

# **The Political Polarization of US Media: Empirical Evidence from Newspapers**

Alex Williamson

Northwestern University

## I. Introduction

A recurring theme in modern American politics is polarization. Because the effects of polarization are often negative, as could have been the case with the fiscal cliff incident, there is a growing interest in the source of sharp political division. The media industry inevitably enters the discussion, either as a divider drawing sharp lines, a sensationalist machine creating a false sense of polarization, or as a political victim to the rapid pace of technology. In each case, the discussion uses centers around large news networks and online news, two categories in which the political slant can be both very large and vary significantly from outlet to outlet. It is difficult to determine whether biased news is the result of political imbalance or changing technology. An examination of long-existing newspapers, however, could prove to be more fruitful. If print news has become more polarized, then there is evidence of the media leading the nation in political division; if this is not the case, then there perhaps polarization is overstated.

To this end, I use longitudinal data from 248 Newspapers to determine whether biases, or media slants, have changed in the last 20 years. If newspapers are being polarized, then the distribution of media slants will have shifted towards the ends of the political spectrum. I wrote a natural language processing algorithm to find the most politically charged words from the congressional record of each year in the sample. From this, I created an index for each newspaper based on the frequency with they used each of the phrases. The indexes are then compared over

time to gauge whether newspapers are moving further away from the center of the political spectrum.

## II. Literature Review

The last twenty years have brought vast changes in the media industry; this is in no small part due to near-universal access to the Internet; in 2010, 75% of American adults accessed news via the internet.<sup>1</sup> Because a strong and vigilant news industry has long been considered vital to holding various factions of society accountable (Gentzkow, Glaeser and Goldin), easier access to larger quantities of news may appear strictly beneficial. Evidence suggests, however, that an increasing polarization of the US population is accompanying the growth of the internet (Fiorina and Abrams). A reasonable explanation offered by the literature is that as people abandon more traditional forms of news in favor of sources that more closely match their personal ideologies across more dimensions, they are seeking a niche (Gaskins and Jerit). These sources are capable of reaching an audience that is scattered geographically but clustered ideologically. Rather than adjusting personal inclinations with new information, people are able to seek and successfully find information that reaffirms their inclinations.

If the growing ideological gap in both media and media consumers is detrimental, with the eminence and volume of online news at least partially to blame, then amplifying the presence and guaranteeing the financial viability of politically neutral news sources, like print news, could be beneficial. If, however, traditional news sources are being forced away from the political center by the need

---

<sup>1</sup> (Gaskins and Jerit 2012)

to stay competitive, the solution is less clear. Furthermore, if long-existing newspapers are also becoming more ideologically separated, then the direction of causality must also be questioned. An important preliminary question in the discourse on media and political polarization is then whether or not traditional news has become more polarized in recent years.

### III. Creating the Data

#### A. Overview

To track how the biases of newspapers have changed over time, I created a media slant index based on how many times a publication uses certain politically charged phrases in a given year. The phrases are selected as those in the congressional record most likely to be used by the most conservative or the most liberal congressmen (See Section III-E). Borrowing from the statistical methodology suggested by Gentzkow, et. Al (2010), I wrote several automated Ruby scripts that together created the index and compiled relevant information for each publication in my sample. The data was stored and manipulated through PostGres3.

#### B. Web Scrape, Info, Parse

Using the Watir gem for ruby, I created a script to find online and download each passage from the congressional record and identify its speaker. For every speaker that had been entered in to the database, I ran a script to download information like Party, Branch, District, and State. Each Passage was then split into a set of every two and three word sequences, removing stop words<sup>2</sup> and using the Stemmify ruby gem to reduce each word to its root.

---

<sup>2</sup> <http://norm.al/2009/04/14/list-of-english-stop-words/>

C. Phrase Selection

D. News search

E. Index creation

#### IV. Analysis and Results

The coding is finished for every step up to the News Search; because of the sheer size of the congressional record text, however, I will not be able to run the program on my computer. From April 5-9 the program will be running on a local server, from which I will execute the remainder of index generation. Once the data has been created and made computationally mobile, I will move it back to my computer for analysis. I will report summary statistics from each stage of the index algorithm. The estimates of the following model will be the key result of the paper:

$$\Delta Slant_{it} = \beta_0 + \beta_1 t + \beta_2 C_{it-1} + \beta_3 (C_{it-1} \times t) + \beta_4 Circulation_{it} + \beta_5 RegionPol_{it} + \beta_6 WebArticles_{it} + u_i$$

Here  $C_{it}$  denotes the slant of paper  $i$  at time  $t$  minus the mean slant of the papers, yielding positive numbers for the most conservative half and negative numbers for the least conservative half. Circulation size, regional conservatism as measured by the percentage of the popular vote in encompassing congressional districts going to the Presidential Candidate in the prior election, and Online articles are also controlled for.

#### V. Future Steps/Conclusions

A. Concluding Analysis

TBD

### B. If I Had More Time

I will expand here when I know how much I was able to complete.

### C. Other Future Researchers

This paper will provide empirical evidence as to whether or not newspapers have become more polarized: either of these conclusions will have strong implications for both normative and positive research. In the case that polarization is verified, future research may take as given that there is evidence of political polarization in the media and may then assess the effects and ways to circumvent them. If newspapers have remained ideologically stable, then researchers may explore explanations as well as ways to leverage stable media outlets to maintain political balance.

### Works Cited

- Fiorina, Morris P. and Samuel J. Abrams. "Political Polarization in the American Public." Annual Review of Political Science 11 (2008): 563-588.
- Gaskins, Benjamin and Jennifer Jerit. "Internet News : Is It a Replacement for Traditional Media Outlets?" The International Journal of Press/Politics 17.2 (2012): 190-2013.
- Gentzkow, Matthew and Jessie M. Shapiro. "What Drive Media Slant? Evidence From U.S. Daily Newspapers." Econometrica 78.1 (2010): 35-71.
- Gentzkow, Matthew, Edward L. Glaeser and Claudia Goldin. "The Rise of the Fourth Estate: How Newspapers Became Informative and Why It Mattered ." National Bureau of Economic Research. Chicago: University of Chicago Press, 2006. 187-230.