Supervised Topic Models

Advanced Machine Learning for NLP
Jordan Boyd-Graber
HIERARCHIES
Motivation: Representing Elected Officials with Ideal Points

An essential tool in political science: distinguish trends and characterize subgroups
Evaluation: Tea Party in the House

The Tea Party

- American political movement for freedom, small government, lower tax
- Disrupting Republican Party and recent elections
- Organizations:
  - Institutional: Tea Party Caucus
  - Other: Tea Party Express, Tea Party Patriots, Freedom Works
- “Conventional views of ideology as a single-dimensional, left–right spectrum experience great difficulty in understanding or explaining the Tea Party.”

Goal

- Explain Tea Partiers in terms of issues and votes
- Identify Tea Partiers from their rhetoric
Not everyone has a voting record

- Ideal points estimated based on voting record
- Not all candidates have a voting record
  - Governors
  - Entertainers
  - CEOs
Not everyone has a voting record

• Ideal points estimated based on voting record
• Not all candidates have a voting record
  ◦ Governors
  ◦ Entertainers
  ◦ CEOs
• But all politicians—by definition—talk
Let’s use whatever data we have

A single model that uses:

- Bill text
- Votes
- Commentary

to map political actors to the same continuous space.
Let’s use whatever data we have

A single model that uses:
- Bill text
- Votes
- Commentary

to map political actors to the same continuous space. This work: congressional floor speeches
Outline

1. Ideal Point Review
2. Hierarchical Ideal Point Topic Model
3. Predicting Membership
4. How They Vote
5. How They Talk
One-dimensional Ideal Point using Votes

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>YEA</td>
<td>NAY</td>
<td>NAY</td>
<td>YEA</td>
</tr>
<tr>
<td>NAY</td>
<td>YEA</td>
<td></td>
<td>YEA</td>
</tr>
<tr>
<td>YEA</td>
<td>YEA</td>
<td></td>
<td>NAY</td>
</tr>
</tbody>
</table>
One-dimensional Ideal Point using Votes

Legislator a votes 'Yea' on bill b with probability

\[ p(v_{a,b} = \text{Yea}) = \Phi(u_a x_b + y_b) \]

\[ \Phi(\alpha) = \frac{\exp(\alpha)}{\exp(\alpha) + 1} \]
### One-dimensional Ideal Point using Votes

Legislator a votes 'Yea' on bill b with probability

$$p(v_{a,b} = \text{Yea}) = \Phi(u_a x_b + y_b)$$

- **One-dimensional ideal point** of legislator a
- **Polarity** of bill b
- **Popularity** of bill b
### One-dimensional Ideal Point using Votes

#### Legislator a votes 'Yea' on bill b with probability

![Table showing votes for YEA and NAY]

<table>
<thead>
<tr>
<th></th>
<th>YEA</th>
<th>NAY</th>
<th>NAY</th>
<th>YEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAY</td>
<td>YEA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>YEA</td>
<td></td>
<td>YEA</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[ p(v_{a,b} = \text{Yea}) = \Phi(u_a x_b + y_b) \]

- **One-dimensional ideal point** of legislator a
- **Polarity** of bill b
- **Popularity** of bill b
One-dimensional Ideal Point using Votes

Legislator a votes 'Yea' on bill b with probability

\[ p(v_{a,b} = \text{Yea}) = \Phi(u_a x_b + y_b) \]

One-dimensional ideal point of legislator a

Polarity of bill b

Popularity of bill b
Multi-dimensional Ideal Point using Votes

**Legislator a votes 'Yea' on bill b with probability**

\[ p(v_{a,b} = \text{Yea}) = \Phi \left( \sum_{k=1}^{K} u_{a,k} x_{b,k} + y_b \right) \]
Multi-dimensional Ideal Point using Votes

Legislator a votes 'Yea' on bill b with probability

\[ p(v_{a,b} = \text{Yea}) = \Phi \left( \sum_{k=1}^{K} u_{a,k} x_{b,k} + y_b \right) \]

Multi-dimensional ideal point of legislator a

K dimensions
Multi-dimensional Ideal Point using Votes

Legislator a votes 'Yea' on bill b with probability

\[ p(v_{a,b} = \text{Yea}) = \Phi \left( \sum_{k=1}^{K} u_{a,k} x_{b,k} + y_b \right) \]

\( K \) dimensions

<table>
<thead>
<tr>
<th>YEA</th>
<th>NAY</th>
<th>NAY</th>
<th>YEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>YAY</td>
<td>YEA</td>
<td>YEA</td>
<td></td>
</tr>
<tr>
<td>YEA</td>
<td>YEA</td>
<td>NAY</td>
<td></td>
</tr>
</tbody>
</table>

Multi-dimensional ideal point of legislator a
Multi-dimensional Ideal Point using Votes

Legislator a votes 'Yea' on bill b with probability

\[
p(v_{a,b} = \text{Yea}) = \Phi \left( \sum_{k=1}^{K} u_{a,k} x_{b,k} + y_{b} \right)
\]

Multi-dimensional ideal point of legislator a

K dimensions

Dimensions are difficult to interpret
Multi-dimensional Ideal Point using Votes & Text

Bill Text

<table>
<thead>
<tr>
<th>YEA</th>
<th>NAY</th>
<th>NAY</th>
<th>YEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAY</td>
<td>YEA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YEA</td>
<td>YEA</td>
<td>NAY</td>
<td></td>
</tr>
</tbody>
</table>

Dimensions are difficult to interpret
Multi-dimensional Ideal Point using Votes & Text

Legislator a votes 'Yea' on bill b with probability

\[ p(v_{a,b} = \text{Yea}) = \Phi \left( x_b \sum_{k=1}^{K} u_{a,k} \vartheta_{b,k} + y_b \right) \]
Multi-dimensional Ideal Point using Votes & Text

Legislator a votes 'Yea' on bill b with probability

\[ p(v_{a,b} = \text{Yea}) = \Phi \left( x_b \sum_{k=1}^{K} u_{a,k} \vartheta_{b,k} + y_b \right) \]

Multi-dimensional ideal point of legislator a

Topic proportion of bill b estimated from its text

Dimensions are difficult to interpret
Multi-dimensional Ideal Point using Votes & Text

**Legislator a votes 'Yea' on bill b with probability**

\[
p(v_{a,b} = \text{Yea}) = \Phi \left( x_b \sum_{k=1}^{K} u_{a,k} \theta_{b,k} + y_b \right)
\]

Multi-dimensional ideal point of legislator a

Topic proportion of bill b estimated from its text

- **ObamaCare, Patient, Doctor, Insurance, Affordable Care, Hospital**
- **Balance Budget, Debt Ceiling, Cap, Cut Spend, Raise Tax**
- **Debt Limit, Nation Debt**
- **Employ, Hire, Job Creator, Union, NLRB, BOE, Labor, Business Owner**
Outline

1. Ideal Point Review
2. Hierarchical Ideal Point Topic Model
3. Predicting Membership
4. How They Vote
5. How They Talk
What are your thoughts on the issue of immigration?

- path to citizenship
- self-deportation
- $*/#7!
Hierarchical Ideal Point Topic Model: Overview

Using both votes and text to learn

- Two-level topic hierarchy:
- Ideal points in multiple interpretable dimensions
Hierarchical Ideal Point Topic Model: Overview

Using both votes and text to learn

- **Two-level topic hierarchy:**
  - First-level nodes map to agenda issues
  - Second-level nodes map to issue-specific frames
- Ideal points in multiple interpretable dimensions

![Diagram of topic hierarchy](image-url)

- **Topic Hierarchy**
  - First-level nodes ~ Agenda issues
  - Second-level nodes ~ Issue-specific frames

Hierarchical Ideal Point Topic Model
Hierarchical Ideal Point Topic Model: Overview

Using both votes and text to learn

- Two-level topic hierarchy: Use existing labeled data to learn priors for interpretable issues
- Ideal points in multiple interpretable dimensions

Use prior to learn interpretable issue topics
Hierarchical Ideal Point Topic Model: Overview

Using both votes and text to learn

- Two-level topic hierarchy: Ideal points for frames for predictions using text only
- Ideal points in multiple interpretable dimensions
Hierarchical Ideal Point Topic Model: Overview

Using both votes and text to learn

- Two-level topic hierarchy:
- Ideal points in multiple interpretable dimensions

Multi-dimensional Ideal Points
Hierarchical Ideal Point Topic Model: Inputs

- A collection of votes \( \{v_{a,b}\} \)
- A collection of \( D \) speeches \( \{w_d\} \), each of which is given by legislator \( a_d \)
- A collection of \( B \) bill text \( \{w'_b\} \)
Hierarchical Ideal Point Topic Model

Modeling bill text

- Each bill text $b$ is a mixture over $K$ issues $\theta_b$
- Each bill token generated from topic at first-level issue node

**Health**
- obamacare, patient, doctor, physician, afford_care, hospit, insur, replac, mandat, exchang, health_insur, coverag, medicaid, patient_protect, board

**Macroeconomics**
- balanc_budget, borrow, debt_cell, cap, cut_spend, nation_debt, grandchildren, rais_tax, entitl, white_hous, debt_limit, prosper

**Frame H1**
- afford_care, exchang, patient_protect, human_servic, public_health, slush_fund, ppaca, mandatior,mandatior_spend, governor, hospit, health_center, flexibi, teach_health,

**Frame H2**
- patient, doctor, physician, hospit, medicaid, board, georgia, save_medicar, nurs, tennessee, page, bureaucrat, advesori_board, medicin, independ_payment

**Frame H3**
- obamacare, replac, mandat, insur, health_insur, coverag, social_secur, premium, repeal_obamacar, entitl, govern_takeov, purchas, unconstitut, preexist_condit

**Frame M1**
- white_hous, shut, continu_resolut, mess, hous_republican, novemb, govern_shutdown, senat_reid, harri_reid, vision, shutdown, liber, arriv, republican_parti, blame

**Frame M2**
- balanc_budget, debt_cell, cap, cut_spend, debt_limit, spend_cut, fiscal_hous, grandchild, guarente, default, august, obama, deficit_spend, rein, feder_budget

**Frame M3**
- borrow, nation_debt, rais_tax, entitl, prosper, chart, grandchildren, spend_monei, size, gdp, tax_increas, cent, govern_spend, social_secur
Hierarchical Ideal Point Topic Model: Generative Process

- Each speech $d$ also has a distribution $\theta_d$ over $K$ issues.
- Each issue $k$, each speech $d$ has distribution over frames $\psi_{d,k}$.
- Each speech token from topic at second-level frame node.
Hierarchical Ideal Point Topic Model

Hierarchical Ideal Point Topic Model: Modeling votes

- **Legislator** $a$ votes ‘Yea’ on bill $b$ with probability
  \[ p(v_{a,b} = \text{Yea}) = \Phi(x_b \sum_{k=1}^{K} \theta_{b,k} u_{a,k} + y_b) \]

- **Ideal point** $u_{a,k} \sim \mathcal{N}\left(\sum_{j=1}^{J} \eta_{k,j} \psi_{a,k,j} + \rho\right)$

---

**Health**
- obamacar, patient, doctor, physician, afford_care, hospit, insur, replac, mandat, exchang, health_insur, coverag, medicaid, patient_protect, board

**Frame H1**
- afford_care, exchang, patient_protect, human_servic, public_health, slush_fund, ppaca, mandatori, mandatori_spend, governor, hospit, health_center, flexibl, teach_health

**Frame H2**
- patient, doctor, physician, hospit, medicaid, board, georgia, save_medicar, nurs, tennesse, page, bureaucrat, advisori_board, medicin, independ_payment

**Frame H3**
- obamacar, replac, mandat, insur, health_insur, coverag, social_secur, premium, repeal_obamacar, entitl, govern_takeov, purchas, unconstitu, preexist_condit

**Macroeconomics**
- balanc_budget, borrow, debt_cell, cap, cut_spend, nation_debt, grandchildren, rais_tax, entitl, white_hous, debt_limit, prosper

**Frame M1**
- white_hous, shut, continu_resolut, mess, hous_republican, novemb, govern_shutdown, senat_reid, harri_reid, vision, shutdown, liber, arriv, republican parti, blame

**Frame M2**
- balanc_budget, debt_cell, cap, cut_spend, debt_limit, spend_cut, fiscal_hous, grandchildren, guarante, default, august, obama, deficit_spend, rein, feder_budget

**Frame M3**
- borrow, nation_debt, rais_tax, entitl, prosper, chart, grandchildren, spend_monei, size, gdp, tax_increas, cent, govern_spend, social_secur
Hierarchical Ideal Point Topic Model: Modeling votes

- Legislator $a$ votes ‘Yea’ on bill $b$ with probability
  $$p(v_{a,b} = \text{Yea}) = \Phi(x_b \sum_{k=1}^{K} \theta_{b,k} u_{a,k} + y_b)$$

- Ideal point $u_{a,k} \sim \mathcal{N}(\sum_{j=1}^{J} \eta_{k,j} \psi_{a,k,j}, \rho)$

**Health**
- obamacare, patient, doctor, physician, afford_care, hospit, insur, replac, mandat, exchang, health_insur, coverag, medicaid, patient_protect, board

**Macroeconomics**
- balanc_budget, borrow, debt_cell, cap, cut_spend, nation_debt, grandchildren, rais_tax, entitl, white_hous, debt_limit, prosper

Advanced Machine Learning for NLP | Boyd-Graber
 Supervised Topic Models | 13 of 29
Hierarchical Ideal Point Topic Model: Modeling votes

- Legislator \( a \) votes ‘Yea’ on bill \( b \) with probability
  \[
p(v_{a,b} = \text{Yea}) = \Phi(x_b \sum_{k=1}^{K} \vartheta_{b,k} u_{a,k} + y_b)
  \]

- Ideal point \( u_{a,k} \sim \mathcal{N}(\sum_{j=1}^{J} \eta_{k,j} \psi_{a,k,j}, \rho)
  \)
Hierarchical Ideal Point Topic Model: Modeling votes

- Legislator $a$ votes ‘Yea’ on bill $b$ with probability
  \[ p(v_{a,b} = \text{Yea}) = \Phi(x_b \sum_{k=1}^{K} \vartheta_{b,k} u_{a,k} + y_b) \]

- Ideal point $u_{a,k} \sim \mathcal{N}(\sum_{j=1}^{J_k} \eta_{k,j} \psi_{a,k,j}, \rho)$
Hierarchical Ideal Point Topic Model: Modeling votes

- Legislator $a$ votes ‘Yea’ on bill $b$ with probability
  \[ p(v_{a,b} = \text{Yea}) = \Phi(x_b \sum_{k=1}^{K} \theta_{b,k} u_{a,k} + y_b) \]

- Ideal point $u_{a,k} \sim \mathcal{N}\left(\sum_{j=1}^{J} \eta_{k,j} \psi_{a,k,j} \rho\right)$
Nested Chinese Restaurant Process

- Start at a CRP, choose a table
- That table has not just a dish (distribution over words) but also business card
- That card tells you which restaurant to go to next
- You do this $L$ times
Topic Hierarchies
Hierarchical Ideal Point Topic Model

Generative Process

\[ \Phi_{1,1} \]

Start at the root node
Generative Process

Need to choose which table to sit at
Generative Process

This is a CRP! (Can create new table too.)
Hierarchical Ideal Point Topic Model

Generative Process

Repeat
Your path then becomes the set of topics you use for this document
Warning: Probably don’t want to only use one path per document (but useful explanation)
Evaluation: Tea Party in the House

The Tea Party

- American political movement for freedom, small government, lower tax
- Disrupting Republican Party and recent elections
Evaluation: Tea Party in the House

The Tea Party

• American political movement for freedom, small government, lower tax
• Disrupting Republican Party and recent elections

Data

• 240 Republican Representatives in the 112th U.S. House
• 60 are members of the Tea Party Caucus (self-identified)
• 60 key votes selected by Freedom Works (2011-2012)
• Speeches, bill text and voting records from the Library of Congress
Outline

1. Ideal Point Review
2. Hierarchical Ideal Point Topic Model
3. Predicting Membership
4. How They Vote
5. How They Talk
### Tea Party Caucus Membership Prediction

#### Experiment setup

- **Task:** Binary classification of whether a legislator is a member of the Tea Party Caucus
- **Evaluation metric:** AUC-ROC
- **Classifier:** SVM
- **Five-fold stratified cross-validation**

**Features**

- **Text-based features:** normalized term frequency (TF) and TF-IDF
- **Vote:** binary features
- **HIPTM:** features extracted from our model including
  - $K$-dim ideal point $\mathbf{u}_a, k$ estimated from both votes and text
  - $K$-dim ideal point estimated from text only $\mathbf{\eta}_k^\psi$
  - $B$ probabilities estimating $a$'s votes $\Phi(x_b \sum_{k=1}^{K} \vartheta_b^k, k \mathbf{u}_a, k + y_b)$
Predicting Membership

**Tea Party Caucus Membership Prediction**

**Experiment setup**
- **Task**: Binary classification of whether a legislator is a member of the Tea Party Caucus
- **Evaluation metric**: AUC-ROC
- **Classifier**: SVM\(^{light}\)
- **Five-fold stratified cross-validation**

**Features**
- **Text-based features**: normalized term frequency (TF) and TF-IDF
- **Vote**: binary features
- **HIPTM**: features extracted from our model including
  - \(K\)-dim ideal point \(u_{a,k}\) estimated from both votes and text
  - \(K\)-dim ideal point estimated from text only \(\eta_k^T \hat{\psi}_{a,k}\)
  - \(B\) probabilities estimating \(a\)’s votes \(\Phi(x_b \sum_{k=1}^{K} \theta_{b,k} u_{a,k} + y_b)\)
Tea Party Caucus Membership Prediction: Votes & Text

AUCROC

0.60
0.65
0.70
0.75

TF
TFIDF
Vote
HIPTM
Vote-TF
Vote-TF-IDF
Vote-HIPTM
All
Tea Party Caucus Membership Prediction: Votes & Text

AUCROC

0.60
0.65
0.70
0.75

TF
TFIDF
Vote
HIPTM
Vote-TF
Vote-TF-IDF
Vote-HIPTM
All

Text-based Features
Tea Party Caucus Membership Prediction: Votes & Text

AUCROC

0.60
0.65
0.70
0.75

TF
TFIDF
Vote
HIPTM
Vote-TF
Vote-TF-IDF
Vote-HIPTM
All

Text-based Features

Vote Features
Tea Party Caucus Membership Prediction: Votes & Text

AUCROC

0.75

0.70

0.65

0.60

TF  TFIDF  Vote  HIPTM  Vote-TF  Vote-TF-IDF  Vote-HIPTM  All

Text-based Features  Vote Features  Our Features
Tea Party Caucus Membership Prediction: Votes & Text

<table>
<thead>
<tr>
<th>Features</th>
<th>AUCROC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Text-based</td>
<td></td>
</tr>
<tr>
<td>Vote</td>
<td></td>
</tr>
<tr>
<td>Our Features</td>
<td></td>
</tr>
<tr>
<td>Combining Vote with TF/TF-IDF</td>
<td></td>
</tr>
</tbody>
</table>

The graph shows the AUC-ROC for different feature combinations:

- **TF**
- **TFIDF**
- **Vote**
- **HIPTM**
- **Vote-TF**
- **Vote-TF-IDF**
- **Vote-HIPTM**
- **All**

The graph indicates that combining vote features with TF/TF-IDF provides the best performance, as evidenced by the highest AUC-ROC value.
Tea Party Caucus Membership Prediction: Votes & Text

- Predicting Membership
- AUCROC
  - 0.60
  - 0.65
  - 0.70
  - 0.75
- TF
- TFIDF
- Vote
- HIPTM
- Vote-TF
- Vote-TF-IDF
- Vote-HIPTM
- All

- Text-based Features
- Vote Features
- Our Features
- Combining Vote with TF/TF-IDF
- Combining Vote with Our Features
Tea Party Caucus Membership Prediction: Text Only

<table>
<thead>
<tr>
<th>Method</th>
<th>AUCROC</th>
</tr>
</thead>
<tbody>
<tr>
<td>TF</td>
<td>0.61</td>
</tr>
<tr>
<td>TF-IDF</td>
<td>0.63</td>
</tr>
<tr>
<td>HIPTM</td>
<td>0.65</td>
</tr>
</tbody>
</table>
Tea Party Caucus Membership Prediction: Text Only

<table>
<thead>
<tr>
<th>Method</th>
<th>Training</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>TF</td>
<td>text only</td>
<td>text only</td>
</tr>
<tr>
<td>TF-IDF</td>
<td>text only</td>
<td>text only</td>
</tr>
<tr>
<td>HIPTM</td>
<td>text and votes</td>
<td>text only</td>
</tr>
</tbody>
</table>

AUCROC

0.61
0.63
0.65
Outline

1 Ideal Point Review
2 Hierarchical Ideal Point Topic Model
3 Predicting Membership
4 How They Vote
5 How They Talk
One-dimensional Ideal Points

- Judy Biggert
- Ander Crenshaw
- Jim Gerlach
- Michael Grimm
- Christopher H. Smith
- Harold Rogers
- Peter T. King
- Patrick Meehan
- Rodney M. Alexander
- Bob Dold
- Steve Stivers
- Jon Runyan
- Ileana Ros-Lehtinen
- Dave G. Reichert
- Mario Diaz-Balart

- Trent Franks
- Trey Gowdy
- Tom McClintock
- David Schweikert
- Doug Lamborn
- Scott Garrett
- Tom McClintock
- David Schweikert
- Trey Gowdy
- Trent Franks
- Paul C. Broun
- Joe Walsh
- Mick Mulvaney
- Justin Amash
- Jeff Flake
- Jeff Duncan
- Raúl Labrador
- Jim Jordan
- Tom Graves
- Justin Amash
- Mick Mulvaney
- Paul C. Broun
- Joe Walsh
- Scott Garrett
- Doug Lamborn
- Tom McClintock
- David Schweikert
- Trey Gowdy
- Trent Franks
- Paul C. Broun
- Joe Walsh
- Scott Garrett
- Doug Lamborn
- Tom McClintock
- David Schweikert
- Trey Gowdy
- Trent Franks
One-dimensional Ideal Points

Tea Party Caucus a Member a Nonmember

Advanced Machine Learning for NLP | Boyd-Graber
One-dimensional Ideal Points

- **Alexander** and **Crenshaw**'s votes only agree with Freedom Works 48% and 50% respectively.
- Both voted for raising the debt ceiling and are listed as “traitor”.

<table>
<thead>
<tr>
<th>House Tea Party Caucus members</th>
<th>how they voted on debt ceiling increase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sandy Adams, Florida</td>
<td>traitor</td>
</tr>
<tr>
<td>Robert Aderholt, Alabama</td>
<td>traitor</td>
</tr>
<tr>
<td>Todd Akin, Missouri</td>
<td>no</td>
</tr>
<tr>
<td>Rodney Alexander, Louisiana</td>
<td>traitor</td>
</tr>
<tr>
<td>Michele Bachmann, Minnesota, Chairman</td>
<td>no</td>
</tr>
<tr>
<td>Rob Bishop, Utah</td>
<td>no</td>
</tr>
<tr>
<td>Ander Crenshaw, Florida</td>
<td>traitor</td>
</tr>
<tr>
<td>Michael C. Burgess, Texas</td>
<td>traitor</td>
</tr>
</tbody>
</table>

**John T. Reed on Headline News**
points and perspectives not offered elsewhere
• **Flake** and **Amash** didn’t self-identify as members of the Tea Party Caucus but have been endorsed by other Tea Party organizations.

**NEW REPUBLIC**

“Some 46 House members and six senators had been [Tea Party] … In addition, there were about 18 other House members like Trey Gowdy, Mark Meadows, and **Justin Amash**, and several senators, including **Jeff Flake** and Pat Toomey, who owed their election to support from the Tea Party and its Washington allies.”
Freedom Works’ key votes on most highly polarized dimensions are about government spending.
Outline

1 Ideal Point Review

2 Hierarchical Ideal Point Topic Model

3 Predicting Membership

4 How They Vote

5 How They Talk
Framing Healthcare

Health
- obamacare, patient, doctor, physician, afford_care, hospit, insur, replac, mandat, exchang, health_insur, coverag, medicaid, patient_protect, board

Frame H1
- afford_care, exchang, patient_protect, human_servic, public_health, slush_fund, ppaca, mandatori, mandatori_spend, governor, hospit, health_center, flexibl, teach_health, unlimit

Frame H2
- patient, doctor, physician, hospit, medicaid, board, georgia, save_medicar, nurs, tennesse, page, bureaucrat, advisori_board, medicin, independ_payment

Frame H2
- obamacare, replac, mandat, insur, health_insur, coverag, social_secur, premium, repeal_obamacare, entitl, govern_takeov, purchas, unconstit, preexist_condit, employ

-1.13 0.04 0.56
Framing Macroeconomics

Macroeconomics
- balanc_budget, borrow, debt_ceil, cap, cut_spend,
  nation_debt, grandchildren, rais_tax, entitl

Frame M1
- white_hous, shut, continu_resolut, mess,
  hous_republican, novemb,
  govern_shutdown

Frame M2
- balanc_budget, debt_ceil, cap, cut_spend, debt_limit,
  spend_cut, fiscal_hous, grandchildren, guarante

Frame M2
- borrow, nation_debt, rais_tax, entitl, prosper,
  chart, grandchildren, spend_monei, size, gdp

-.57 -.24 .56
## Polarization

<table>
<thead>
<tr>
<th>Ideal Point Distributions</th>
<th>Not</th>
<th>Polarized</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distribution of Issue Frames</td>
<td>Civil Rights, Minority Issues, Civil Liberties</td>
<td>Banking and Finance; Transportation</td>
</tr>
<tr>
<td>Polarized</td>
<td>Health; Public Lands and Water Management</td>
<td>Macroeconomics; Government Operations</td>
</tr>
</tbody>
</table>

---

**How They Talk**

**Polarization**

**Ideal Point Distributions**

- Not
- Polarized

**Distribution of Issue Frames**

- Civil Rights, Minority Issues, Civil Liberties
- Health; Public Lands and Water Management
- Banking and Finance; Transportation
- Macroeconomics; Government Operations

---

**Advanced Machine Learning for NLP**

| Supervised Topic Models | Boyd-Graber | 28 of 29 |
### Polarization

<table>
<thead>
<tr>
<th>Ideal Point Distributions</th>
<th>Not</th>
<th>Polarized</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distribution of Issue Frames</td>
<td>Not</td>
<td>Civil Rights, Minority Issues, Civil Liberties</td>
</tr>
<tr>
<td>Polarized</td>
<td>Health; Public Lands and Water Management</td>
<td>Macroeconomics; Government Operations</td>
</tr>
</tbody>
</table>

![Diagram showing the distribution of issue frames and ideal points, with 'YES' and 'NO' boxes.]
Polarization

<table>
<thead>
<tr>
<th>Ideal Point Distributions</th>
<th>Not</th>
<th>Polarized</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distribution of Issue Frames</td>
<td>Not</td>
<td>Civil Rights, Minority Issues, Civil Liberties</td>
</tr>
<tr>
<td></td>
<td>Polarized</td>
<td>Banking and Finance Transportation</td>
</tr>
<tr>
<td></td>
<td><strong>Health; Public Lands and Water Management</strong></td>
<td>Macroeconomics; Government Operations</td>
</tr>
</tbody>
</table>

---

**How They Talk**
## Polarization

<table>
<thead>
<tr>
<th>Ideal Point Distributions</th>
<th>Not</th>
<th>Polarized</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distribution of Issue Frames</td>
<td>Not</td>
<td>Civil Rights, Minority Issues, Civil Liberties, Banking and Finance, Transportation</td>
</tr>
<tr>
<td>Polared</td>
<td>Health; Public Lands and Water Management</td>
<td>Macroeconomics; Government Operations</td>
</tr>
</tbody>
</table>

**YES** | **NO**
Hierarchies are Cool

- For a sweep of single parameter, BNP not that useful
- Complex structures are more realistic applications
- Combining with supervised objective
- Unsolved problem: good prediction with interpretable structure