

PS1, Part 2

Yilun Dai

**Paper critiqued:**

Grimmer, J. (2016). Measuring Representational Style in the House: The Tea Party, Obama, and Legislators' Changing Expressed Priorities. In R. Alvarez (Ed.), *Computational Social Science: Discovery and Prediction (Analytical Methods for Social Research)*, pp. 225-245). Cambridge: Cambridge University Press. doi:10.1017/CBO9781316257340.010

**1. Research Question:**

This paper studies how legislators adopt their ways of communication with constituents and how they strategically change their approaches in response to changes in political environment. It utilizes computational tools to study representation.

**2. Data used:**

The author adopted a text as data method and used a large collection of Congressional text. Specifically, the author used a collection of nearly 170,000 House press releases: every press release, from each House office, from 2005 to 2010, and studied the language pattern of the text.

**3. Theory referenced:**

The paper has no reference to theory. It uses a statistical model. To be more specific, the author adopts a Topic model with a two-layer hierarchy which nests granular topics into a set of coarse topics. This Topic model is used to analyze speech patterns.

**4. Classification:**

This paper is a combination of a descriptive study and an identification exercise. It configures and uses a new dataset (collection of texts), and highlights the change in speech patterns. It also identifies the constituent priorities that are correlated with speech patterns and real-world events that are concurrent with the shift in press representation.

## 5. **Computational Methods:**

The paper first preprocesses texts to transform contents as numbers. “For each legislator-year  $i$  ( $i = 1, \dots, 2,587$ ), each press release  $j$  ( $j = 1, \dots, N_i$ ) is represented as a  $W = 2,727$  element-long count vector  $y_{ij} = (y_{ij1}, y_{ij2}, \dots, y_{ij2727})$ .<sup>1</sup> Each  $y_{ijw}$  counts the number of times token  $w$  occurs in document  $j$  from legislator  $i$ ”(pp. 8) .

After preprocessing texts, the paper uses a Topic Model with a two-layer hierarchy which nests granular topics into a set of coarse topics to analyze speech patterns and how they change across time, in response to political events.

To validate the model’s accuracy, the paper uses regression to check the correlation between the proportion of press releases on a specific topic and constituents’ priority on that topic.

The paper also uses data visualization to present the changes of speech over the years.

**Results:** The paper validates the accuracy of the Topic model, and demonstrates that legislators change their way of communication in response to constituents’ priority and political environment. The paper finds that after Obama’s election, the Republicans stops credit claiming and turns to opposing health care reform, while the Democrats increases their credit claiming. The paper also shows that computational tools can be useful in studying legislators’ speech patterns.

## 6. **Suggestions:**

- 1) For part 5, pp. 18, I suggest that you add a heat map reflecting the different proportion of press release in farming in different district, so that you could demonstrate that farming district have a higher frequency of press release on agriculture.
- 2) For part 5, I suggest that you add at least one figure with spikes corresponding to real-world events (similar to the ones in part 4, pp. 14) to strengthen your argument. Besides figure 3, a visualization of spikes that indicate real-world events would explain the changes in press release even better.