The Exaggerated Life of Death Panels: The Limits of Persuasion in the 2009-2017 Health Care Debate

Daniel J. Hopkins
Associate Professor
University of Pennsylvania

January 12th, 2018

---

Robert Biemesderfer, Patrick Gavin, Douglas Kovel, Jake Sticka, Anton Strezhnev, Clare Tilton, and Marzena Zukowska provided invaluable research assistance for this project. Will Hobbs serves as my co-author on some of this research.
Puzzle: How does the rhetoric of politicians, other political elites influence public opinion?
Puzzle: How does the rhetoric of politicians, other political elites influence public opinion?

Widespread presumption that elite messaging matters
Puzzle: How does the rhetoric of politicians, other political elites influence public opinion?

Widespread presumption that elite messaging matters

2000 Congressional candidate: I’m able to “best articulate and frame the issues that are most important to voters”
Puzzle: How does the rhetoric of politicians, other political elites influence public opinion?

- Widespread presumption that elite messaging matters
- 2000 Congressional candidate: I’m able to “best articulate and frame the issues that are most important to voters”
Puzzle: How does the rhetoric of politicians, other political elites influence public opinion?

Widespread presumption that elite messaging matters

2000 Congressional candidate: I’m able to “best articulate and frame the issues that are most important to voters” —Barack Obama
Example: “The America I know and love is not one in which my parents or my baby with Down Syndrome will have to stand in front of Obama’s ‘death panel’ so his bureaucrats can decide, based on a subjective judgment of their ‘level of productivity in society,’ whether they are worthy of health care. Such a system is downright evil.” — Former AK Gov. Sarah Palin
We asked: “As you may know, a health reform bill was signed into law in 2010. Given what you know about the health reform law, do you have a generally favorable or generally unfavorable opinion of it?”

Figure: Attitudes towards health care reform by month (Kaiser Family Foundation).
Framing in the Health Care Debate

What explains initial decline in public support for health care reform?
What explains initial decline in public support for health care reform?
Focus here: issue framing / messaging / persuasive rhetoric
What explains initial decline in public support for health care reform?
Focus here: issue framing / messaging / persuasive rhetoric
  Widely invoked by commentators; significant political science literature on issue framing, persuasion
What explains initial decline in public support for health care reform?

Focus here: issue framing / messaging / persuasive rhetoric
- Widely invoked by commentators; significant political science literature on issue framing, persuasion

Research questions: Did political leaders’ health care arguments influence public opinion—and can automated content analysis help us measure that influence? How did public opinion on ACA evolve over time?
Past issue framing research: primarily experimental; commonly finds persuasive impact of messages
Automated Methods to Measuring Framing Effects

- Past issue framing research: primarily experimental; commonly finds persuasive impact of messages
- Challenges in real-world setting:
  
  1. Measurement of Elites' Arguments: difficult to identify salient arguments in fractured media environment
  2. Measurement of Public Opinion: prior research employs closed-ended survey questions, making it difficult to identify subtle shifts in public opinion
  3. Causal Inference: possibility of reverse causation as politicians adopt words used by public (e.g. Page and Shapiro 2000)

Automated content analysis can improve measurement, facilitate study of persuasion
Past issue framing research: primarily experimental; commonly finds persuasive impact of messages

Challenges in real-world setting:

1. Measurement of Elites’ Arguments: difficult to identify salient arguments in fractured media environment
Past issue framing research: primarily experimental; commonly finds persuasive impact of messages

Challenges in real-world setting:

1. Measurement of Elites’ Arguments: difficult to identify salient arguments in fractured media environment
2. Measurement of Public Opinion: prior research employs closed-ended survey questions, making it difficult to identify subtle shifts in public opinion
Past issue framing research: primarily experimental; commonly finds persuasive impact of messages

Challenges in real-world setting:

1. **Measurement of Elites’ Arguments**: difficult to identify salient arguments in fractured media environment
2. **Measurement of Public Opinion**: prior research employs closed-ended survey questions, making it difficult to identify subtle shifts in public opinion
3. **Causal Inference**: possibility of reverse causation as politicians adopt words used by public (e.g. Page and Shapiro 2000)
Past issue framing research: primarily experimental; commonly finds persuasive impact of messages

Challenges in real-world setting:

1. Measurement of Elites’ Arguments: difficult to identify salient arguments in fractured media environment

2. Measurement of Public Opinion: prior research employs closed-ended survey questions, making it difficult to identify subtle shifts in public opinion

3. Causal Inference: possibility of reverse causation as politicians adopt words used by public (e.g. Page and Shapiro 2000)

Automated content analysis can improve measurement, facilitate study of persuasion
Empirical analyses:
Empirical analyses:

1. use Latent Dirichlet Allocation (LDA) to measure elite health care arguments
Empirical analyses:

1. use Latent Dirichlet Allocation (LDA) to measure elite health care arguments
2. use topic models, pivot scaling to measure public opinion via open-ended survey questions
Empirical analyses:

1. Use Latent Dirichlet Allocation (LDA) to measure elite health care arguments.
2. Use topic models, pivot scaling to measure public opinion via open-ended survey questions.
3. Use distribution of words, distance metrics to measure similarity in elite, mass word choice over time.
Take-home Points:

1. Elite arguments → limited but discernible effects on public word choice
Take-home Points:

1. Elite arguments → limited but discernible effects on public word choice
2. Public → shifts to emphasize more concrete aspects of ACA over time
Take-home Points:

1. Elite arguments → limited but discernible effects on public word choice
2. Public → shifts to emphasize more concrete aspects of ACA over time
3. Use of natural language can improve measurement, allow direct comparison between elite, mass speech
Issue frame: rhetorical structure which calls attention to a subset of the relevant considerations (Kinder 1998; Chong and Druckman 2010)
Framing

- Issue frame: rhetorical structure which calls attention to a subset of the relevant considerations (Kinder 1998; Chong and Druckman 2010)
- Tell-tale sign of frames: distinctive vocabulary (e.g. ‘death panels’)
Framing

• Issue frame: rhetorical structure which calls attention to a subset of the relevant considerations (Kinder 1998; Chong and Druckman 2010)

• Tell-tale sign of frames: distinctive vocabulary (e.g. ‘death panels’)

• Political science studies framing, persuasion primarily through survey experiments (e.g. Iyengar and Kinder 1987; Chong and Druckman 2007, 2010)
Challenges to external validity of framing, persuasion experiments:
Limits of Persuasive Rhetoric

Challenges to external validity of framing, persuasion experiments:
1. Coordination problems among elites (Bush, Cruz, Rubio et al. 2016)
Challenges to external validity of framing, persuasion experiments:

1. Coordination problems among elites (Bush, Cruz, Rubio et al. 2016)
2. Elites choose arguments already in public use (e.g. Jacobs and Shapiro)
Challenges to external validity of framing, persuasion experiments:

1. Coordination problems among elites (Bush, Cruz, Rubio et al. 2016)
2. Elites choose arguments already in public use (e.g. Jacobs and Shapiro)
3. Fractured media environment
Empirical goal #1: measure changes in elite rhetoric over time
Empirical goal #1: measure changes in elite rhetoric over time

Unsupervised techniques (see also Grimmer and King 2011; Spirling 2011); allow for identification of many, potentially unanticipated arguments
Empirical goal #1: measure changes in elite rhetoric over time

Unsupervised techniques (see also Grimmer and King 2011; Spirling 2011); allow for identification of many, potentially unanticipated arguments

LDA (Blei, Ng, and Jordan 2003): represent each document as finite mixture of $K$ topics; each topic is a probability distribution over words
Empirical goal #1: measure changes in elite rhetoric over time

Unsupervised techniques (see also Grimmer and King 2011; Spirling 2011); allow for identification of many, potentially unanticipated arguments

LDA (Blei, Ng, and Jordan 2003): represent each document as finite mixture of $K$ topics; each topic is a probability distribution over words

This: approximation of conception of frames (reflect distinctive vocabulary)
Press releases: clear statement of officials’ desired frames before journalists’ filtering (see also Grimmer 2010)
Press releases: clear statement of officials’ desired frames before journalists’ filtering (see also Grimmer 2010)

Data: 1,488 press releases on health care by U.S. Senators in 2009, 2010
Press releases: clear statement of officials’ desired frames before journalists’ filtering (see also Grimmer 2010)

Data: 1,488 press releases on health care by U.S. Senators in 2009, 2010

Analyze 2,043 word stems that appear in more than 1% of press releases
Figure: This figure depicts the number of press releases by month and identifies key events in the legislative timeline.
Figure: Sample cluster 1. Model: LDA, fit to 1,488 press releases.
Figure: Sample cluster 2. Model: LDA, fit to 1,488 press releases.
Figure: Variation in Topics Over Time. Model: LDA, fit to 1,488 press releases.
- Dems: public option wanes; emphasize extending coverage
Notes on LDA Results

- Dems: public option wanes; emphasize extending coverage
- GOP: procedure; tax increase; January ’10 Medicare spike
Dems: public option wanes; emphasize extending coverage

GOP: procedure; tax increase; January ’10 Medicare spike

Punctuation: mean standard errors over time are 0.08 (D), 0.07 (R)
Dems: public option wanes; emphasize extending coverage
GOP: procedure; tax increase; January ’10 Medicare spike
Punctuation: mean standard errors over time are 0.08 (D), 0.07 (R)
Core point: elite frames vary with events
Empirical goal #2: measure changes in public opinion over time
Measuring Public Opinion

- Empirical goal #2: measure changes in public opinion over time
- Kaiser Family Foundation conducts monthly health care tracking polls of American adults
Empirical goal #2: measure changes in public opinion over time

Kaiser Family Foundation conducts monthly health care tracking polls of American adults

Typical empirical strategy: examine relationship between rhetoric, closed-ended survey responses
Here: apply LDA to open-ended responses in seven surveys about health care reform (July 2009 to November 2011)
Here: apply LDA to open-ended responses in seven surveys about health care reform (July 2009 to November 2011)

Measure effects more subtle than opinion change
Here: apply LDA to open-ended responses in seven surveys about health care reform (July 2009 to November 2011)

Measure effects more subtle than opinion change

Open-ended questions cover period from before health care reform became salient until after 2010 mid-term elections
Here: apply LDA to open-ended responses in seven surveys about health care reform (July 2009 to November 2011)

Measure effects more subtle than opinion change

Open-ended questions cover period from before health care reform became salient until after 2010 mid-term elections

Number of topics=6
Figure: Model: LDA, fit to 6,355 open-ended survey responses in 7 surveys.
LDA on Open-ended Responses

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>government</td>
<td>people</td>
<td>care</td>
<td>afford</td>
<td>money</td>
<td>dont</td>
</tr>
<tr>
<td>2</td>
<td>dont</td>
<td>insurance</td>
<td>health</td>
<td>people</td>
<td>cost</td>
<td>bill</td>
</tr>
<tr>
<td>3</td>
<td>control</td>
<td>help</td>
<td>access</td>
<td>cant</td>
<td>medicare</td>
<td>trying</td>
</tr>
<tr>
<td>4</td>
<td>business</td>
<td>companies</td>
<td>system</td>
<td>lot</td>
<td>pay</td>
<td>country</td>
</tr>
<tr>
<td>5</td>
<td>involved</td>
<td>coverage</td>
<td>americans</td>
<td>help</td>
<td>tax</td>
<td>healthcare</td>
</tr>
<tr>
<td>6</td>
<td>run</td>
<td>conditions</td>
<td>affordable</td>
<td>insurance</td>
<td>taxes</td>
<td>doesnt</td>
</tr>
<tr>
<td>7</td>
<td>believe</td>
<td>able</td>
<td>provide</td>
<td>health</td>
<td>cut</td>
<td>law</td>
</tr>
<tr>
<td>8</td>
<td>private</td>
<td>existing</td>
<td>universal</td>
<td>care</td>
<td>dont</td>
<td>medicine</td>
</tr>
<tr>
<td>9</td>
<td>choice</td>
<td>preexisting</td>
<td>believe</td>
<td>dont</td>
<td>lot</td>
<td>time</td>
</tr>
<tr>
<td>10</td>
<td>federal</td>
<td>age</td>
<td>reform</td>
<td>able</td>
<td>class</td>
<td>socialized</td>
</tr>
<tr>
<td>11</td>
<td>care</td>
<td>premiums</td>
<td>country</td>
<td>helps</td>
<td>doctors</td>
<td>system</td>
</tr>
<tr>
<td>12</td>
<td>buy</td>
<td>children</td>
<td>free</td>
<td>buy</td>
<td>people</td>
<td>citizens</td>
</tr>
<tr>
<td>13</td>
<td>purchase</td>
<td>que</td>
<td>costs</td>
<td>pay</td>
<td>paying</td>
<td>public</td>
</tr>
<tr>
<td>14</td>
<td>forced</td>
<td>income</td>
<td>canada</td>
<td>benefit</td>
<td>middle</td>
<td>change</td>
</tr>
<tr>
<td>15</td>
<td>individual</td>
<td>covered</td>
<td>coverage</td>
<td>time</td>
<td>healthcare</td>
<td>passed</td>
</tr>
<tr>
<td>16</td>
<td>program</td>
<td>covers</td>
<td>countries</td>
<td>healthcare</td>
<td>costs</td>
<td>senior</td>
</tr>
<tr>
<td>17</td>
<td>tell</td>
<td>pre</td>
<td>quality</td>
<td>expensive</td>
<td>social</td>
<td>read</td>
</tr>
<tr>
<td>18</td>
<td>freedom</td>
<td>low</td>
<td>american</td>
<td>sick</td>
<td>dollars</td>
<td>didnt</td>
</tr>
<tr>
<td>19</td>
<td>socialism</td>
<td>helps</td>
<td>change</td>
<td>elderly</td>
<td>debt</td>
<td>feel</td>
</tr>
<tr>
<td>20</td>
<td>choose</td>
<td>health</td>
<td>world</td>
<td>issues</td>
<td>medicaid</td>
<td>thats</td>
</tr>
</tbody>
</table>

Table: This table presents the most commonly occurring words in each of the six clusters of open-ended responses identified through LDA.
Alternative Approach: Pivot Scores

- Use pivot scores (Hobbs 2017)
Alternative Approach: Pivot Scores

- Use pivot scores (Hobbs 2017)
- Low-dimensional representation of key differences in text
July 2009
Empirical goal #3: evaluate elite rhetoric, mass word usage on same scale
Empirical goal #3: evaluate elite rhetoric, mass word usage on same scale

Can use vectors of word probabilities to compare elite \( p_{pr}(w_n) \), mass language \( p_{sur}(w_n) \) over time
Empirical goal #3: evaluate elite rhetoric, mass word usage on same scale

Can use vectors of word probabilities to compare elite ($p_{pr}(w_n)$), mass language ($p_{sur}(w_n)$) over time

Address alternative that elites → adopt mass-level word choice
Figure: Difference in word usage for supporters, opponents.
1. Identify 100 word stems that meet criteria for both press releases, open-ended responses
2. Calculate distance between press release language, language in prior open-ended responses
3. Calculate distance between press release language, language in subsequent open-ended responses
4. Calculate (3) - (2)
5. Calculate uncertainty of (4) via bootstrapping
Public: Distribution of Words

Difference in Words:
Press Releases
KL Distance
Dems
Jul→Nov
P=0.95
Dems
Nov→May
P=0.925
GOP
Jul→Nov
P=0.004
GOP
Nov→May
P=0.003

Difference in Words:
Press Releases
Euclidean Distance
Dems
Jul→Nov
P=0.795
Dems
Nov→May
P=0.143
GOP
Jul→Nov
P=0.893
GOP
Nov→May
P=0.143

Correlations in Words:
Press Releases
Pearson's Correlation
Dems
Jul→Nov
P=0.004
Dems
Nov→May
P=0.01
GOP
Jul→Nov
P=0.795
GOP
Nov→May
P=0.01

Figure: Change in the distance between words in press releases, survey responses
“This may be your last chance to weigh the consequences of taking the first step toward establishment of socialized medicine in the United States... When costs get out of line – and let me assure you, they will – there are three possible courses of action. The first is to reduce the benefits; the second is to increase taxes; the third is to impose government controls of the services in an attempt to control costs.”
“This may be your last chance to weigh the consequences of taking the first step toward establishment of socialized medicine in the United States... When costs get out of line – and let me assure you, they will – there are three possible courses of action. The first is to reduce the benefits; the second is to increase taxes; the third is to impose government controls of the services in an attempt to control costs.” — Dr. Donovan F. Ward, A.M.A. President, 1965
Rhetoric on health care → appears similar across decades; limits scope for actor-specific changes.
Rhetoric on health care → appears similar across decades; limits scope for actor-specific changes

Elite-level frames: do vary over time
Conclusions

- Rhetoric on health care appears similar across decades; limits scope for actor-specific changes.
- Elite-level frames: do vary over time.
- Opinions, language around health care reform: generally stable; some evidence of adoption of elite-level word choice.

Daniel Hopkins (U. of Pennsylvania)
The Exaggerated Life of Death Panels
January 12th, 2018
Conclusions

- Rhetoric on health care → appears similar across decades; limits scope for actor-specific changes
- Elite-level frames: do vary over time
- Opinions, language around health care reform: generally stable; some evidence of adoption of elite-level word choice
- Rationales shift to be increasingly concrete
Conclusions

- Rhetoric on health care → appears similar across decades; limits scope for actor-specific changes
- Elite-level frames: do vary over time
- Opinions, language around health care reform: generally stable; some evidence of adoption of elite-level word choice
- Rationales shift to be increasingly concrete
- Substantive conclusion: limited but real evidence of framing
Conclusions

- Rhetoric on health care → appears similar across decades; limits scope for actor-specific changes
- Elite-level frames: do vary over time
- Opinions, language around health care reform: generally stable; some evidence of adoption of elite-level word choice
- Rationales shift to be increasingly concrete
- Substantive conclusion: limited but real evidence of framing
- “Death Panels” — almost never used by Members of Congress or the public
Consider convergence among sub-groups of respondents
Next Steps

- Consider convergence among sub-groups of respondents
- Joint modeling of elite frames, open-ended opinions
Next Steps

- Consider convergence among sub-groups of respondents
- Joint modeling of elite frames, open-ended opinions
- Extend definition of political elites (e.g. Glenn Beck, etc.); consider activists, blog posts, Twitter, etc.
Next Steps

- Consider convergence among sub-groups of respondents
- Joint modeling of elite frames, open-ended opinions
- Extend definition of political elites (e.g. Glenn Beck, etc.); consider activists, blog posts, Twitter, etc.
- Study impact of paid advertising
Next Steps

- Consider convergence among sub-groups of respondents
- Joint modeling of elite frames, open-ended opinions
- Extend definition of political elites (e.g. Glenn Beck, etc.); consider activists, blog posts, Twitter, etc.
- Study impact of paid advertising
- Apply to other issues
Latent Dirichlet Allocation

- Vocabulary \( \{1, \ldots, V\} \) of words
- Corpus: \( M \) documents \( \mathbf{D} = \{W_1, W_2, \ldots, W_M\} \)
- For each document \( W_m \):
  1. Choose \( N \sim \text{Poisson}(\xi) \)
  2. Choose \( \theta \sim \text{Dirichlet}(\alpha) \)
  3. For each of \( N \) words \( w_n \):
     - Choose topic \( z_n \sim \text{Multinomial}(\theta) \)
     - Choose word \( w_n \sim \text{Multinomial}(z_n, \beta) \)

- \( \beta \) is a \( K \times V \) matrix of word probabilities
- Mixed membership model
LDA and CTM

- Can estimate LDA via MCMC (partially collapsed Gibbs sampler; van Dyk and Park 2008) or Variational Inference (Jordan et al. 1999; Grimmer 2011)
- Dirichlet model of topics → topics are uncorrelated
- Can instead use Correlated Topic Model (Blei and Lafferty 2006)
- Draw $K - 1$ topic proportions from logistic normal; formally model covariance of topics
  - $\eta \sim Normal(\mu, \Sigma)$
  - $f(\eta_i) = \frac{\exp(\eta_i)}{\sum_j \exp(\eta_j)}$
- CTM → loses conjugacy between Dirichlet, Multinomial; necessitates estimation via Variational Expectation-Maximization
- Structural Topic Model (STM; Roberts et al. 2014)
- Builds on CTM:
  - $\eta \sim Normal(\gamma X, \Sigma)$
  - (Also allow covariates $U$ to influence word frequencies)
- Documents = own prior distributions governed by $X$
**Results: keywords/pivots**

<table>
<thead>
<tr>
<th>Keywords</th>
<th>Dimension 1</th>
<th></th>
<th>Dimension 2</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pro</td>
<td>Anti</td>
<td>Pro</td>
<td>Anti</td>
</tr>
<tr>
<td>preexisting conditions</td>
<td>government</td>
<td>unconstitutional</td>
<td>universal</td>
<td>premiums</td>
</tr>
<tr>
<td>parents</td>
<td>run</td>
<td>involved</td>
<td>health</td>
<td>deductibles</td>
</tr>
<tr>
<td>children</td>
<td>us</td>
<td>us</td>
<td>access</td>
<td>money</td>
</tr>
<tr>
<td>stay</td>
<td>us</td>
<td>us</td>
<td>everyone</td>
<td>working</td>
</tr>
<tr>
<td>pre</td>
<td>economy</td>
<td>direction</td>
<td>step</td>
<td>even</td>
</tr>
<tr>
<td>young</td>
<td>direction</td>
<td>direction</td>
<td>income</td>
<td></td>
</tr>
<tr>
<td>age</td>
<td>peoples</td>
<td>needs</td>
<td>middle</td>
<td></td>
</tr>
<tr>
<td>opportunity</td>
<td>business</td>
<td>americans</td>
<td>high</td>
<td></td>
</tr>
<tr>
<td>coverage</td>
<td>much</td>
<td>affordable</td>
<td>medicare</td>
<td></td>
</tr>
<tr>
<td>helps</td>
<td>control</td>
<td>provide</td>
<td>going</td>
<td></td>
</tr>
<tr>
<td>year</td>
<td>dont</td>
<td>right</td>
<td>low</td>
<td></td>
</tr>
<tr>
<td>now</td>
<td>medicine</td>
<td>preexisting</td>
<td>less</td>
<td></td>
</tr>
<tr>
<td>able</td>
<td>president</td>
<td>single</td>
<td>fine</td>
<td></td>
</tr>
<tr>
<td>insurance</td>
<td>end</td>
<td>every</td>
<td>taxes</td>
<td></td>
</tr>
<tr>
<td>can</td>
<td>everything</td>
<td>conditions</td>
<td>elderly</td>
<td></td>
</tr>
<tr>
<td>uninsured</td>
<td>taxes</td>
<td>country</td>
<td>higher</td>
<td></td>
</tr>
<tr>
<td>allows</td>
<td>socialized</td>
<td>provides</td>
<td>everything</td>
<td></td>
</tr>
<tr>
<td>condition</td>
<td>socialism</td>
<td>issues</td>
<td>doctors</td>
<td></td>
</tr>
<tr>
<td>still</td>
<td>every</td>
<td>system</td>
<td>put</td>
<td></td>
</tr>
</tbody>
</table>

**Figure: Pivot Scaling**
Open-Ended Opinions (3)

Figure: Pivot Scaling
# Figure: Pivot Scaling

1. Standardize word co-occurrences $G$ with diagonal $D_g$:

$$X = D_g^{-1} G; \quad G = M^T M$$

2. Weight out-of-sample data $W$ by word counts:

$$Y = D_g^b W^{\alpha a}$$

2b. (optional) Predict usage with knowledge embeddings:

$$W = W_{W_{ik} CCA(W_{W_{ik}}, W_{T_{wi}})_{left}}$$

3. Run CCA between $X$ and $Y$ with regularization $k$:

$$\max_{\phi_x, \phi_y} \frac{\phi_x^T C_{xy} \phi_y}{\sqrt{\phi_x^T (C_{xx} + k \sigma I) \phi_x} \sqrt{\phi_y^T C_{yy} \phi_y}}.$$  

$$\frac{1}{e^{-\lambda + 1}} \propto \|\phi_y^{proj}\|; \quad \lambda = 2b \left( \ln \left( \frac{P_i}{P_r} \right) - c \right)$$

if $\ln \left( \frac{P_i}{P_r} \right) < 0$ then $\frac{1}{e^{-\lambda + 1}} \to 0$

max $\left( \phi_y^{proj} \right) \propto \ln \left( \frac{P_i^b}{P_r} + 1 \right) \to \text{rectifier}$

4. Correct for pivots $\phi_x^{fin}$:

$$\phi_x^{proj} = \phi_x^{fin}$$

$$\|\phi_y^{proj}\| = \frac{\phi_x^{fin}}{\|\phi_y^{proj}\| + 1}$$

5. Apply projections to term-document matrix $M$:

$$M \phi_x^{fin}$$
Figure: Pivot Scores
Figure: Model: LDA, fit to 12,191 open-ended survey responses in 12 surveys.
Figure: Model: LDA, fit to 12,191 open-ended survey responses in 12 surveys.
**Granger Tests**

**Table:** This table reports the results of Granger tests when examining the relationship between the “Medicare,” “Coverage,” and “Taxes” frames and the opinions of relevant sub-groups.

<table>
<thead>
<tr>
<th></th>
<th>F Statistic</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medicare, GOP-Dem</td>
<td>0.367</td>
<td>0.553</td>
</tr>
<tr>
<td>Medicare, GOP</td>
<td>0.235</td>
<td>0.635</td>
</tr>
<tr>
<td>Coverage, Dem-GOP</td>
<td>0.349</td>
<td>0.563</td>
</tr>
<tr>
<td>Coverage, Dem</td>
<td>1.818</td>
<td>0.196</td>
</tr>
<tr>
<td>Taxes, Dem-GOP</td>
<td>0.697</td>
<td>0.416</td>
</tr>
<tr>
<td>Taxes, GOP</td>
<td>0.071</td>
<td>0.793</td>
</tr>
</tbody>
</table>
### Table: Hand-coding

This table details how each open-ended response in select surveys was hand-coded. Cells indicate percentage points.

<table>
<thead>
<tr>
<th>Name</th>
<th>July '09</th>
<th>Nov. '09</th>
<th>Oct. '10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access (+)</td>
<td>17</td>
<td>16</td>
<td>11</td>
</tr>
<tr>
<td>Gov’t Cost (-)</td>
<td>13</td>
<td>14</td>
<td>8</td>
</tr>
<tr>
<td>General (+)</td>
<td>10</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>Gov’t Role (-)</td>
<td>8</td>
<td>11</td>
<td>6</td>
</tr>
<tr>
<td>Affordability (+)</td>
<td>7</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>General (-)</td>
<td>6</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>Process (-)</td>
<td>5</td>
<td>2</td>
<td>4</td>
</tr>
</tbody>
</table>
Variational inference:

- Introduce variational distribution $q(\Theta)$ with independence assumed between blocks; factorized distribution.
- Minimize KL divergence between approximating distribution, posterior of interest (via iteratively averaging over other parameters using the approximating distribution).
- Understates posterior variability.
Variational inference:

- Deterministic optimization algorithm (e.g. Jordan et al. 1999; Grimmer 2011)

Introduce variational distribution $q(\Theta)$ with independence assumed between blocks; factorized distribution.

Minimize KL divergence between approximating distribution, posterior of interest (via iteratively averaging over other parameters using the approximating distribution)

Understates posterior variability
Variational inference:

- Deterministic optimization algorithm (e.g. Jordan et al. 1999; Grimmer 2011)
- Introduce variational distribution $q(\Theta)$ with independence assumed between blocks; factorized distribution
Variational inference:

- Deterministic optimization algorithm (e.g. Jordan et al. 1999; Grimmer 2011)
- Introduce variational distribution $q(\Theta)$ with independence assumed between blocks; factorized distribution
- Minimize KL divergence between approximating distribution, posterior of interest (via iteratively averaging over other parameters using the approximating distribution)
Inference: CTM

- Variational inference:
  - Deterministic optimization algorithm (e.g. Jordan et al. 1999; Grimmer 2011)
  - Introduce variational distribution $q(\Theta)$ with independence assumed between blocks; factorized distribution
  - Minimize KL divergence between approximating distribution, posterior of interest (via iteratively averaging over other parameters using the approximating distribution)

- Understates posterior variability
Blei and Lafferty develop Variational Expectation-Maximization procedure to estimate CTM
Blei and Lafferty develop Variational Expectation-Maximization procedure to estimate CTM

Variational Expectation Maximization:

- **E Step** – Maximize probability bound on posterior w.r.t. variational parameters; perform variational inference using coordinate ascent on each document
- **M Step** – Maximize bound on posterior w.r.t. model parameters

CTM: tolerance parameters $\leq 10^{-5}$
Blei and Lafferty develop Variational Expectation-Maximization procedure to estimate CTM

Variational Expectation Maximization:
- E Step – Maximize probability bound on posterior w.r.t. variational parameters; perform variational inference using coordinate ascent on each document
Blei and Lafferty develop Variational Expectation-Maximization procedure to estimate CTM

Variational Expectation Maximization:
- E Step – Maximize probability bound on posterior w.r.t. variational parameters; perform variational inference using coordinate ascent on each document
- M Step – Maximize bound on posterior w.r.t. model parameters
Blei and Lafferty develop Variational Expectation-Maximization procedure to estimate CTM

Variational Expectation Maximization:
- E Step – Maximize probability bound on posterior w.r.t. variational parameters; perform variational inference using coordinate ascent on each document
- M Step – Maximize bound on posterior w.r.t. model parameters

CTM: tolerance parameters $\leq 10^{-5}$
LDA: 30,000 burn-in iterations; 200,000 iterations; thinning: 1/1000, \( \alpha = \frac{50}{k} = 4.1667 \)
LDA: 30,000 burn-in iterations; 200,000 iterations; thinning: 1/1000, $\alpha = 50/k = 4.1667$

Partially collapsed Gibbs sampler: draw some steps from conditional distribution of a marginal distribution of the joint posterior
Inference: Gibbs sampler

- LDA: 30,000 burn-in iterations; 200,000 iterations; thinning: 1/1000, $\alpha = 50/k = 4.1667$
- Partially collapsed Gibbs sampler: draw some steps from conditional distribution of a marginal distribution of the joint posterior
- Marginalize over $K \times V$ word parameters
Assessing Convergence

Figure: This figure shows the cumulative average as the number of iterations increases for each of the 12 topics.
Figure: Variation in Topics Over Time. Model: CTM, fit to 1,488 press releases.
Figure: Use of emotional and threatening words in press releases
Figure: Coefficients predicting belief that health care reform will help country as a whole compared to 2004/8 exit polls.
Figure: Distances (left, middle) or correlations (right) between words used in Republican or Democratic Senators’ press releases and survey responses.