Political Ideology and Polarization of Policy Positions: A Multi-dimensional Approach

Barea Sinno*[1], Bernardo Oviedo*[2], Katherine Atwell*[3], Malihe Alikhani[3], Junyi Jessy Li[4]

[1]Political Science, Rutgers University; [2]Computer Science, [4]Linguistics, The University of Texas at Austin;
[3]Computer Science, University of Pittsburgh

*denotes equal contribution

RUTGERS

TEXAS
The University of Texas at Austin

University of Pittsburgh

Motivation: A nuanced perspective on media bias

- We argue that stance in text is different from ideology
- We study **ideology** as a preference for a policy position
  - For example: “support for universal healthcare”
- We also offer measures of polarization over time

Example

The nuanced co-existence of stance and ideology can be illustrated in the following excerpt:

“Republicans and Joe Biden are making a huge mistake by focusing on cost. The implication is that government-run health care would be a good thing–a wonderful thing!– if only we could afford it." (The Federalist, 9/27/2019)

Existing Approaches to study ideology

- Often, ideology is conceptualized broadly in two classes: conservative and liberal
  - Polarization is then measured as distance between these two positions
- Fine-grained attempts add to the magnitude of these classes
  - i.e *mildly conservative*

Our Approach

- We follow the political science literature and instead increase the number of **dimensions** of ideology
    - i.e. *economic, social, and foreign*
- We measure polarization as a pairwise correlation between dimensions

Outline

1. **Data collection and annotation**
2. Measure polarization over time
3. Predicting ideology
4. Conclusions

Annotation Task

We follow political science literature and use **three different dimensions** for ideology

- **Economic**
- **Social**
- **Foreign**

A paragraph can discuss multiple dimensions

Each dimension is annotated as **conservative, liberal, or neutral**

Socially and economically liberal
Secretary of Defense Robert S. threw his full support today behind the Administration's drive against poverty. Citing figures showing that, about a third of the nation's youths fail either mental or physical examinations given by, the Selective Service, Mr. Mc-Namara said : “It is the youth that we can expect to be the most immediate beneficiaries of the war on poverty."

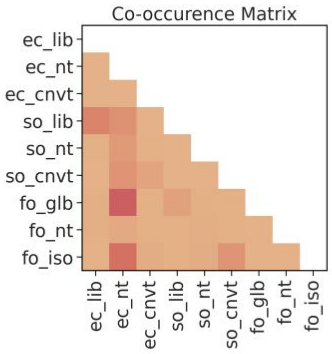
Data Collection and Annotation

- News articles on the federal budget from **1947 to 1975**
- **7.5 articles** per year
- **721 paragraphs**
- Annotators **adjudicated** their differences to create gold labels
- Total annotation time: ~ **150 hours**

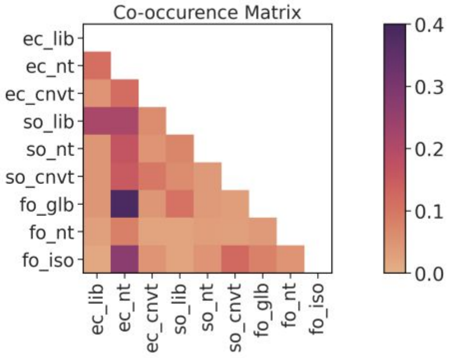
Richness of information in a multi-dimensional approach

In our annotation sample, some policy positions co-occur more across dimensions than others.

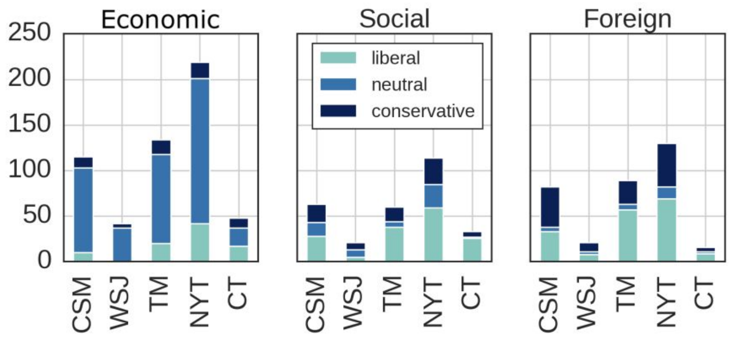
In the aggregate, the distribution of dimensions of ideology in sources is different from the AllSides media bias measurements.



Paragraph Level



Article Level



Outline

1. Data collection and annotation
2. **Measure polarization over time**
3. Predicting ideology
4. Conclusions

Case study: analyses of polarization

We can use our annotated data to **study broad trends in polarization over time**

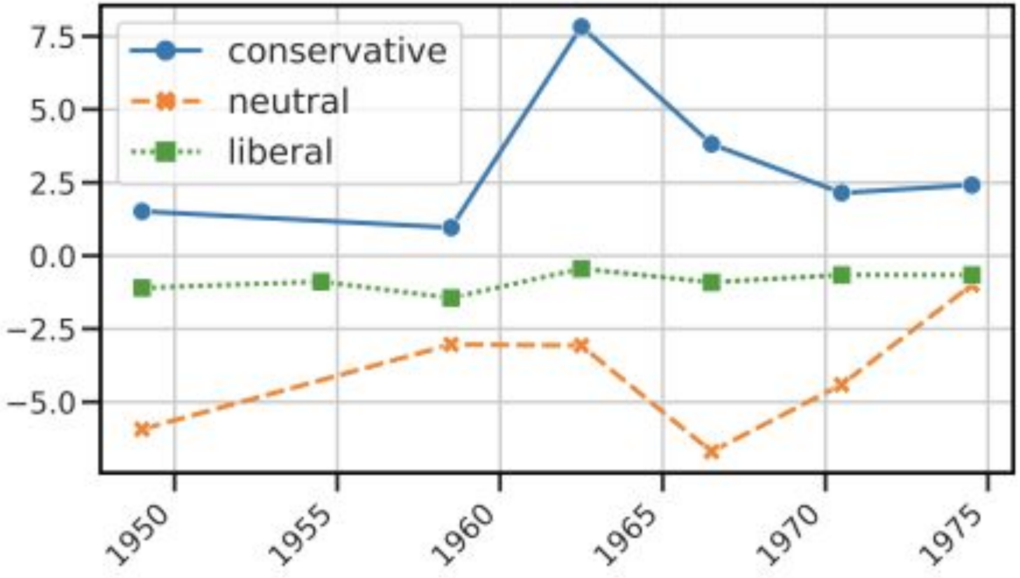
Three important metrics:

- **Sorting measure** - to what extent articles deviate from their publication's proclaimed ideology
- **Issue constraint** - how closely associated ideologies are across dimensions
- **Ideological divergence** - the distance between two ideological groups on a single dimension

Sorting Measure

Difference between **proclaimed ideological bias** of a news outlet and the **ideology of annotated articles** from the outlet

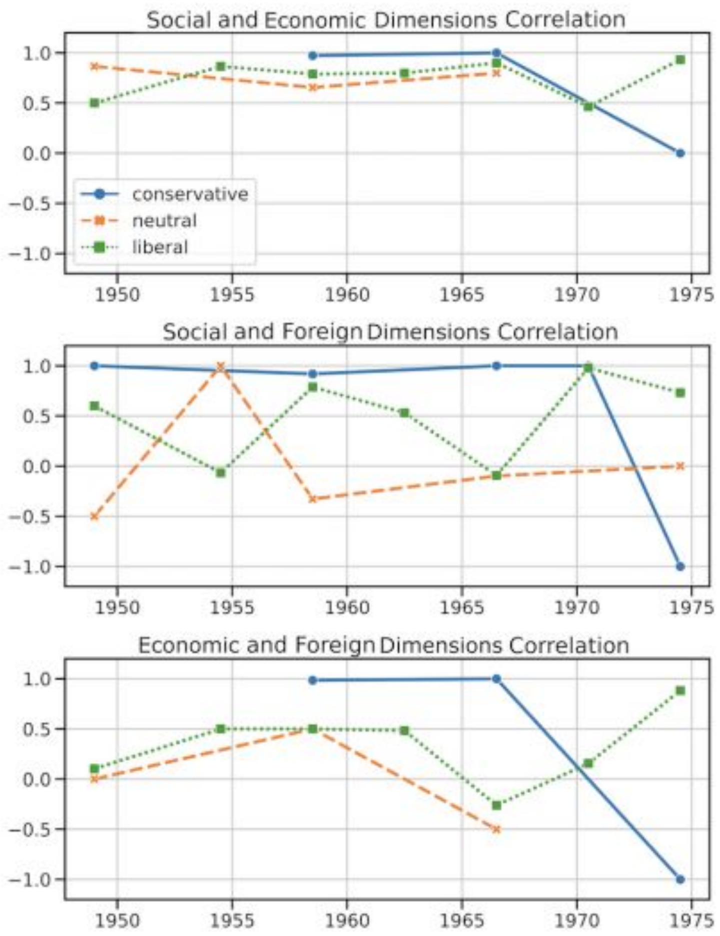
- **Left-leaning outlets** were closest to their proclaimed ideological bias measure over time
- **Neutral outlets** were more liberal before 1957 and after 1964
- **Right leaning outlets** were more conservative than their proclaimed ideological bias between 1957 and 1964



Issue Constraint

How closely trends in associated ideological bias levels are across dimensions (e.g. how likely are socially liberal articles to also be economically liberal?)

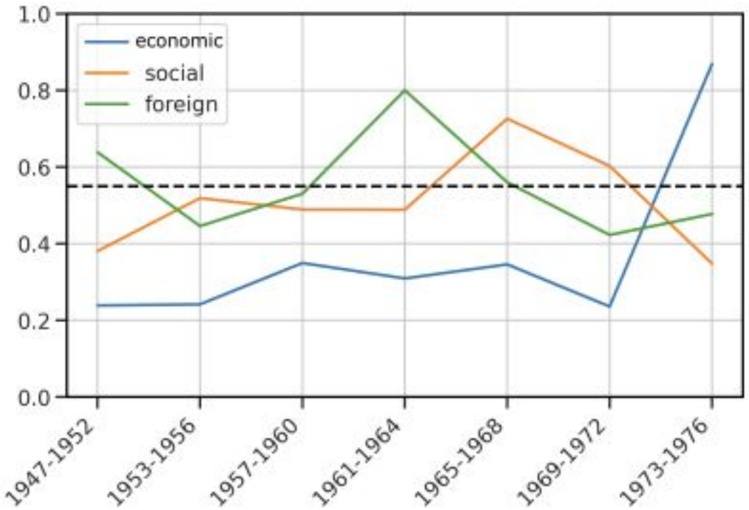
- **Left-leaning and neutral outlets** show fluctuating correlations over time
- **Right leaning outlets** show positive correlations between dimensions before 1967/1970



Ideological Divergence

Distance between two ideological groups on a single dimension

- **The foreign dimension crosses the bimodality threshold between 1956 and 1968.** This means that proclaimed left-leaning and right-leaning outlets grew further apart on foreign issues during this time period.



Outline

1. Data collection and annotation
2. Measure polarization over time
3. **Predicting ideology**
4. Conclusions

Predicting Political Ideology

Majority Class Baseline: simply predicting the majority class for each dimension

Recurrent Neural Network: 2-layer bidirectional LSTM, with sequence length and hidden size of 256, and 100D GloVe embeddings

Pre-trained language models: BERT-base model with and without fine-tuning

- Learning Rate
- # of Epochs
- Gamma
- Batch Size
- Dropout
- Frozen/Finetuned version

Main Results

- The fine-tuned BERT model, with no task-guided pre-training shows the best performance across all 3 ideology dimensions
- All models do better than the majority class baseline

	Econ	**Social**	**Foreign**	**Average**
Majority	0.30	0.23	0.25	0.26
BiLSTM	0.44	0.37	0.33	0.38
BERT (finetuned)	**0.64**	**0.50**	**0.52**	**0.55**

F-1 scores of our baseline, BiLSTM & BERT models

Ablations

- **Task-guided pre-training:** we labeled the ideology of each article following the ideology of its source in www.allsides.com
- The pre-training task generally decreased f-1 scores compared to vanilla/finetuned BERT
- This decrease seems to support our more nuanced analysis of media stance vs. ideology in text

	Econ	Social	Foreign	Average
BERT	0.46	0.31	0.53	0.44
+pre-training	0.42	0.32	0.46	0.40
BERT (finetuned)	**0.64**	**0.50**	**0.52**	**0.55**
+pre-training	0.56	0.47	0.46	0.49

F-1 scores of our BERT models with/without pretraining

	Econ	**Social**	**Foreign**	**Average**
Majority	0.30	0.23	0.25	0.26
BiLSTM	0.44	0.37	0.33	0.38
BERT	0.46	0.31	0.53	0.44
+pre-training	0.42	0.32	0.46	0.40
BERT (finetuned)	**0.64**	**0.50**	**0.52**	**0.55**
+pre-training	0.56	0.47	0.46	0.49
-focal loss	0.61	**0.50**	0.50	0.54

F-1 scores of our BiLSTM models and BERT models

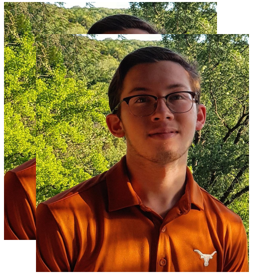
Outline

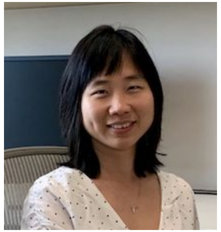
1. Data collection and annotation
2. Measure polarization over time
3. Predicting ideology
4. **Conclusions**

Conclusions and Takeaways

- We demonstrate that ideology like stance is important in text
- Our annotations show the potential of a fine-grained approach for the study of ideology and polarization
- But predicting ideology is a difficult task!

Thank you!





NSF

RUTGERS

TEXAS
The University of Texas at Austin

University of Pittsburgh

DARPA