows one fiscal year, and the dotted lines show the zero-point for both axes. To the extent that media content points in the same direction as actual spending, we expect dots to appear in either the bottom left quadrant (where spending is decreasing, and media content suggests a decrease), or in the upper right quadrant (where spending is increasing, and media content suggests an increase). Dots in other quadrants signal years in which the media signal is in conflict with the direction of spending. The right panels then show over-time trends in both policy change and the media signal, where the black line shows changes in defense appropriations, in billions of constant dollars, from 1980 to the present. The red line shows our basic media "policy signal" over the same time period, with missing data in quarters for which Factiva stories are missing subject codes.

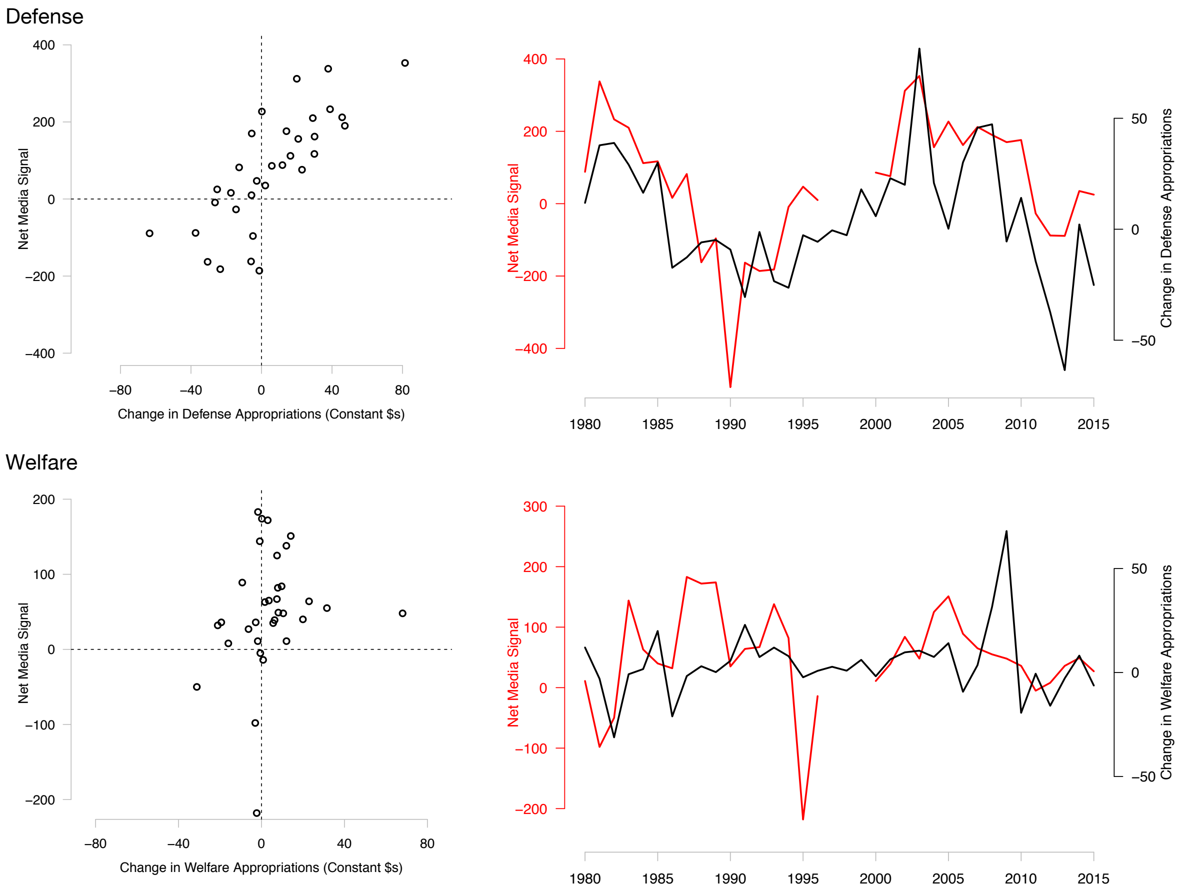
---

[8] More specifically, of the 68,873 sentences on defense spending, 35,207 also include direction keywords; of the 55,003 sentences on welfare spending, 28,350 including spending keywords.

[9] This approach is far from perfect. Perhaps most importantly, it takes the frequency of co-occurrences as an indication of magnitude. A month in which there are many co-occurrences in a positive direction will show a strongly positive signal, while a month in which there are only a few co-occurrences in a positive direction will show a weakly positive signal, for instance. There is however no reason to expect that the magnitude of spending change will be systematically matched to the number of mentions of upward or downward movement. A highly salient but fiscally small change in policy is a problem for our current measure, then. We thus do not expect a perfect correspondence between our current media signal and fiscal policy. But we regard this work as an initial test only.

[10] These data have the advantage of using functional definitions that are more temporally consistent than the standard OMB Historical Tables. OMB classification actually changes over time, though probably less often than in other countries, where reliable spending data can be even more difficult to identify (Soroka, Wlezien and McLean 2006).

Figure 1. Media Cues and Budgetary Policy



Results for defense suggest a rather strong relationship between media content and actual policy. There are very few off-diagonal dots in the first panel; and the correlation between the two series is 0.68.[11] It is clear that media cues are pretty consistent with the direction of spending change, both in terms of direction and magnitude. The case for welfare is weaker – the correlation between the series is just 0.19. Even so, the number of years in which the media signal points in the opposite direction as spending change (see the bottom right panel of Figure 1) is, again relatively limited: there are seven or eight years in which the direction of the media signal is at odds with the direction of spending, and this is actually very close to what we see in defense. Viewed in this way, the media signal in the welfare domain is relatively accurate – in terms of the direction of change at least, if not also the magnitude.

What accounts for the differences between defense and welfare? There are many possible reasons, and we cannot easily adjudicate between them here. One possibility is that our content analysis simply does not capture what it needs to in welfare. It could be that our spending and/or direction keywords are not ideal in this domain. It could be that media content on welfare is less focused on "spending," and more focused on increases or

[11] Results do not differ dramatically if we look at spending in current versus constant dollars, or percentage versus absolute change. See Soroka and Wlezien (2015).

decreases in other types of policy action. It could also be that media content provides an amalgam of federal and state-level policy in this domain, and we do not adequately separate out the federal material here. Each of these possibilities is testable, relying on a larger body of human-coded data. This is one objective of future work. For the time being, we take these results as strong indications that (1) mass media content contain a good number of informational cues about the direction of policy change, and (2) on average, those cues reflect what actually happens to policy (although this may be more true in defense than in welfare). We accordingly proceed to the next step.

Can Citizens Identify Mediated Policy Cues?

Our first test of whether citizens can identify spending cues in media content comes in the form of a simple coding exercise, fielded in MTurk. Note that we do not require that MTurkers are broadly representative in this instance – we regard them only as non-expert coders, and examine their ability to identify policy cues in media content.

We select a sample of media stories as follows. We begin with a random draw of 120 stories on welfare and 120 stories on defense (40 from each tercile, based on the total number of spending keywords), from our 35-year database, but excluding stories with no spending keywords. This last constraint is intended to ensure that we have stories with a least a minimal amount of policy content. A single expert coder read through all stories to ensure that they were relevant, i.e., that they dealt with either welfare or defense. We then took a random sample of 40 defense and 40 welfare articles from all articles deemed relevant. In the process of data cleaning we dropped three articles that where compilations of a variety of news articles on different topics. This left 77 stories which we then inserted into our online coding exercise, built in Qualtrics.

The instructions to MTurkers were straightforward: "We are interested in understanding information in news content. On the next page you will be presented with a newspaper article about defense or social welfare policy from the past 35 years. Please take your time reading the article. When you are done, click the 'next' button and we will ask a few questions about the article." Respondents were then presented with one article randomly drawn from the set of 77. Following the article, they answered several questions, including the following:[12]

> *Policy Change*: Now, thinking about the article you have just read: Did this news article offer any information about changes in government spending on either defense or welfare (yes, no, I am sure);
>
> *Direction*: Did the article indicate whether spending was increasing or decreasing (increasing, decreasing, unsure);
>
> *Magnitude*: How would you describe the size of this spending change? On a scale from 1 to 5, would you say that the spending change is… very small to very large;

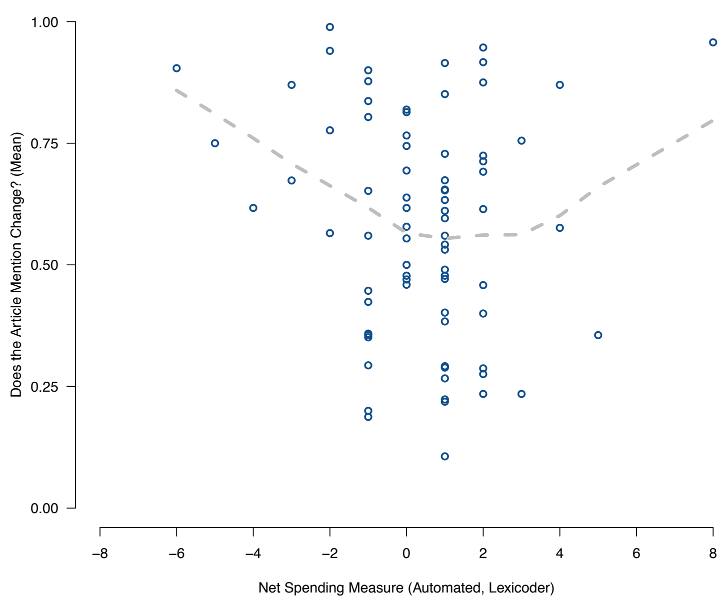
Note that the *Direction* and *Magnitude* questions were conditional on previous answers. And note that for this first article, respondents received no information about what they should be looking for in the article. We thus see this first coding attempt as being highly realistic, in the sense that respondents are reading for no particular reason. They then code a second article, preceded by "Now we would like you to read another article. After reading it you will be asked the same set of questions we asked for the previous article."

[12] The titles of questions were not part of the survey; they are included for the sake of exposition here.

Now our participants are trained, minimally at least. The extent to which there are differences between the first and second articles is thus of some interest. This procedure thus provides us with responses on two articles from roughly 1800 unique US-based MTurkers, that is over 3600 assessments of our variables across 77 articles.

We begin with some aggregate results, based on mean ratings for each of the 77 articles. Note first that the tendency to report *Policy Change* increases with the number of spending mentions in an article. This seems relatively obvious, perhaps, but we consider it an important test of whether the cues we capture using Lexicoder are of relevance when humans read news articles. They clearly are. Figure 2 shows the net spending measure, derived using Lexicoder, along the x-axis. The y-axis shows the proportion of MTurk responses to the Policy Change question (where 0: no change, 0.5: unsure, 1: change). There is a clear u-shaped relationship between the two. Articles which feature a greater number of either upward or downward change phrases are more regularly identified as indicating change.[13]

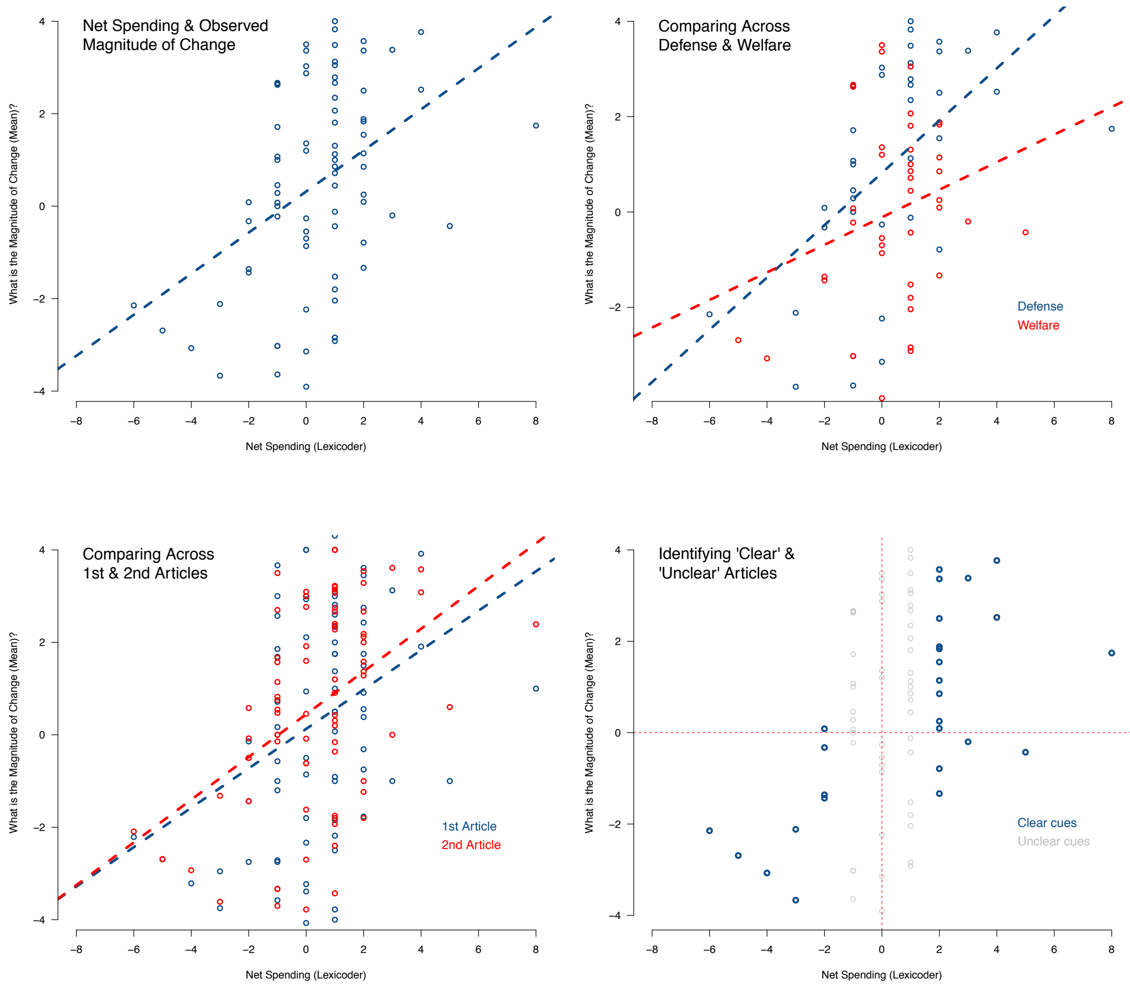
Figure 2. Media Cues and Aggregated MTurk Views on the Presence of Policy Change



Can readers identify not just that there is change, but the direction of that change? Here we focus on a combined analysis of *Policy Change*, *Direction* and *Magnitude*. The top left panel of Figure 3 plots the net spending measure from Lexicoder (on the y-axis) against the mean for *Magnitude*, which effectively incorporates earlier questions on *Policy Change* and *Direction*.

[13] Note that variation on the y-axis makes clear that there are some articles for which most respondents identify policy change, and others for which very few respondents identify change. This has implications for the reliability and variability in the direction/magnitude measures used below. Note, for instance, that we find significant results below in spite of our inclusion of articles for which direction/magnitude is assessed by only a few respondents (while the rest see no policy change).

Figure 3. Media Cues and Aggregated MTurk Views on the Magnitude of Policy Change



The regression slope in the top left panel of Figure 3 (B = 0.44, p<.001) makes clear the strong relationship between Lexicoder-based results and aggregated MTurk responses. Specifically, increases in the net spending measure are associated with aggregate perceptions that spending is increasing. Similarly, a model that captures the impact of upward and downward cues independently suggests that both are associated with perceptions that spending is increasing or decreasing. (For upward cues, the B = 0.39, p<.001; for downward ones, the B = -0.47, p<.001. See Appendix Table 2.)

There is some difference across the two issue areas. This can be seen in the top right panel of Figure 3, which separates out defense in blue from welfare in red, and plots the corresponding OLS regression lines. That the connection between MTurk ratings and automated coding is weaker for welfare between the slopes is insignificant based on a Wald test (B for welfare is 0.29; for defense, 0.55; p = .34). Nonetheless, we suspect that the connection between welfare stories and welfare policy may be stronger than we identify above; that is, we possibly can improve the welfare policy dictionary. Even so, we take these results as strong evidence that untrained readers can readily identify policy change in news content.

Recall that our coding exercise includes one "untrained" reading of a first story, followed by a "trained" reading of a second. The bottom left panel of Figure 3 compares results across these first and second readings of each article; and the difference in the regression lines for first (blue) versus second (red) readings indicates the extent to which training improves readers' ability to identify spending direction. Note that the regression slopes are very similar, which we consider in more detail below. For now, let us say that we see this as evidence of the ease with which readers are able to identify policy direction.

That said, the task is more difficult for articles that do not contain a clear policy signal. The bottom right panel of Figure 3 shows the same data, focusing on the clear instances of policy change, in this case, the articles with net spending measures either below -2 or above +2. We indicate the zero-points on both axes to make clear that there are very few of these "clear" cases in which the mean MTurk rating is not in the right direction. The grayed-out "unclear" cases produce much more varied ratings, however. Our analyses suggest that these 50 cases are evenly split across what we could characterize as "limited" cases, which include just one single mention of spending change, and "mixed" cases, which include a good number of both upward *and* downward mentions. (See Appendix Figure 1 for a graphic showing the number of upward and downward cues in all "unclear" cases.) It makes good sense that our respondents do not converge on a clear answer for these articles, of course. We accordingly take this as additional evidence that untrained readers are readily able to identify spending cues in media content.

We can explore these same data at the individual level as well, of course, and do so using a stacked dataset for which each respondent-article combination is a case. To be clear: we now turn to analyses of a dataset including two cases for each respondent (and where coefficient standard errors are clustered accordingly). Note that because direction and magnitude variables are conditional on respondents identifying spending change, these models rely on roughly 2,000 observations of the 77 articles.[14]

Models 1 and 2 regress the direction and magnitude of change variable (-2 to +2) on the Lexicoder-derived net spending measure, and then upward and downward spending cues respectively. Results in Model 1 make clear the strong relationship between the automated cues and individuals' perceptions of policy change; Model 2, which shows almost identical coefficients for upward and downward change, makes clear that respondents are equally able to identify upward and downward cues. (We regard this as much a test of our direction dictionaries as it is a test of asymmetry in cues amongst readers.) Models 3 and 4 are a straightforward replication of Model 2, though in these cases we use responses from first articles only, and then second articles only. There are no large differences in these models – upward and downward cues appear to be equally effective even before our respondents receive instructions and practice.

This is not to say that there aren't slight improvements once our readers are "trained" to know what to look for. Consider the slightly larger *R*-squared in Model 4 versus 3, for instance – an indication that spending cues explain slightly more variance in the second

[14] This is rather than the entire sample of 3,600, which includes 1,600 instances in which respondents do not identify policy change.

round of coding. Also consider this somewhat better illustration: If we take "correct" answers as the ability to identify the Lexicoder-identified direction of spending change in articles where that direction is "clear,"[15] 40% of the 659 first readings were correct.[16] The percent of correct responses in the second round is 46%. A t-test of means indicates that this difference is significant at p = .02.[17]

Table 2. The Direction and Magnitude of Change and Spending Cues

	Model 1 *all articles*	Model 2 *all articles*	Model 3 *1st articles*	Model 4 *2nd articles*	Model 5 *all articles*
Net spending	.516***				
	(.001)				
Upward spending mentions		.532***	.536***	.527***	.259***
		(.042)	(.062)	(.058)	(.061)
Downward spending mentions		-.508***	-.478***	-.542***	-.859***
		(.032)	(.046)	(.044)	(.061)
Upward * Downward					.110***
					(.016)
Constant	-.157**	.102	-.090	.122	.683***
	(.072)	(.116)	(.162)	(.167)	(.151)
N	1,966	1,966	999	967	1,966
R2	.122	.122	.116	.130	.139

Cells contain regression coefficients and robust standard errors from an OLS panel estimation. * p < .10; ** p < .05; *** p < .01.

Returning to Table 2, Model 5 recombines respondents' first and second codings but adds an interaction between upward and downward mentions. This allows us to see whether the impact of cues in one direction depend on the existence of cues in the other direction. The magnitude of the coefficients for upward and downward changes thus shift here, as they should; with the inclusion of the interaction term, the direct effect of upward (downward) mentions captures the impact of an upward (downward) article when the number of downward (upward) mentions is 0. The significant interaction makes clear that the impact of both upward and downward cues is moderated by the presence of cues in the opposite direction. So articles with multiple cues, in multiple directions, make identifying the direction of change more difficult. We view this as another illustration of the complications of multiple cues (seen earlier in the grayed-out dots in Figure 3). We also see it as further evidence that respondents are sensitive to the kinds of spending cues that we are capturing in automated analyses.

Discussion

This project provides a preliminary attempt to unpack the individual-level dynamics that help produce thermostatic responsiveness to policy. Taken together, the evidence from

[15] Note that although we rely on automated content analysis to test the clarity of articles for the time being, we are in the process of a second confirmation using expert coders.
[16] To be clear: there are 1868 MTurk respondents, but only 659 of them received a 'clear' article in their first reading.
[17] Regression models using additional individual-level variables (not shown here) suggest that this upward shift in correctness is not restricted to those with higher education or political interest – the improved ability to identify direction occurs across the board. We take this as evidence that spending cues can be informative to those with quite different backgrounds and interests, though we readily acknowledge that our MTurkers are not at all representative in this regard.

the automated content analysis and the MTurk coding study suggests that spending cues exist, and that citizens can extract these cues from news content (even if they are not consciously looking for them). This suggests that the basic information needed for thermostatic responsiveness to function is readily available. We regard this as a first, and significant, step towards better understanding the role of media content in public responsiveness, and in representative democracy more broadly.

There are several important caveats. First, note that when we look at "correct" responses in our individual-level MTurk data, even in comparatively clear media stories, the percent of "correct" responses is below 50%. It clearly is not the case that everyone easily assimilates information on spending from news content. Aggregation is thus an important part of the process that gets us to thermostatic responsiveness; though the underlying trend evidently is driven by a large (and diverse) minority, receiving relatively reliable mediated cues about policy change.

Second, we have so far not dealt with the fourth pattern mentioned above – when and how thermostatic responsiveness to policy operates at the individual level. The task at hand, then, is to design an experiment that is able to identify those who are, and are not, responding thermostatically to news content. We are considering several possibilities.[18] One involves fielding an exploratory study to gauge opinion distributions on a number of policy domains. These results would then be used to develop experimental manipulations for policy domains with differing degrees of preference heterogeneity. One set of manipulations might then involve designing news articles with varying numbers of spending cues, which will allow us to examine whether individuals shift their preferences thermostatically in response to new information about spending direction in a single-shot experiment. Importantly, this design would allow us to unpack whether spending cues can affect perceptions of past policy change and whether these perceptions, in turn, affect preferences for future spending.

That said, observational data already make clear that thermostatic responsiveness exists, at least across salient policy domains in developed democracies. And we have argued elsewhere that there is no obvious dynamic that can account for this finding except that publics react thermostatically to policy change. It follows that an experimental demonstration of thermostatic responsiveness is less critical to our endeavor that the material presented above. That does not make the prospect of observing individual-level thermostatic responsiveness any less attractive, of course; though note that we see the advantage of being able to do so not much as proof of thermostatic responsiveness, but rather as an opportunity to probe further who is responding, and the kinds of cues to which they are most responsive. In the meantime, we take the evidence above as powerful confirmation of the first three of the four patterns which we believe should be evident if media play a role in facilitating thermostatic responsiveness to policy. There *are* cues, they *are* often indicative of real policy change, and untrained individuals *are* able to identify those cues in real media stories.

[18] We have in fact fielded one study on MTurk as an initial, preliminary test, which did not reveal thermostatic responsiveness. That experiment relied on basic spending queues on a singular issue -- spending on maintaining and upgrading nuclear weapons. More details are available upon request

Appendix

Additional Tables

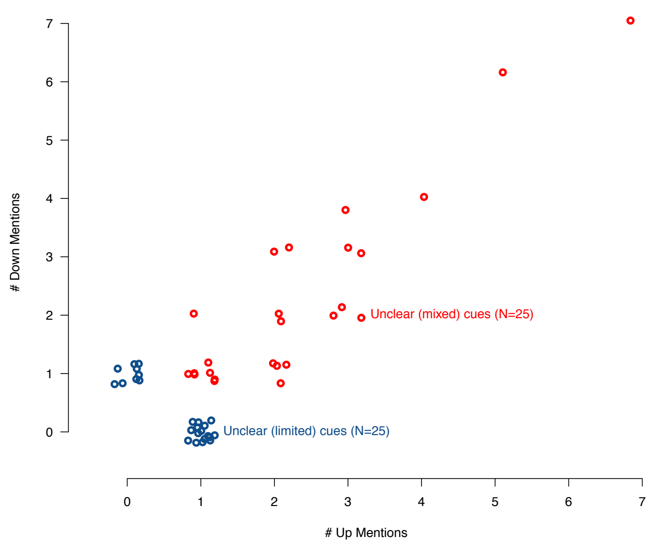
Appendix Table 2: Aggregate-Level Models of Direction and Magnitude

	Model 1	Model 2	Model 3	Model 4
Net spending	.444***		.548***	.289*
	(.111)		(.138)	(.171)
Upward spending mentions		.391***		
		(.143)		
Downward spending mentions		-.471***		
		(.120)		
Constant	.317	.475	.820**	-.111
	(.231)	(.352)	(.326)	(.309)
N	77	77	37	40
R2	.177	.181	.310	.070

* p < .10; ** p < .05; *** p < .01.

Additional Figures

Appendix Figure 1. Upward and Downward Cues in Unclear Articles



Media Content

We rely on Factiva subject keywords to identify articles: GWELF, and GDEF. These topics were selected based on preliminary searches; they tend to be policy-focused (rather than just focused on wars in the case of GDEF, for instance).

Note that the use of Factiva keywords by no means perfectly captures only relevant articles. We invariably miss some relevant articles; and our analyses identify a good volume of irrelevant material captured in our searches as well. The keyword search is nevertheless more efficient than a full-text keyword search. (Consider what we might get when searching for articles using the word "welfare," for instance – a set of articles dealing not just with social welfare, but also child welfare, animal welfare, etc.) Of course, Factiva's assignment of topics is most likely a function of their own dictionary-based word search, but our assumption is that their search is more developed than ours would be. We suspect that our use of their keywords means that we err on the side of Type I rather than Type II errors. That said, we have to sort through irrelevant material as well, and one aim of the analyses that follow is to improve our ability to identify relevant material.

A more critical issue with subject codes in Factiva is that they have not been applied to all years – there are in particular years in the late 1990s when no subject codes exist for either of the newspapers that we focus on, and we thus have missing data in these years. For this preliminary work, we simply work around these missing data. That said, we do not lack for media content. Our database includes roughly 16,600 welfare articles and 52,800 defense articles. The analyses that follow are thus based on nearly 70,000 articles in total. The basic breakdown of articles is shown in Appendix Table 1, which reports the total number of articles downloaded by topic, newspaper and decade.

Appendix Table 1. Sample of News Articles, by Topic, Decade and Newspaper

	Defense		Welfare	
	NYT	*WPost*	*NYT*	*WPost*
1980s	15312	0	7415	0
1990s	3382	14	3885	0
2000s	10667	10346	2237	2063
2010s	5572	5807	580	424

The table makes clear the gaps that result from the intermittent non-existence of subject codes in the Factiva database. There are many more defense articles than welfare articles (at least based on these Factiva keywords); and we have no *Washington Post* articles before 2000. (There are also articles missing from Table 1 due to miscodes for date and/or news source, but these cases are relatively rare.) For the time being, however, we work with these data, with the belief that they will give us a sufficient sense for whether policy cues are available, and retrievable, in media content.

Content-Analytic Dictionaries

Dictionaries used above are implemented in Lexicoder, and constructed from our own reading of *kwic* retrievals, augmented by thesaurus searches. They are as follows:

```
SPENDING: allocate*, appropriation*, budget*, cost*, earmark*, expend*, fund*, grant*,
outlay*, resourc*, spend*

UP: accelerat*, accession, accru*, accumulat*, arise*, arose, ascen*, augment*, boom*,
boost, climb*, elevat*, exceed*, expand*, expansion, extend*, gain*, grow*, heighten*,
higher, increas*, increment*, jump*, leap*, more, multiply*, peak*, rais*, resurg*,
rise*, rising, rose, skyrocket*, soar*, surg*, escalat*, up, upraise, upsurge, upward

DOWN: collaps*, contract*, cut*, decay*, declin*, decompos*, decreas*, deflat*,
deplet*, depreciat*, descend*, diminish*, dip*, drop*, dwindl*, fall*, fell, fewer,
less, lose, losing, loss, lost, lower*, minimiz*, plung*, reced*, reduc*, sank, sink*,
scarcit*, shrank, shrink*, shrivel*, shrunk, slash*, slid*, slip*, slow*, slump*,
sunk*, toppl*, trim*, tumbl*, wane, waning, wither*
```

Note that not all direction words will be used in relation to spending; just as not all spending mentions will co-occur with direction keywords. We start by casting a relatively wide net in both instances, however; and expect that the concurrence of these two dictionaries will identify sentences that are related to policy direction in each domain.

References

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

Altheide, David L. 1997. “The News Media, the Problem Frame, and the Production of Fear.” *Sociological Quarterly* 38(4): 647–68.

Ball-Rokeach, Sandra J. and M.L. DeFleur. 1976. “A dependency model of mass-media effects.” *Communication Research* 3 (1): 3–21.

Barabas, Jason and Jennifer Jerit. 2009. “Estimating the Causal Effects of Media Coverage on Policy Specific Knowledge.” *American Journal of Political Science* 53(1): 73-89.

Bennett, Stephen. 1988. “’Know-Nothings’ Revisited: The Meaning of Political Ignorance Today. *Social Science Quarterly* 69:476-490.

Berelson, Bernard R., Paul F. Lazarsfeld, and William N. McPhee. 1954. *Voting*. Chicago: University of Chicago Press.

Converse, Philip E. 1970. “Attitudes and Non-Attitudes: Continuation of a Dialogue. In Edward R. Tufte (ed.), *The Quantitative Analysis of Social Problems*. Reading, Mass: Addison-Wesley.

-----. 1964. “The Nature of Belief Systems in Mass Publics.” In David Apter (ed.), *Ideology and Discontent*. New York: Free Press.

Dahl, Robert. 1971. *Polyarchy: Participation and Opposition*. New Haven: Yale University Press.

Davie, William R., and Jung Sook Lee. 1995. “Sex, Violence, and Consonance/Differentiation: An Analysis of Local TV News Values.” *Journalism and Mass Communication Quarterly* 72(1): 128–38.

Delli Carpini, Michael and Scott Keeter. 1996. *What American Know about Politics and Why it Matters*. New Haven: Yale University Press.

Druckman, James N. and Lawrence R. Jacobs. 2015. *Who Governs? Presidents, Public Opinion and Manipulation*. Chicago: University of Chicago Press.

Druckman, James N. and Thomas J. Leeper. 2012. “Learning more from political communication experiments: Pretreatment and its effects.” *American Journal of Political Science* 56(4): 875-896.

Enns, Peter and Christopher Wlezien (eds.). 2011. *Who Gets Represented?* New York: Russell Sage Foundation.

Eichenberg, Richard, and Richard Stoll. 2003. “Representing Defence: Democratic Control of the Defence Budget in the United States and Western Europe.” *Journal of Conflict Resolution* 47: 399-423.

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

Franklin, Mark and Christopher Wlezien. 1997. “The Responsive Public: Issue Salience, Policy Change, and Preferences for European Unification.” *Journal of Theoretical Politics* 9:347-63.

Geer, John Gray. 1996. *From Tea Leaves to Opinion Polls: A Theory of Democratic Leadership*. New York: Columbia University Press.

Jennings, Will. 2009. “The Public Thermostat, Political Responsiveness and Error Correction: Border Control and Asylum in Britain, 1994-2007.” *British Journal of Political Science* 39:847-870.

Lawrence, Regina G. "Game-Framing the Issues: Tracking the Strategy Frame in Public Policy News." *Political Communication* 17(2): 93-114.

Lichter, S. Robert and Richard Noyes. 1995. *Good Intentions Make Bad News*. Lanham MD: Rowman and Littlefield.

Meyrowitz, Joshua. 1985. *No Sense of Place: The Impact of Electronic Media on Social Behavior*. New York: Oxford University Press.

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

———. 1983. "Effects of public opinion on policy." *American Political Science Review* 77: 175–90.

Popkin, Samuel and Michael Dimock. 1999. Political Knowledge and Citizen Competence. In Stephen Elkin and Karol Salton (eds), *Citizen Competence and Democratic Institutions*. University Park: Pennsylvania State University.

Patterson, Thomas. 1994. *Out of Order*. New York: Vintage.

Sabato, Larry, 1991. *Feeding Frenzy*. New York: Free Press.

Soroka, Stuart. 2014. *Negativity in Democratic Politics: Causes and Consequences*. Cambridge Studies in Public Opinion and Political Psychology, Cambridge University Press.

Soroka, Stuart N. and Christopher Wlezien. 2015. "On the Sources of Public Responsiveness to Policy." Paper presented at the Annual Meeting of the American Political Science Association, San Francisco CA. Available at http://degreesofdemocracy.net.

———. 2010. *Degrees of Democracy: Politics, Public Opinion and Policy.* Cambridge University Press.

Stuart Soroka, Christopher Wlezien and Iain MacLean. 2006. "Public Expenditure in the UK: How Measures Matter," *Journal of the Royal Statistical Society (Series A)*, volume 169.

Turcotte, Jason. 2014. "Debates and the Disincentives of Democracy: How News Norms and Values Influence Moderator Agendas." *Journalism Practice 8(6): 772-788*

Wlezien, Christopher. 2004. "Patterns of Representation: Dynamics of Public Preferences and Policy." *Journal of Politics* 66:1-24.

———. 1996. "Dynamics of Representation: The Case of US Spending on Defense." *British Journal of Political Science* 26:81-103.

———. 1995. "The Public as Thermostat: Dynamics of Preferences for Spending." *American Journal of Political Science* 39:981-1000.

Wlezien, Christopher and Stuart Soroka. 2013. "Measuring and Modeling Public Preferences for Policy: The Case of Spending on Welfare in the US."

———. 2012. "Political Institutions and the Opinion-Policy Link." *West European Politics* 35(6): 1407-1432.

———. 2003. "Measures and Models of Budgetary Policy." *Policy Studies Journal* 31(2): 273-286.