A Network Model of Political Argument and Opinion Change

Nick Beauchamp, New York University   nick.beauchamp@nyu.edu   nickbeauchamp.com

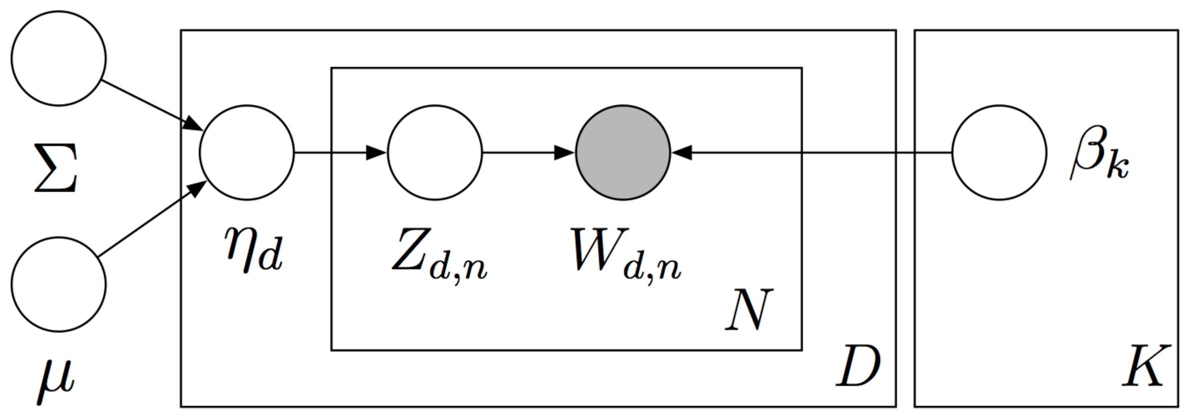
Introduction

Political opinions form and change under the ceaseless barrage of speech, conversation and argument. Online forums such as Dailykos.com record the arguments between hundreds of thousands of participants over many years. A correlated topic model of these exchanges allows us to represent the patterns of debate, predict strategic speech made in response to earlier statements, and ultimately understand how complex opinions change over time.

Correlated topic model

D documents. N unique words. K topics.
Bag of words: $W_{d,n}$ = % of doc d that is word n.
Each document has distribution η_d over K topics.

	Text	*Topics*
Documents:	$W = D \times N$	$\eta = D \times K$
Topic of word:	$Z = D \times N$	$\mu = K$
Word dist per topic:	$\beta = K \times N$	$\Sigma = K \times K$



$$\eta_d \sim \mathcal{N}(\mu, \Sigma) \quad (1)$$

$$Z_{d,n} \sim \mathcal{M}(f(\eta_d)) \quad (2)$$

$$W_{d,n} \sim \mathcal{M}(\beta_{Z_{d,n}}) \quad (3)$$

$$f(\eta_d) \equiv \theta_d = \frac{\exp(\eta_d)}{\Sigma_k \exp(\eta_{d,k})} \quad (4)$$

Coordinate ascent

Multinomials $\mathcal{M}$ allow direct, fast maximization:

$$Z_{d,n} \to \phi_{d,n,k}$$

$$\beta_{k,n} = \frac{\Sigma_d W_{d,n} \phi_{d,n,k}}{\Sigma_{d,n} W_{d,n} \phi_{d,n,k}} \quad (5)$$

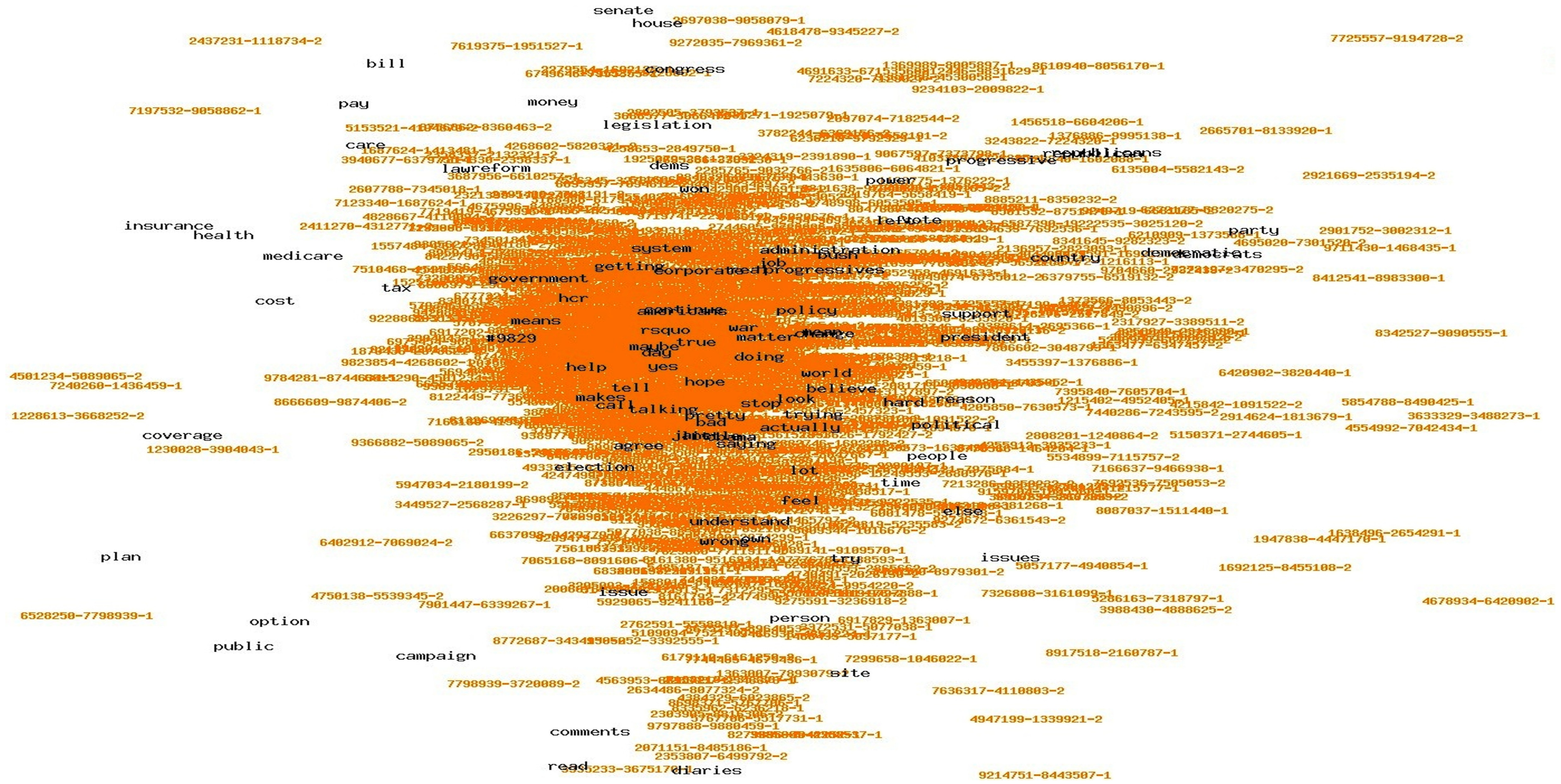
$$\theta_{d,k} = \frac{\Sigma_n W_{d,n} \phi_{d,n,k}}{\Sigma_{n,k} W_{d,n} \phi_{d,n,k}} \quad (6)$$

$$\phi_{d,n,k} = \frac{\theta_{d,k} \beta_{k,n}}{\Sigma_k \theta_{d,k} \beta_{k,n}} \quad (7)$$

Maximize via: (5) $\to$ (6) $\to$ (7) $\to$ (5) $\to$ ...

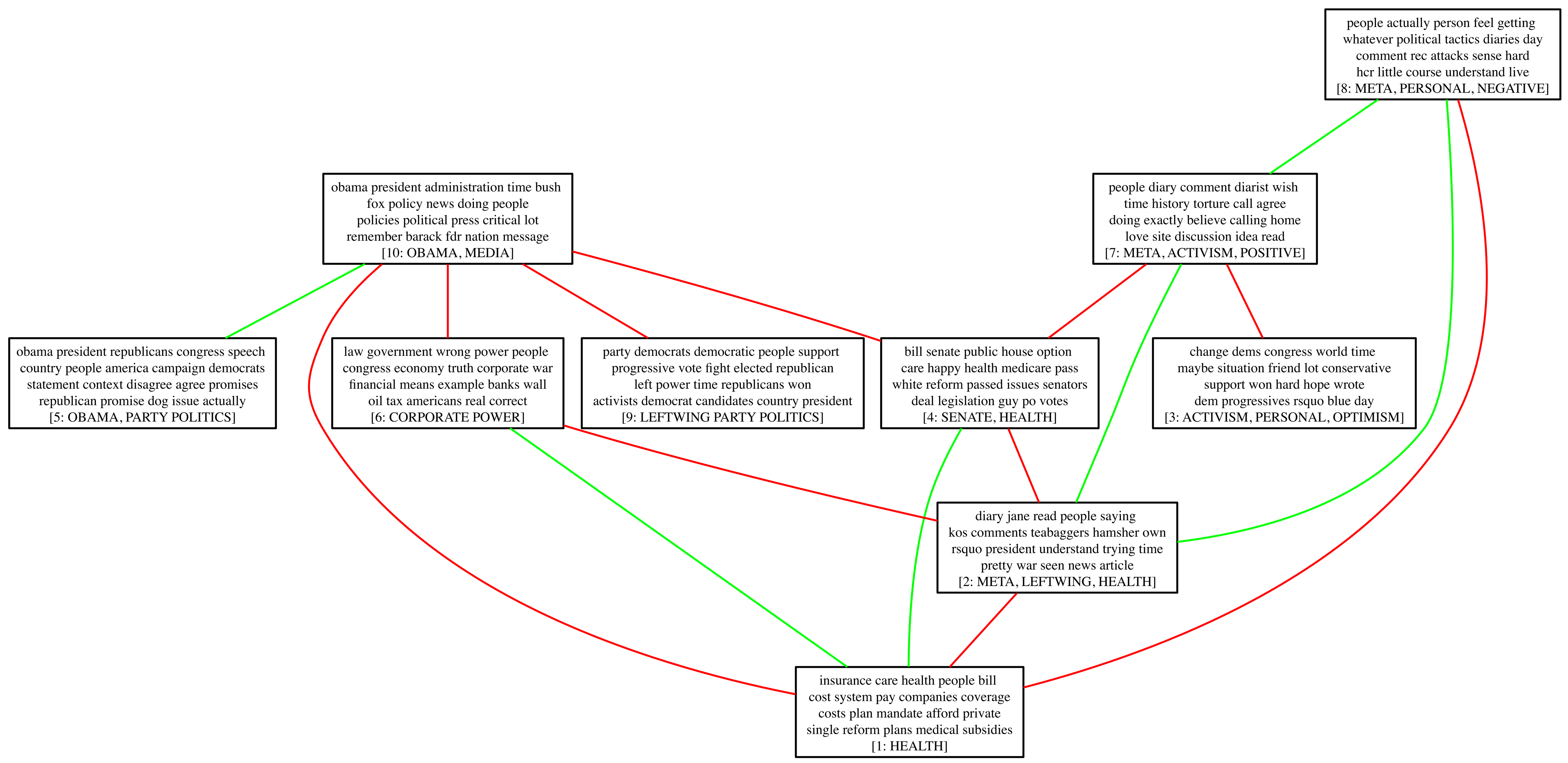
Motivation: Debate on the left is not low dimensional

Comments and words (W) can be scaled in the same space using a variant of principal component analysis. However, the top eigenvalues are roughly equal, and plots of word scores show there are no obvious left-right oppositions, just clusters of related topics – hence the need for a topic model.



Network of correlated topics

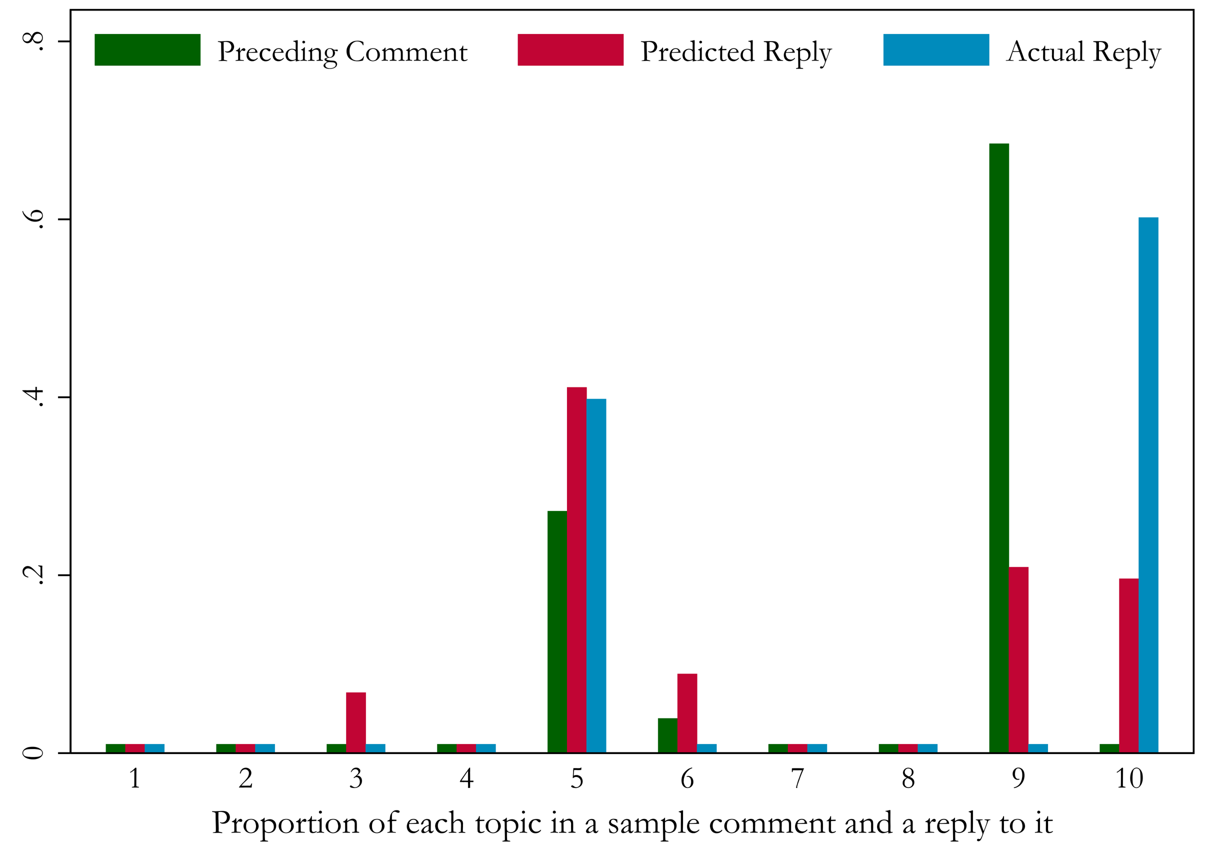
In 2010, the major issues on dailykos.com revolved around health care and Obama. After fitting the model, the K=10 topics can be represented in a least-energy network, with green or red links between positively or negatively correlated topics (Σ). Each topic is characterized by the top words from β_k (with interpretations in brackets).



Conclusion

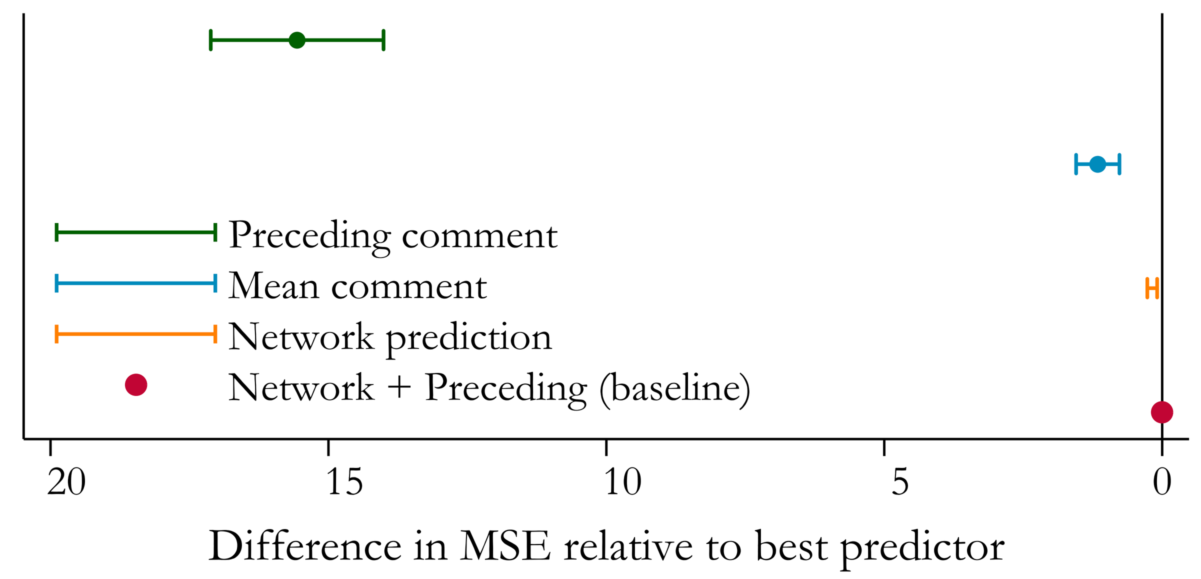
Correlated topic models represent a debate as a network of correlated facts, ideas, and arguments that speakers selectively deploy when a topic appears relevant to preceding speech, yet missing from it. This non-spatial, network-based model of opinion allows us to predict speech content, and will next be used to model and predict opinion change as listeners add new topics to their mental network.

Example: Topic distribution



Topic distributions for a comment and reply, along with the predicted reply distribution based on the correlated network (see below). The network correctly predicts a reply that shifts from topic 9 to 10, while sticking with 5.

Predicting replies



The topic distribution of a reply can best be predicted by replacing topic covariances Σ with regression coefficients A. If v_1 is the topic distribution of the initial post, $v' = Av_1$ is the distribution of topics connected via the network to v_1. A reply v_2 is best predicted by $\alpha v' + (1-\alpha)v_1$, where $\alpha \approx 0.9$.