

# Perspectives Research Problem Set 2

*Julian McClellan*

*Due 4/19/17*

## Measuring Representational Style in the House: The Tea Party, Obama and Legislators' Changing Expressed Priorities

### (a) State the Research question

This paper seeks to answer a *general question*:

“How do House Representatives define the type of representation they provide to constituents?”

With this specific research question in mind, the paper can answer *more specifically*:

“In regards to the type of representation they provide to constituents, how did House Republicans respond to the drastic change in institutional and electoral context after the 2008 election?”

### (b) What data did the paper use?

From Grimmer (2014) himself:

I use a collection of nearly 170,000 House press releases: every press release, from each House office, from 2005 to 2010.

### (c) What theory was referenced in order to interpret the data?

Generally, Grimmer asserts the importance of Congressional communication:

Communication is a central component of representation (Mansbridge, 2003; Disch, 2012). Legislators invest time and resources in crafting speeches in Congress, composing press releases to send to newspapers, and in distributing messages directly to their constituents (Yiannakis, 1982; Quinn et al., 2010; Lipinski, 2004; Grimmer, 2013).

Grimmer references two of his own works to justify his methods. One reference is to the general method he uses, a *text as data* method (Grimmer and Stewart 2013), and another reference is the specific method he uses on the press releases, a Bayesian hierarchical model for political texts that he had previously used on Senate press releases to measure legislators' expressed priorities: the attention they allocate to topics and issues when communicating to constituents (Grimmer 2010). Grimmer also generally cites the use of topic models as increasingly common tools for studying large bodies of text.

Additionally, Grimmer also justifies the source of the texts he has chosen for analysis:

There is increasing evidence that press releases are a reliable and useful source for capturing how legislators communicate with their constituents. Grimmer (2013) shows that press releases contain politically relevant content not found in floor speeches and that press releases have a direct effect over the content of newspaper stories and constituent evaluations.

### (d) Was the paper a descriptive study, an identification exercise, a

numerical solution to system of equations study, or some combination of the three?

On the one hand, Grimmer's paper has aspects of a descriptive study. Returning to the more general research question, the paper seeks to *describe* how House Representatives represent themselves to constituents.

It utilizes hierarchical modelling, an unsupervised method, on the 2005-2010 House press releases as a means of reducing the dimensionality of more succinctly describing the information contained in the press releases. Grimmer labeled topics by reading a small selection of press releases that were grouped together, and then validated them by checking to see timelines of the topics, and whether the "spikes" in the topics made sense. For example, financial crisis topics spiked around the financial crisis, as expected. Grimmer also checked the output of the model against human coding; for the "coarse" topic "credit claiming", numbers of documents for each legislator that were labeled credit claiming correlated highly (.93) with the numbers attained from human coders.

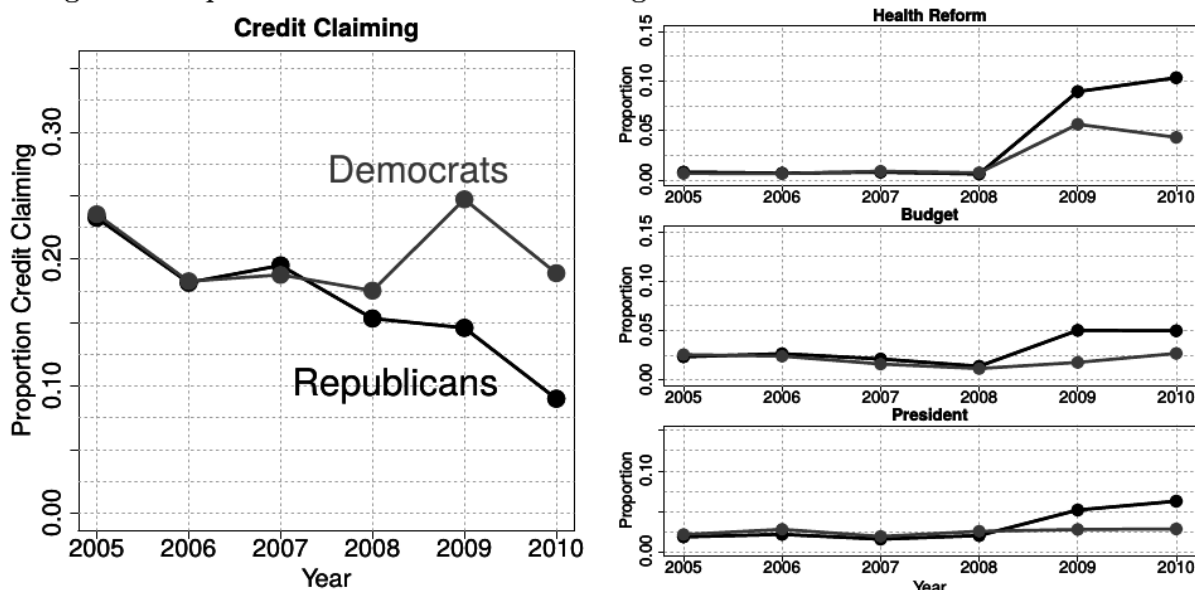
However, with this description in mind, Grimmer then goes on to carry out something more in line with an identification study, seeing how the press releases changed in response to changing electoral and institutional pressures (loss of majorities to Democrats, election of Obama, etc.).

(e) What computational methods did this paper use to answer the research question? What was their result or answer to the question?

As mentioned above, Grimmer utilized a hierarchical topic model with 8 coarse topics (think branches), with 44 total granular topics (think leaves). The topics are then manually labeled, with validation procedures to make sure that these labels are appropriate. With these topics, the paper answers the more general research question.

Then the paper seeks to answer the more specific research question by identifying electoral and institutional pressures that occurred during the span of time of the press releases, and seeing how the distribution of these topics changed over time for political parties. See below:

Figure 3: Republicans Avoid Credit Claiming and Instead Attack Presidential Policies



Republicans avoid claiming credit for spending and instead focus on criticism Obama administration policies

Figure 1:

From here, we see that after the 2008 presidential election, Republicans released fewer credit claiming press releases (while Democrats engage in more), and instead released more press releases concerning health reform, the budget, and the president. Grimmer asserts that the Republican increases in the 3 aforementioned areas are criticisms of presidential policy, citing excerpts from a few press releases as an example.

However, citing himself again, Grimmer notes:

But the response is on the margin and a deviation from a longer run strategy that legislators develop over the course of their career (Grimmer, 2013). And as a result, there remains a strong over time relationship in legislators' expressed priorities.

Thus, the change is less characterized as an “about-face” and more as an adjustment to a changing reality.

**(f) Give two suggestions to the author of the paper of things the authors might do to improve their results or strengthen their evidence for the answer to the question.**

**Suggestion 1:**

As mentioned before Grimmer utilizes 44 granular topics nested under 8 coarse topics for his study. Grimmer justifies his use of 44 granular topics by pointing out that it is a number used in previous studies of Congressional communication, and justifies his use of 8 coarse topics by citing that the number was determined after some initial experiments with a subset of the documents (Grimmer 2013, 10)

Grimmer could have done a better job in justifying the number of granular and coarse topics selected. When Grimmer says previous studies have used 44, he doesn't explicitly cite which studies (or anything at all within the paragraph). Additionally, “initial experiments” is a really vague term. Grimmer's validation of the topics and interpretation of the results has its foundation in the number of topics he chose to a certain degree, and asserting the numbers would have gone a ways in providing more evidence for the answer.

**Suggestion 2:**

One of the main findings of the paper is that Republicans reduced their credit claiming and increased criticism of the Obama administration and Democratic policies. While the graph that has been included shows a clear reduction in credit claiming and increase in health reform, budget, and president related press releases, Grimmer uses evidence of varying concreteness to justify it as criticism.

In regards to health care, Grimmer simply excerpts a few example press releases, for the budget, Grimmer cites another paper, and for the stimulus Grimmer again excerpts from example press releases, but also then uses more “quantitative” comparisons of Republican and Democratic language, contrasting the usage rates of certain words between Democrats and Republicans.

This presentation of evidence seems very uneven. In some cases, Grimmer cites increases he detects in topic usage and cites other papers talking about Republican criticism in regards to that topic to say that criticism was amplified, where in other cases he asserts the evidence of amplified Republican criticism with actual quantitative measures of language usage.

I think the paper would be stronger if Grimmer more evenly utilized the quantitative measure of language difference for *every* topic he claims the Republicans were critical of. It makes a very convincing case for the stimulus debate (which to be fair might cover Budget + President), and I'm sure it would to if it were used for the other topics as well. After all, the paper's central method is a statistical topic model, why not utilize it to back all of the claims it can.

I can imagine it being the case in which the evaluation of language usage actually was not as convincing for the topics which Grimmer does not use them in more explicitly, like the healthcare debate, or the budget specifically, in which case Grimmer ought to explain why he utilized them for the stimulus and not for these topics.

Grimmer, Justin. 2010. "A Bayesian Hierarchical Topic Model for Political Texts: Measuring Expressed Agendas in Senate Press Releases." *Political Analysis*. JSTOR, 1–35.

———. 2013. *Representational Style in Congress: What Legislators Say and Why It Matters*. Cambridge University Press.

———. 2014. "Measuring Representational Style in the House: The Tea Party, Obama and Legislators' Changing Expressed Priorities."

Grimmer, Justin, and Brandon M Stewart. 2013. "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts." *Political Analysis*. JSTOR, 267–97.