Modeling the Effects of Interpersonal Communication on Long-Term Opinion Change using Topic Models

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

Introduction

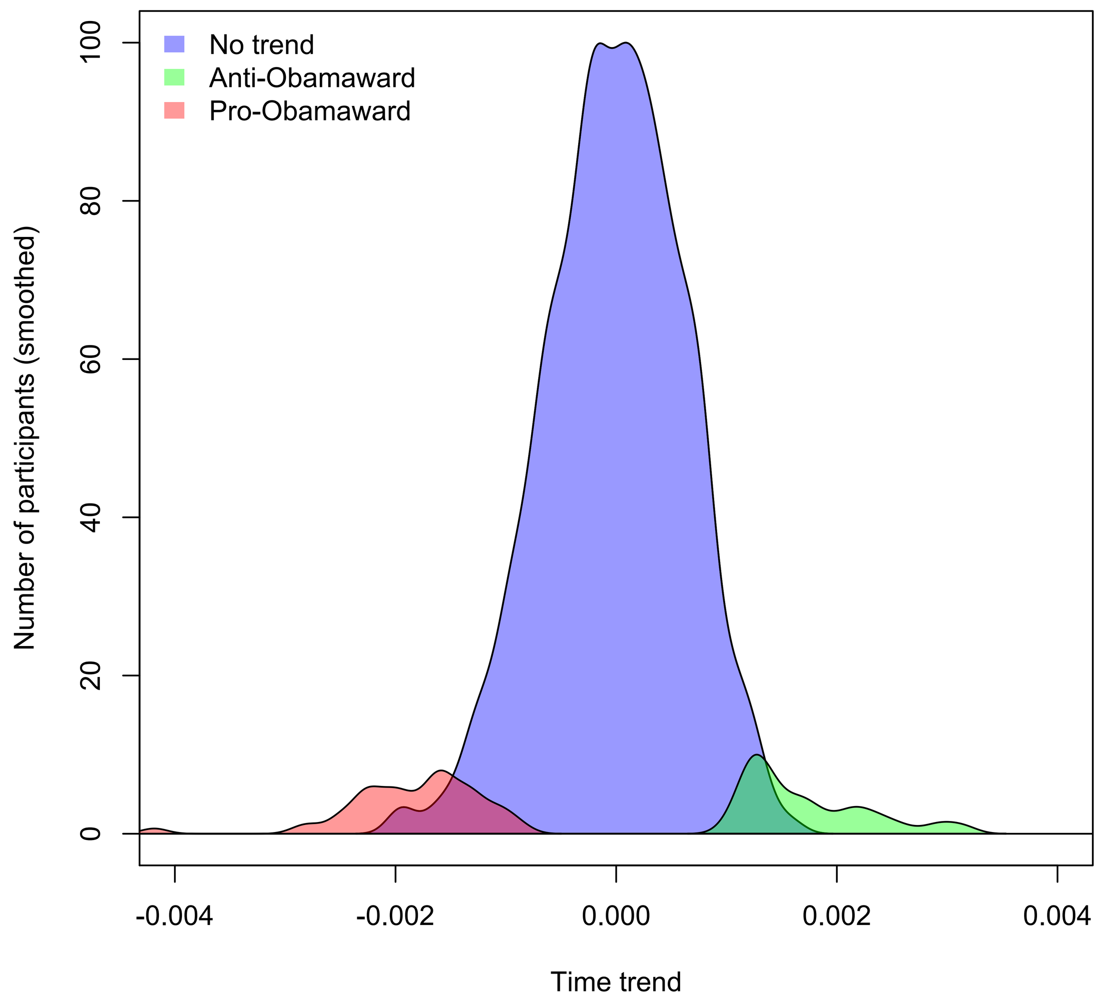
Political opinions form and change under the ceaseless barrage of conversation and argument. Dailykos.com, the largest political forum on the left, records discussions among hundreds of thousands of users over many years. Which "diaries" users vote to "recommend" reveals a prevalent ideological disagreement around President Obama, and topic models allow us to infer the concepts at play in these discussions. By matching those whose opinions shifted between 2009-2010 with similar users who did not shift, the causal effects of speech on opinion change can be inferred.

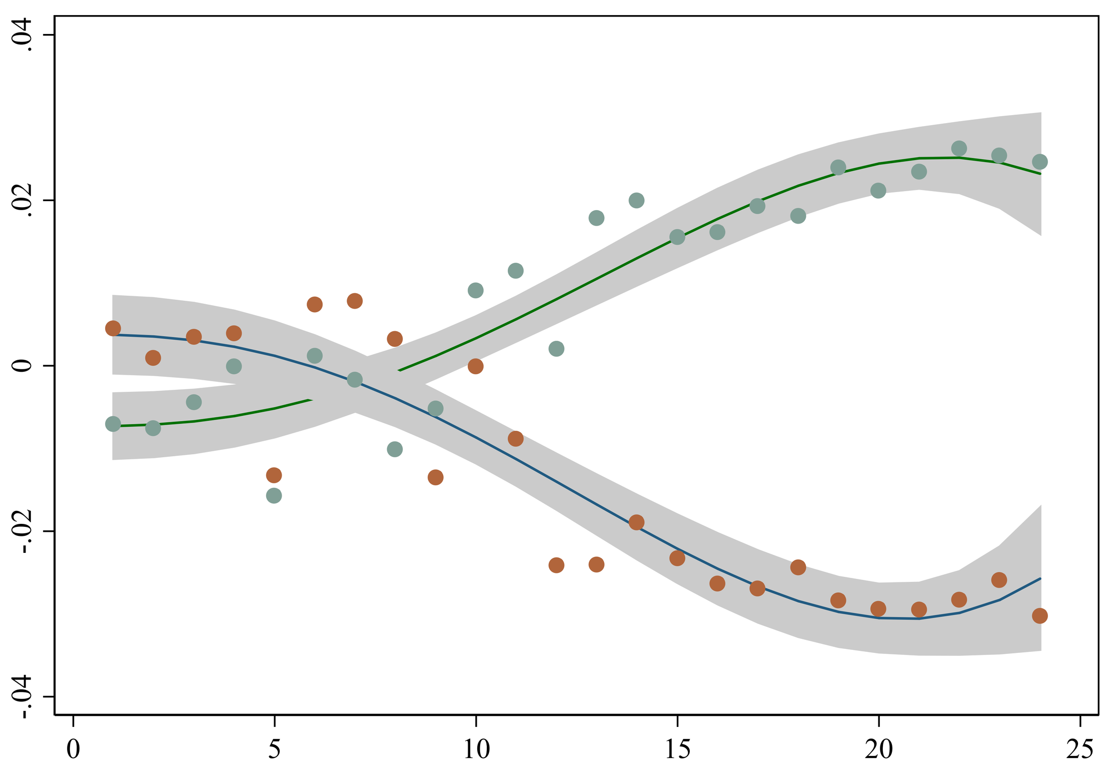
Behavior-based opinion shifts over time

Dailykos.com is viewed by over 2 million activists on the left, and was responsible for millions of dollars of campaign donations in 2008. Users can vote to "recommend" other diaries, which significantly affects the placement of those diaries on the front page, and thus the quantity of attention they receive. This matrix of users, diaries, and recommendations can be used to scale both users and diaries using Principal Component Analysis. The fast Nipals algorithm is used to scale 1000 users and 4,000 diaries between 2009 and 2010, proceeding in month-long chunks. In all months either the first or (occasionally) second principal component corresponds to a pro/anti Obama dimension (anti from the left); this component is extracted for all months.

Each user's vote-based ideology is thus measured over 24 months. To identify those who have significantly trended in a pro- or anti-Obamaward direction, each user's time trend is estimated, with associated p-value. With 1000 measurements, multiple testing corrections are necessary to distinguish true from chance trenders. To minimize the False Discovery Rate (proportion of false positives), the Benjamini and Hochberg (1995) correction is employed: First order all K observations by increasing p-value, then take largest k such that $P_k \leq \frac{k}{K}\alpha$ for a preferred error threshold α, and accept all observations ranked above the kth.

This reveals that about 5% of users trended significantly pro-Obamaward over the course of the two years, with similar numbers trending anti-Obamaward; the rest showed no discernible change in "voting" (recommending) behavior. Notably, most of this change seems to have taken place in 2010, allowing us to match trenders against non-trenders in 2009.





The opinion of pro-Obama (red) and anti-Obama (green) trenders over 24 months

Procedure

1. Estimate user ideology using chained principal component analysis of users' vote-like "recommendations" of diaries over 24 months, 2009-2010.

2. The dominant dimension of disagreement is pro/anti Obama (from the left). Use multiple testing corrections to identify those who shift towards or against Obama during this period.

3. Employ the correlated topic model to estimate an interlinked network of concepts underlying all discussions. Supervised modeling produces topics related to the dominant voting dimension.

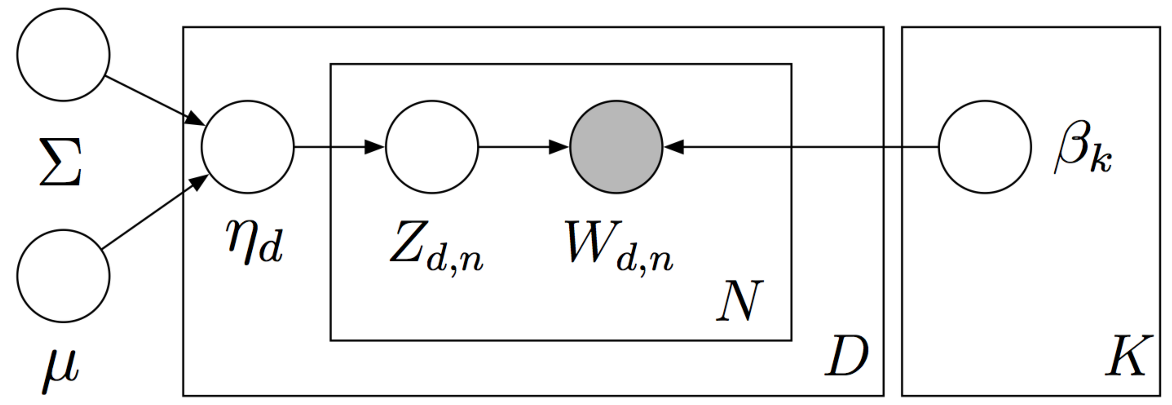
4. Match trenders with non-trenders in 2009 based on vote-based ideology and speech topics. Examine difference-in-differences to discover what received speech may have affected ideology independent of pre-existing ideological or speaking tendencies.

A correlated topic model

In this variant of Blei and Lafferty's Correlated Topic Model (2007), each topic has a characteristic word distribution, and each document has a characteristic topic distribution. The generative process is that a document draws a topic distribution, and then words are generated one at a time, first drawing a topic from the document's topic distribution, then drawing a word from that topic's word distribution. Topics themselves are correlated, in what is interpreted here as a network of concepts.

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: $\beta = K \times N$	$\Sigma = K \times K$



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

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

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

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

Fit is maximized via coordinate ascent: (5) → (6) → (7) → (5) → ...

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

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

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

$$\phi_{d,n,k} = \frac{\theta_{d,k}\beta_{k,n}}{\Sigma_k \theta_{d,k}\beta_{k,n}} \tag{7}$$

Comments and Responses as Markov Process

For threaded online conversations, the conversation tree can be divided into comment pairs: comment $d1$ and a comment in response to it, $d2$. Empirically, we can predict the response's topic distribution $\eta_{d2} \approx \alpha \mathbf{A} \eta_{d1} + (1-\alpha)\eta_{d1}$, where $\alpha \approx 0.9$ and $\mathbf{A}$ is the K x K autoregressive matrix from regressing each topic vector $\eta_{d,k}$ on each other $\forall$ k. That is, the topics of a responding comment are those topics most strongly linked to – but different from – those of the first comment.

We can intensify the empirical finding assuming a hidden markov structure. Unlike previous hidden markov topic models, (1) The distribution of topics is not changing, but instead one document is "causing" the second. (2) We only consider linked pairs rather than the extended chains (modeling full discussion trees is a later project). Analogous to the forward-backward algorithm, we simply draw step (6) towards the assumed structure $\theta_{d2} = \mathbf{A}\theta_{d1}$ for each cycle of the maximization procedure. A strong prior increases topic predictive accuracy at the cost of topic-data fit.

Supervised Topic Model

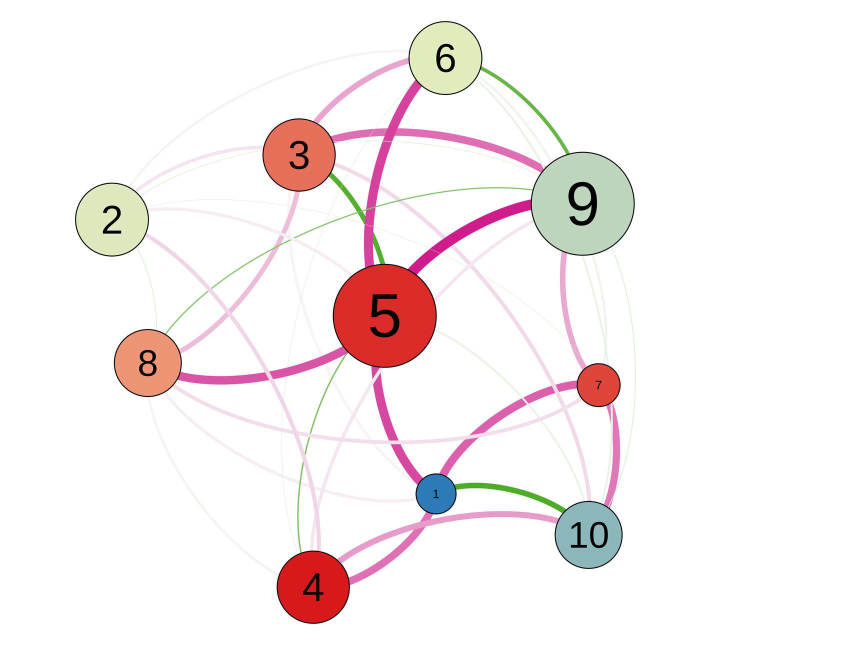
Estimating an unconstrained topic model produces numerous topics with no connection to the dominant pro/anti Obama dimension of disagreement. Estimating (6) - (7) on 125,000 documents at once is computationally infeasible.

Supervised topic modeling instead estimates the parameters using a subset of comments by only the most ideological users (high PCA). β_k is retained from this, and used in estimating all other comments. The resultant topics correlate highly with the first principal component, and estimation is speeded 10x-100x.

Conclusion

The topic modeling allows us to infer an interrelated network of ideas and predict the content of short-term responses. It allows us to identify who is likely to change their opinion and vote-like behavior over time, and to identify potential effects of speech on opinion change. Those who became stronger supporters of Obama in 2010 appear to have been affected by chance encounters with upbeat, non-political discussants, as well as unusually low encounters with discussants raising the BP oil spill and unemployment.

Correlated Topics as a network of interconnected concepts



Green links = positive correlation; red links = negative correlation. Red topics correlated with pro-Obama voters; green topics correlated with anti-Obama voters. Force Atlas network, node and link size by eigenvector centrality. Generated with *Gephi*.

Top words for each topic from $\beta_{k,n}$:

1: bill, insurance, care, system, health, financial, america, companies, economic, money, industry, bank, banks, medicare, buy

2: government, president, americans, oil, private, look, based, world, day, million, post, american, add, water, people

3: wrong, hate, getting, please, win, hope, guy, people, little, mind, fucking, bad, day, doing, won

4: time, people, left, news, diary, story, called, read, personal, agree, days, political, sense, folks, mean

5: people, actually, life, love, family, own, lot, help, little, feel, social, time, believe, god, sorry

6: public, people, obama, care, option, health, change, trying, reform, saying, doing, single, politics, wrong, matter

7: diary, person, diaries, comments, people, comment, support, thanks, diarist, fdr, kos, paid, school, whatever, agree

8: obama, president, bush, torture, war, administration, rights, congress, iraq, vote, law, crimes, legal, political, cheney

9: house, party, democrats, money, white, democratic, vote, line, senate, republicans, time, republican, progressive, support, corporate

10: people, jobs, class, money, tax, war, country, world, military, black, racism, middle, power, pay, poor

Estimating the effects of speech on belief using Matching, 2009-2010

How might we identify causal pathways in dynamic networks of discussion and argument lasting years, when the content of a user's comments naturally affects the content of responses to that user?

Matching those whose opinions have trended in one direction or another with non-trenders with similar (vote-based) ideology and (topic-based) speech patters in year 1 (2009) allows us to estimate potentially exogenous effects that are independent of users' speech or ideological characteristics.

Trenders are matched with non-trending controls on PCA values for each of 12 months plus speech distributions over 10 topics in 2009, using a new matching algorithm, "Blossom." This algorithm outperforms existing matching algorithms such as GenMatch (Sekhon, 2011) often by 10x.

Comparing difference-in-differences, 2009-2010 between trenders and matches shows that anti-Obama trenders began speaking differently for exogenous reasons in 2010, but pro-Obama trenders, though unchanged in speech, may have changed their opinions and voting behavior due to chance variations in received speech.