



Northeastern University

# Auditing the Partisanship of Google Search Snippets

Desheng Hu  
Shan Jiang  
Ronald E. Robertson  
Christo Wilson

[All](#) News Images Videos Books More[Settings](#) Tools

About 639,000,000 results (0.44 seconds)

**Gun control** (or firearms regulation) is the set of laws or policies that regulate the manufacture, sale, transfer, possession, modification, or use of firearms by civilians. Most countries have a restrictive firearm guiding policy, with only a few legislations being categorized as permissive.



### Gun control - Wikipedia

[https://en.wikipedia.org/wiki/Gun\\_control](https://en.wikipedia.org/wiki/Gun_control)[About this result](#) [Feedback](#)

### People also ask

[What does it mean to be for gun control?](#)[What is the purpose of gun control?](#)[What political party is against gun control?](#)[What does the Constitution say about gun control?](#)[Feedback](#)

### Gun control - Wikipedia

[https://en.wikipedia.org/wiki/Gun\\_control](https://en.wikipedia.org/wiki/Gun_control)

Gun control (or firearms regulation) is the set of laws or policies that regulate the manufacture, sale, transfer, possession, modification, or use of firearms by civilians. Most countries have a restrictive firearm guiding policy, with only a few legislations being categorized as permissive.

[Terminology and context](#) · [Regulation of civilian ...](#) · [Studies](#) · [See also](#)

### Gun Control - ProCon.org

<https://gun-control.procon.org/>

Proponents of more gun control laws state that the Second Amendment was intended for militias; that gun violence would be reduced; that gun restrictions have ...

### Gun Control – Just Facts

<https://www.justfacts.com/guncontrol.asp>

Comprehensive and meticulously documented facts about gun control. Learn about ownership rates, crime, background checks, accidents, politics, and more.

Republican: 38% Democrat: 22%

Independent: 31% Male: 45%

[Semi-Automatic Guns](#) · [Violent Crimes](#) · [Gun Show Purchases](#) · [Fatal Accidents](#)[More images](#)

## Gun control



Gun control is the set of laws or policies that regulate the manufacture, sale, transfer, possession, modification, or use of firearms by civilians. Most countries have a restrictive firearm guiding policy, with only a few legislations being categorized as permissive. [Wikipedia](#)

### Gun control organizations

[View 20+ more](#)National  
Rifle  
AssociationEverytown  
for Gun  
SafetyBrady  
CampaignCoalition to  
Stop Gun  
ViolenceAmericans  
for  
Responsible  
Solutions

### People also search for

[View 5+ more](#)

Law

Climate  
change

Rights



Immigration



Education

[Feedback](#)

# Why Snippet Matters?

- Content highlighted may influence readers' perceptions and behaviors

**Weather**  
Today: Partly cloudy.  
High: 53 Low: 41.  
Monday: Partly cloudy.  
High: 62 Low: 45.  
Details, page B2.

**The Washington Post**

**Clinton Impeached**  
House Approves Articles Charging Perjury, Obstruction

*Livingston Quits As Designated House Speaker*

**Mostly Partisan Vote Shifts Drama to Senate**

**Clinton Vows to Finish Out Term**  
President Says He Will Keep Working to 'the Last Hour of the Last Day'

**The Dizzietest Day**  
■ The president was impeached, the vice president was resigning, and Livingston, fighting back tears, "What are we going to do? Line them all up and shoot them?"

Rep. Michael N. Castle (R-Del.) added: "This is so overwhelming. There have been so many bombshells you can hardly turn your back."

Livingston made his decision early yesterday, after a night of soul-searching and consultation with his wife of 33 years, Bonnie. It was his wife who had insisted that he leave his extramarital affair on Thursday, after learning that

See LIVINGSTON, A31 Col. 1  
■ Complete coverage of the GOP search for a new speaker, A29

**Clinton Impeached**  
House Approves Articles Charging Perjury, Obstruction

**Mostly Partisan Vote Shifts Drama to Senate**

**Clinton Vows to Finish Out Term**  
President Says He Will Keep Working to 'the Last Hour of the Last Day'

**The Dizzietest Day**  
■ The president was impeached, the vice president was resigning, and Livingston, fighting back tears, "What are we going to do? Line them all up and shoot them?"

Rep. Michael N. Castle (R-Del.) added: "This is so overwhelming. There have been so many bombshells you can hardly turn your back."

Livingston made his decision early yesterday, after a night of soul-searching and consultation with his wife of 33 years, Bonnie. It was his wife who had insisted that he leave his extramarital affair on Thursday, after learning that

See LIVINGSTON, A31 Col. 1  
■ Complete coverage of the GOP search for a new speaker, A29

**Late Edition**  
New York: Today, sunny, windy, and cool, high 42. Tonight, clear, low 24. Tomorrow, mostly cloudy and less breezy, high 35. Yesterday, high 61, low 44. Weather map, page A17.

**The New York Times**

**CLINTON ACQUITTED DECISIVELY: NO MAJORITY FOR EITHER CHARGE**

**President Says He Is Sorry And Seeks Reconciliation**

**CENSURE IS BARRED**  
But Rebuke From Both Sides of Aisle Dilutes President's Victory

**Senators, How Say You?**

Article	GUILTY		NOT GUILTY	
	REPUBLICANS	DEMOCRATS	REPUBLICANS	DEMOCRATS
PERJURY	45	10	50	5
OBSTRUCTION OF JUSTICE	0	45	0	45
<b>TOTAL</b>	<b>45</b>	<b>55</b>	<b>50</b>	<b>50</b>

**A Dispirited Hyde Opposes Indicting Clinton**

# Politics in Search Engines

- People often use search engines for news-seeking and fact-checking (*Dutton et al., 2017*)
- Political valence of search rankings can shift the preference of undecided voters by 20% or more (*Epstein & Robertson 2015*)
- Prior audits have examined the partisanship of links on search engine result pages (SERPs) but not the text (*Robertson et al., 2018*)



# Research Questions

- 1. Does the political valence of the snippets match the valence of the corresponding page content?**
- 2. Where are the snippets extracted from?**
- 3. Do the results vary by query topics?**
- 4. Do left and right-leaning queries have different results?**



# Study Overview

- First large-scale analysis of partisanship in Google Search snippets, by running 4,570 political queries and their autocomplete suggestions
- Audited the behavior of Google Search's document summarization algorithm
- Found that Google Search's snippets generally **amplify** partisanship relative to their respective webpages
- This effect is **robust** across different types of webpages, query topics, and partisan (left- and right-leaning) queries



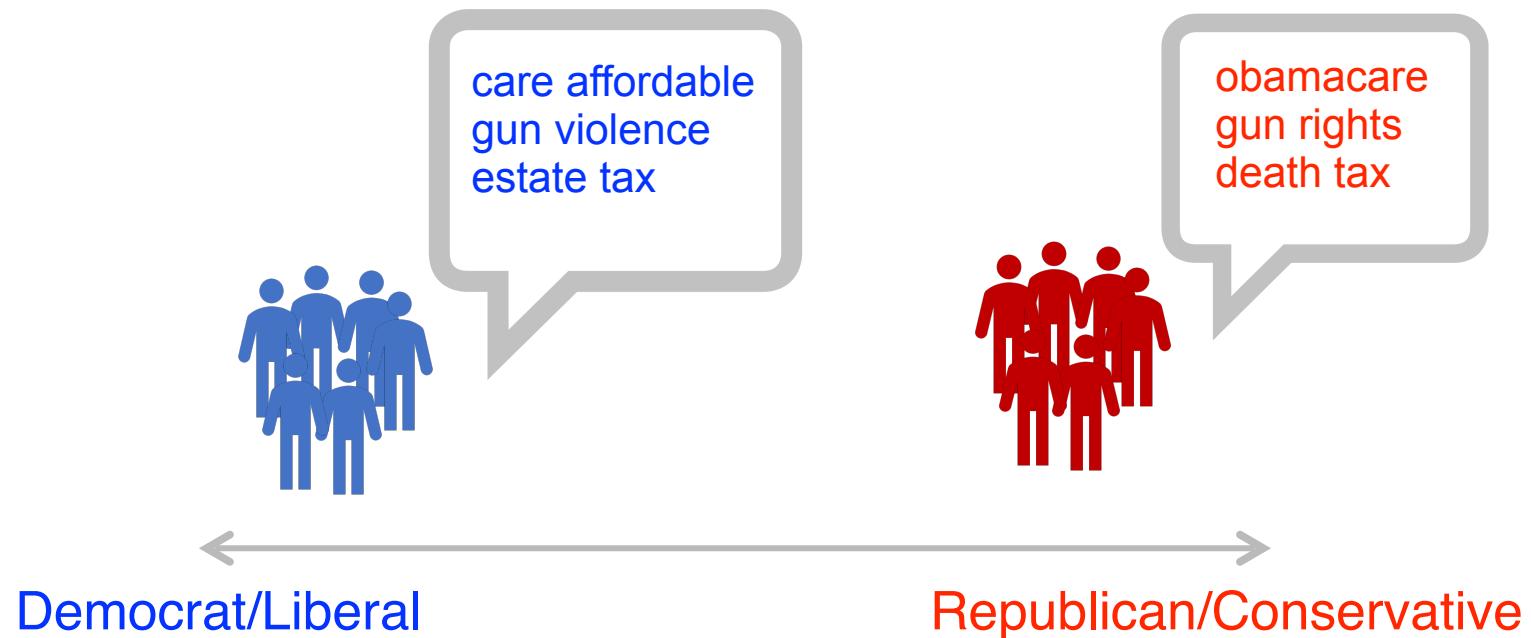
# Outline

- **Dataset**
- Method
- Analysis
- Discussion



# Data overview

## 1. Created a lexicon of partisan terms derived from political speeches



# Data overview

- 1. Created a lexicon of partisan terms derived from political speeches**
- 2. Created a set of search queries using lexicon and autocomplete suggestions**



# Data overview

- 1. Created a lexicon of partisan terms derived from political speeches**
- 2. Created a set of search queries using lexicon and autocomplete suggestions**
- 3. Collected Google SERPs**



# Data overview

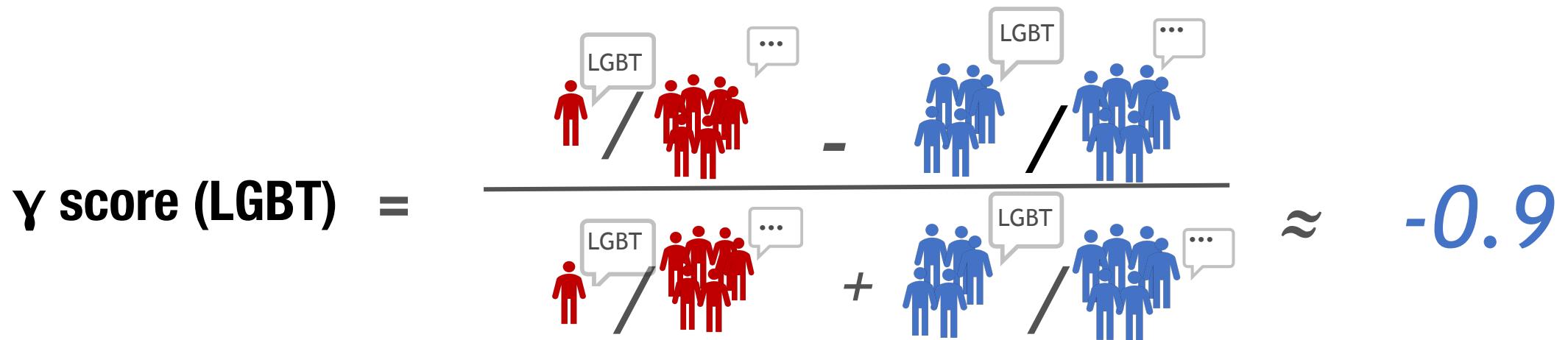
- 1. Created a lexicon of partisan terms derived from political speeches**
- 2. Created a set of search queries using lexicon and autocomplete suggestions**
- 3. Collected Google SERPs**
- 4. Collected webpages on the SERPs**



# Lexicon Dataset

- **Lexicon Construction**

- *Method developed by Gentzkow and Shapiro*
- *Source: transcripts of speeches & letters of politicians from VoteSmart*
- *Time-range: 01/2008 - 08/2018*
- *Preprocessing: downsampling, cleaning, tokenization, filtering*



- **Lexicon Analysis**

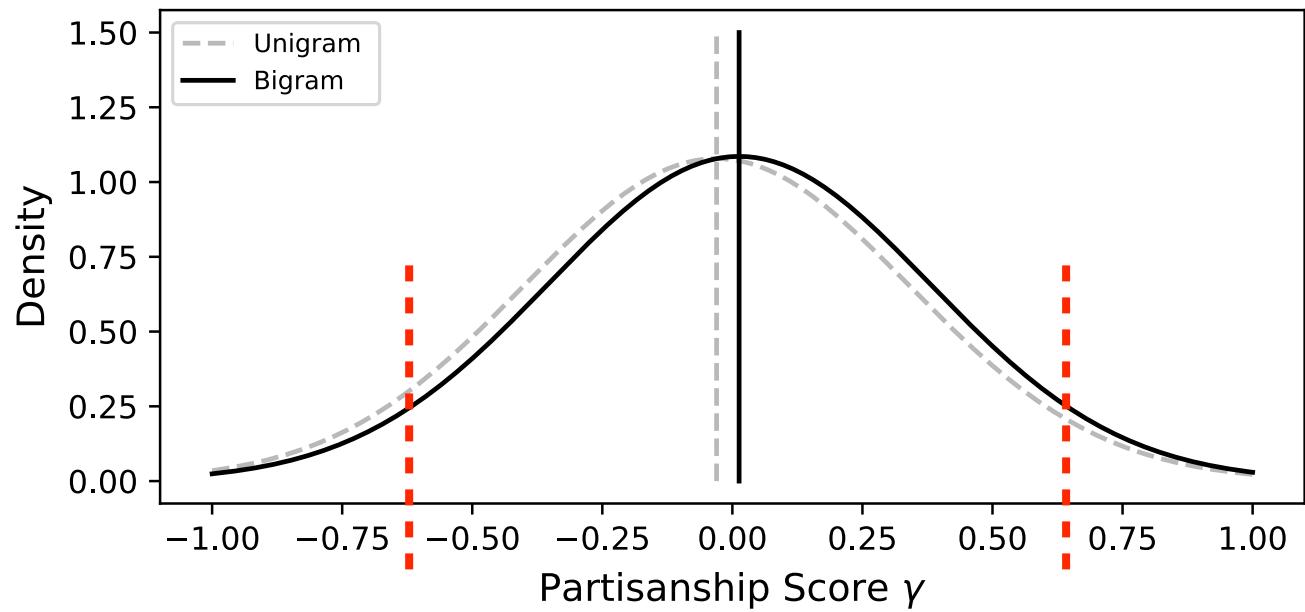


# Lexicon Dataset

- **Lexicon Construction**

- *Method developed by Gentzkow and Shapiro*
- *Source: transcripts of speeches & letters of politicians from VoteSmart*
- *Time-range: 01/2008 - 08/2018*
- *Preprocessing: downsampling, cleaning, tokenization, filtering*

- **Lexicon Analysis**



# Search Query Dataset

- **4570 Root Queries**
  - **3,520 US politician names:** *members of the US House, Senate, Cabinet, Governors, White House staff, 2018 candidates*
  - **1,050 partisan bigrams selected from lexicon**
    - (a) *bigrams with scores < -0.5 or > 0.5*
    - (b) *manually curate terms into valid queries (i.e., filtering, expanding, mutual agreement among raters)*
    - (c) *random downsampling to match same number of bigrams from left and right*



# Search Query Dataset

---

Topic	Root Queries
Gun	conceal weapon, gun rights, gun trace, illegal gun, crime gun, gun violence
Election	military voter, secret ballot, voter assistance, ballot access, voter suppression
President	president health, mandate president, initiative president, president emergency
Pregnancy	elective abortion, pay abortion, unintended pregnancy, emergency contraception
Tax	tax mandate, repeal death tax, burdensome taxes, tax scam, tax giveaway
American	burden american, regulation american, american spirit, eligible american

---



# Search Query Dataset

- **4,570 Root Queries**
  - *3,520 US politician names including: members of the US House, Senate, Cabinet, Governors, White House staff, 2018 candidates*
  - *1,050 partisan bigrams selected from lexicon*
- **Autocomplete suggestions (88,745 unique search queries)**



# SERPs & Webpage Dataset

- **SERPs**
  - *Time range: 10/13/2018 - 10/30/2018*
  - *Queried Google Search for 4,570 Root Queries and all of their autocomplete suggestions, e.g., “donald trump” and its suggestions*
- **Webpage**
  - *Crawled the links on the SERPs*



A screenshot of a Google search results page. The search bar at the top contains the query "dональд трамп". Below the search bar, a list of autocomplete suggestions is displayed, including: donald trump, donald trump twitter, donald trump jr, donald trump age, donald trump net worth, donald trump news, donald trump approval rating, donald trump jr twitter, donald trump education, donald trump signature, and donald trump wife. At the bottom of the screenshot, there are two buttons: "Google Search" and "I'm Feeling Lucky". A small text link "Report inappropriate predictions" is located at the very bottom right.

dональд трамп

donald trump  
donald trump twitter  
donald trump jr  
donald trump age  
donald trump net worth  
donald trump news  
donald trump approval rating  
donald trump jr twitter  
donald trump education  
donald trump signature  
donald trump wife

Google Search I'm Feeling Lucky

Report inappropriate predictions



# Outline

- Dataset
- Method
- Analysis
- Discussion



# Scoring & Comparing

## I. Bias score $\Gamma$ of text:

$S$  — a set of terms in lexicon  
 $\gamma(s)$  — partisan score of term  $s$

$$\Gamma(S) = \frac{\sum_{s \in S} \gamma(s)}{|S|}$$

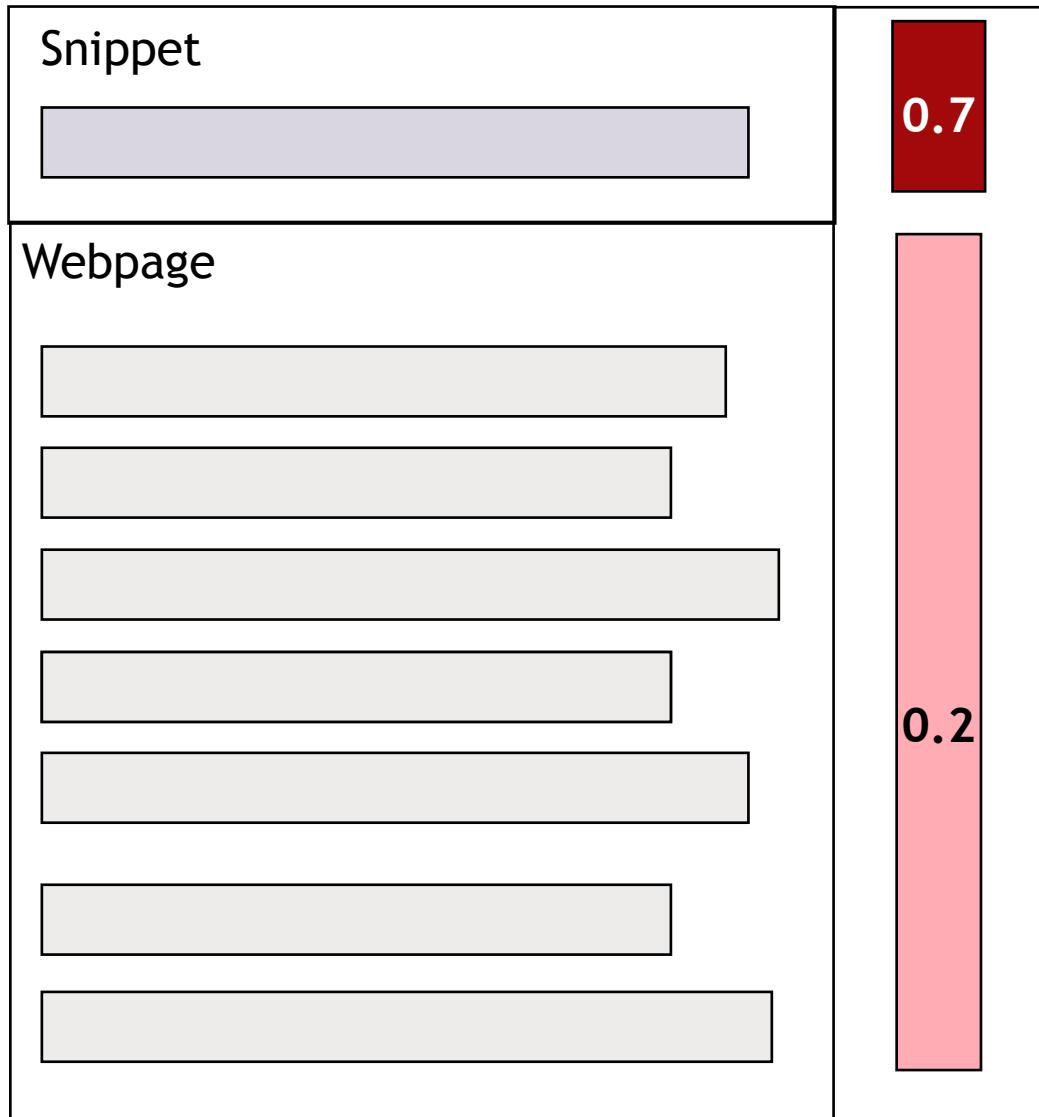
e.g., “equal pay” (-0.90),  
“illegal immigrant” (0.86)  
 $\Gamma(s) = [(-0.90) + (+0.86)]/2 = -0.02$

## II. Bias score comparison:

$\Gamma_s$ — bias score of snippet  
 $\Gamma_w$ — bias score of webpage



# Scoring & Comparing



## II. Bias score comparison:

$\Gamma_s$ — bias score of snippet

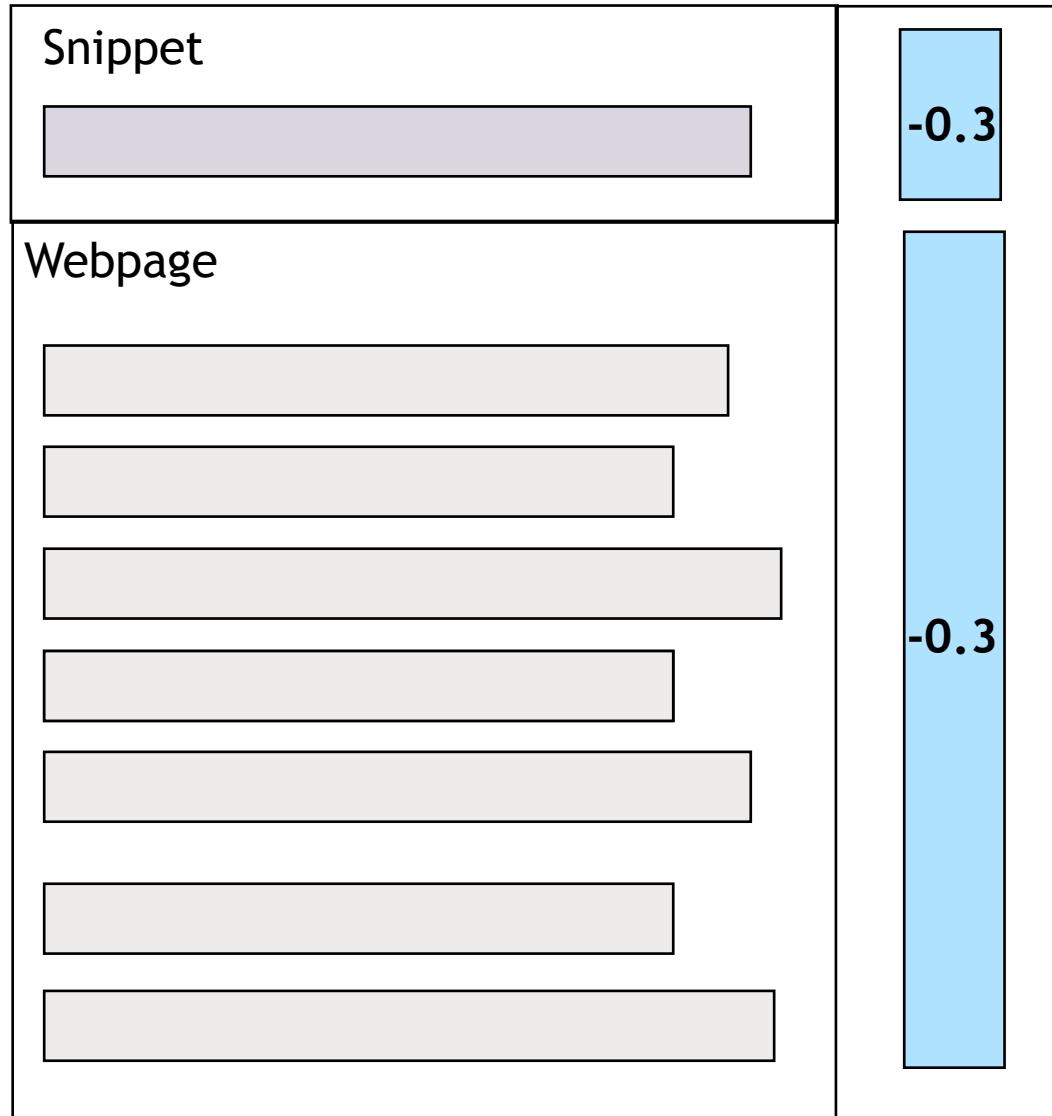
$\Gamma_w$ — bias score of webpage

If  $\Gamma_s * \Gamma_w \geq 0$ :

■ Amplify:  $|\Gamma_s| - |\Gamma_w| > 0.1$



# Scoring & Comparing



## II. Bias score comparison:

$\Gamma_s$ — bias score of snippet

$\Gamma_w$ — bias score of webpage

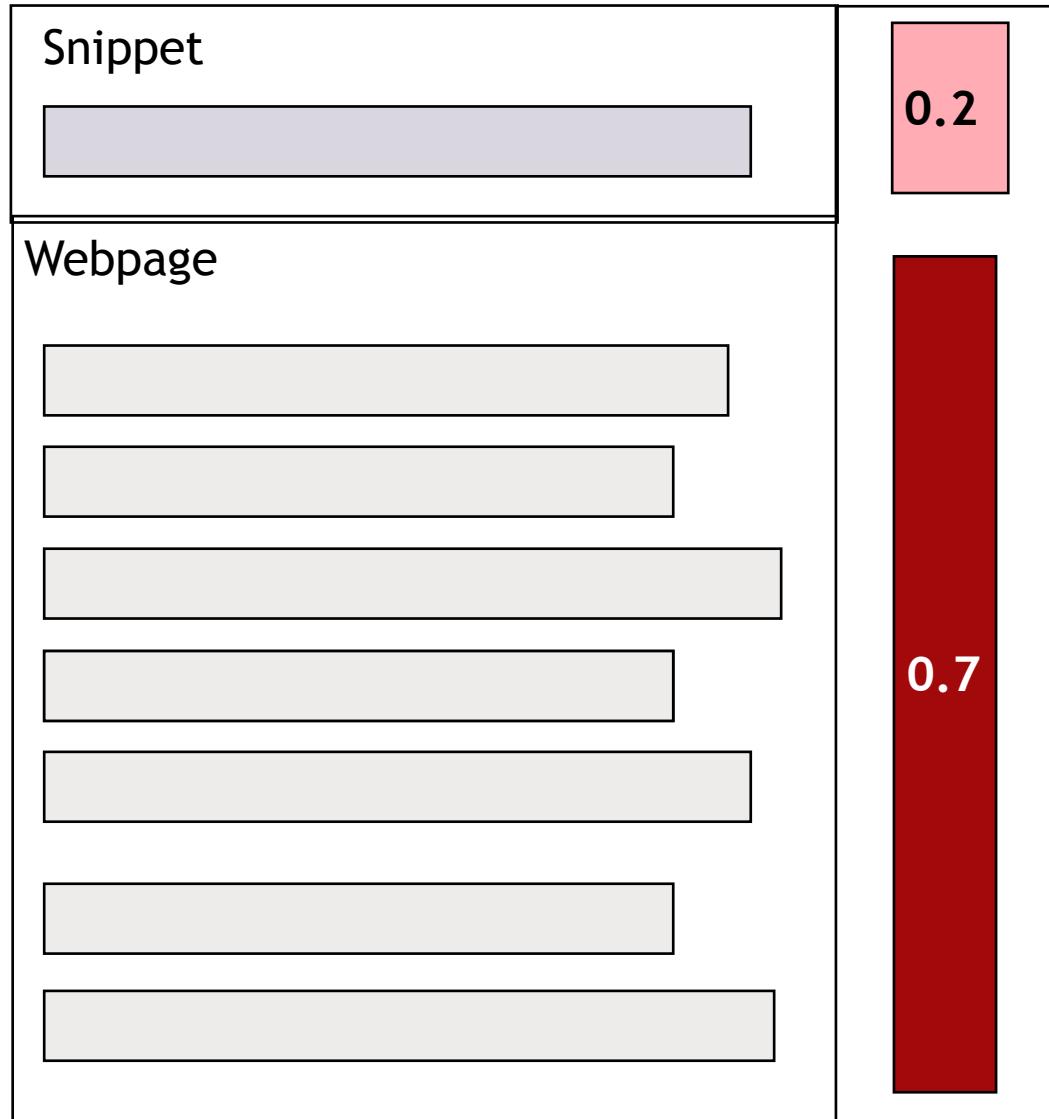
If  $\Gamma_s * \Gamma_w \geq 0$ :

■ Amplify:  $|\Gamma_s| - |\Gamma_w| > 0.1$

■ Unchanged:  $-0.1 \leq |\Gamma_s| - |\Gamma_w| \leq 0.1$



# Scoring & Comparing



## II. Bias score comparison:

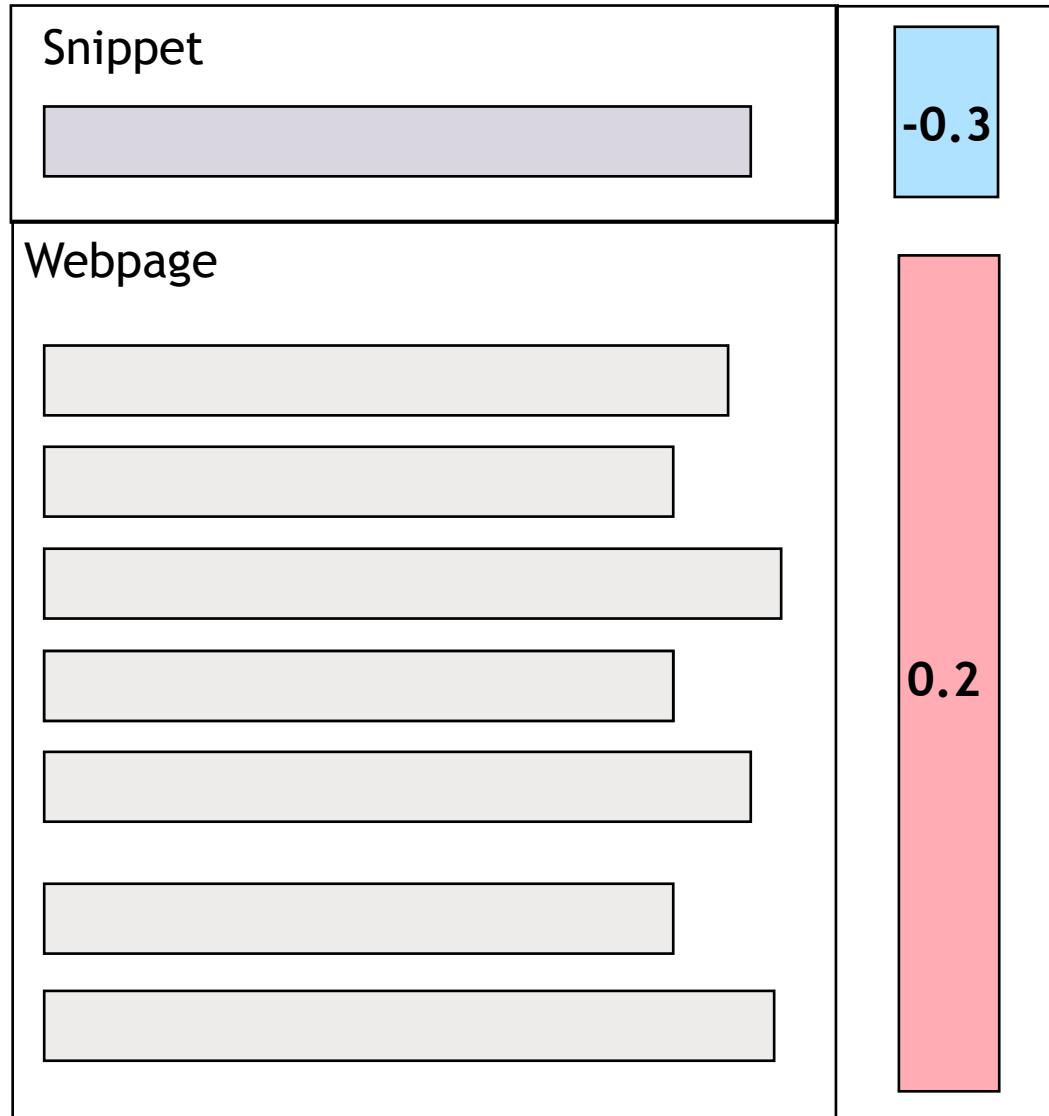
$\Gamma_s$ — bias score of snippet

$\Gamma_w$ — bias score of webpage

If  $\Gamma_s * \Gamma_w >= 0$ :

- Amplify:  $|\Gamma_s| - |\Gamma_w| > 0.1$
- Unchanged:  $-0.1 <= |\Gamma_s| - |\Gamma_w| <= 0.1$
- Decrease:  $|\Gamma_s| - |\Gamma_w| < -0.1$

# Scoring & Comparing



## II. Bias score comparison:

$\Gamma_s$ — bias score of snippet

$\Gamma_w$ — bias score of webpage

If  $\Gamma_s * \Gamma_w \geq 0$ :

- Amplify:  $|\Gamma_s| - |\Gamma_w| > 0.1$
  - Unchanged:  $-0.1 \leq |\Gamma_s| - |\Gamma_w| \leq 0.1$
  - Decrease:  $|\Gamma_s| - |\Gamma_w| < -0.1$
- else:
- Flip: i.e.,  $\Gamma_s * \Gamma_w < 0$



# Outline

- Dataset
- Method
- Analysis
- Discussion



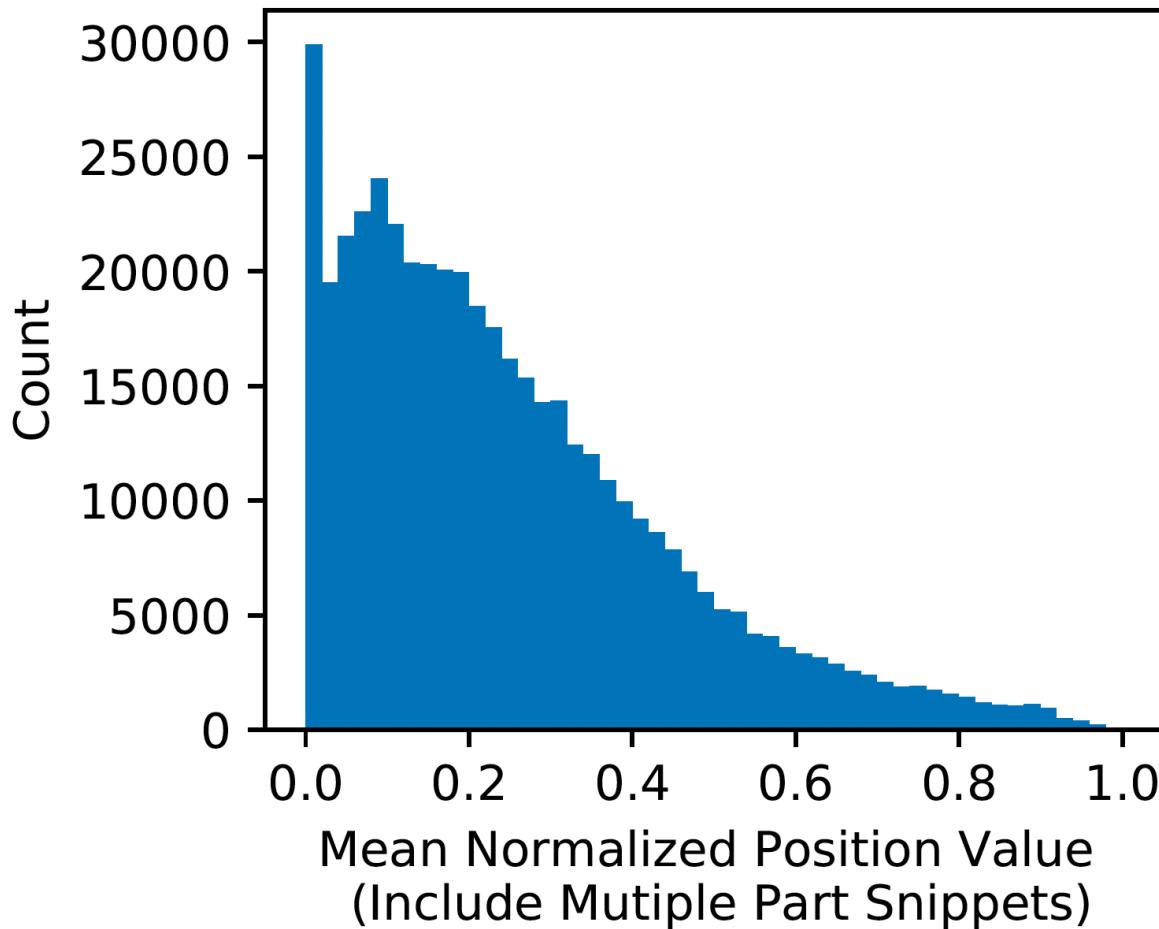
# Overall Results

58% (*amplify*), 14% (*unchaged*), 4% (*decrease*), 24% (*flip*)



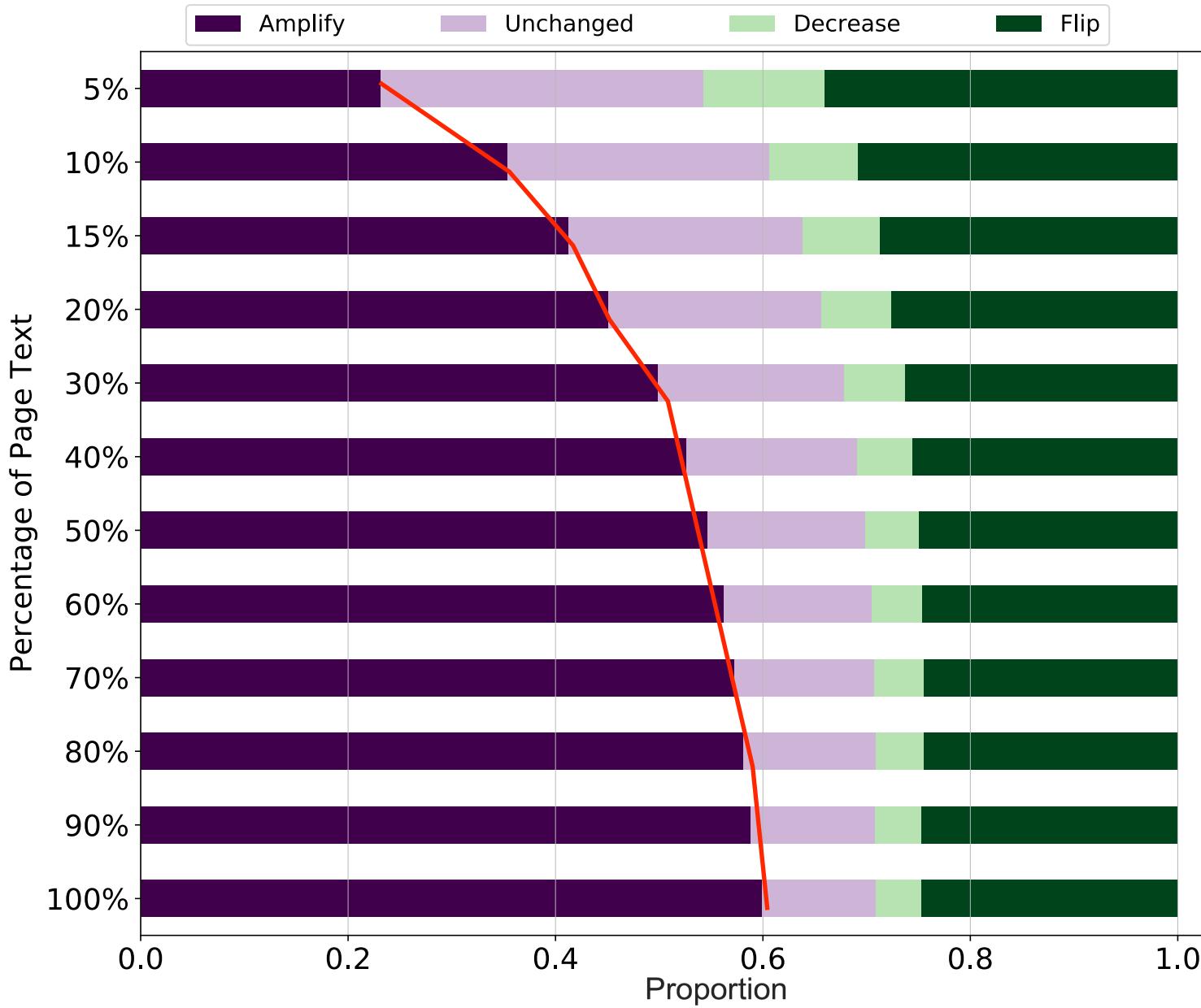
# By Location

- Strong preference for text that is earlier in documents
- May interact poorly with the “inverted pyramid” journalism format



# By Location

- As more webpage text is considered, the amplification effect is stronger
- Possible explanation:
  - (a) Partisan terms of one polarity at the beginning are counteracted by terms with opposite polarity
  - (b) Highly partisan terms at the beginning are averaged out by less partisan terms



# By Query Topics

- 24 topical clusters of partisan phrases
- 6 clusters of politician names

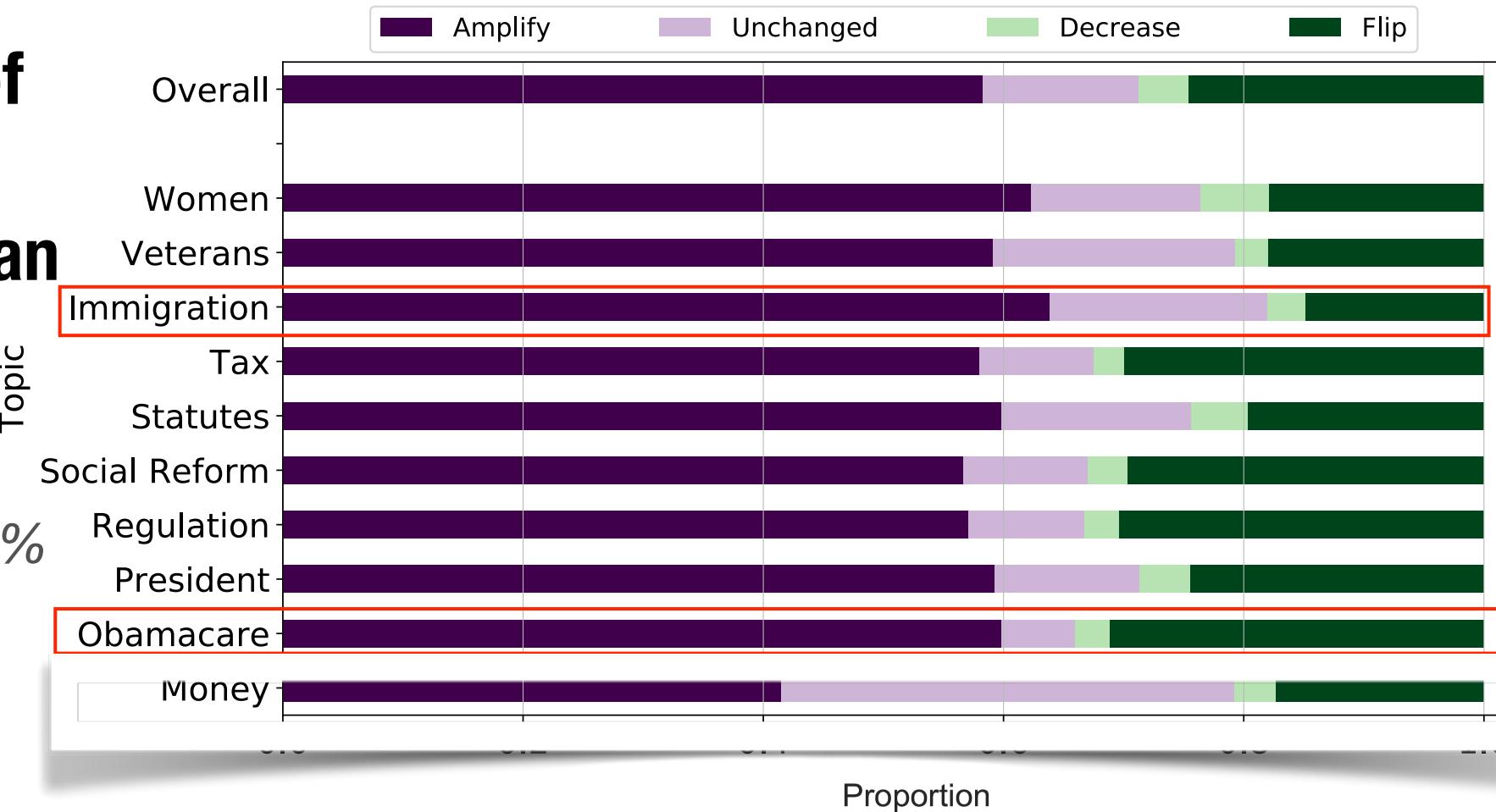
## Greatest effects:

Amplify: “*Immigration*”, 64%

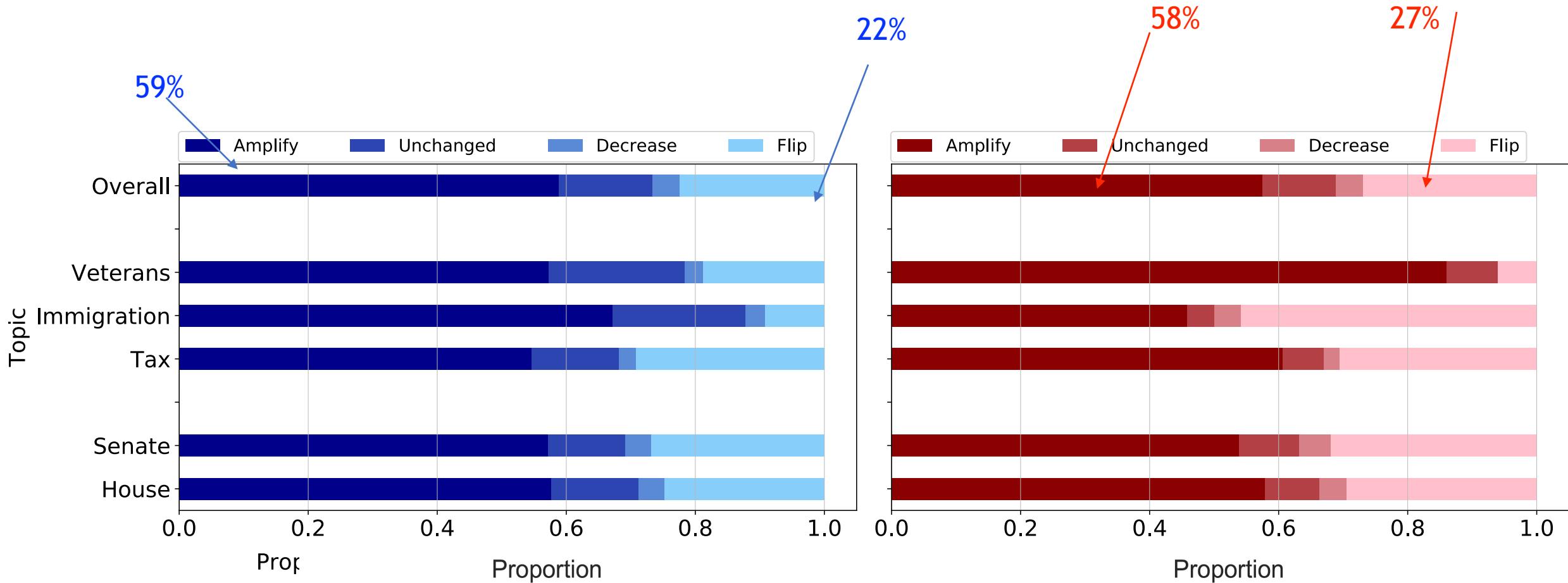
Flip: “*Obamacare*”, 31%

## Most balanced:

“*Money*”(41% amplify,  
38% unchanged)



# By Query Polarity



# Outline

- **Dataset**
- **Method**
- **Analysis**
- **Discussion**



# Limitations

- Contextual semantics aren't considered
- Our lexicon is not completely representative
- Our study is not longitudinal
- People may not recognize all of the n-grams in our lexicon as partisan, or they may not react to them in proportion to their bias scores



# Conclusion — Time to wake up :)

- First evaluation of partisan cues in Google Search snippets
- Snippets tend to amplify partisan cues relative to the original webpages
- May be due to writing style and algorithm preference
- Might impact people's impression or perception of page content
- Future work: longitudinal studies, field studies, lab experiments

