The Impact of Media on Political Polarization:
Evidence from the United States *

February 14, 2022

**JEL Codes:** C15, E60, H12, I18.

**Keywords:** Political Polarization, text-analysis, machine learning

**Acknowledgement:** I am grateful for the suggestions and insights offered by Salvador Ortigueira, Jinhui Bai, and Jia Yan.

*Michael Mahoney, Joesph Navelski, Kennedy Odongo
WSU School of Economic Sciences. Pullman, WA 99163
Email: michael.mahoney@wsu.edu
Phone: (509) 551 8711
ABSTRACT:

Over the last thirty years, political polarization has risen sharply in the United State. Yet, it remains unclear the extent to which political polarization can be explained by changes in the American media environment. In this analysis, we quantify both political and media polarization using millions of lines of text data from political reporting and the proceedings of Congress. We utilize the estimator proposed in Gentzkow, Shapiro, and Taddy (2019) to construct an index representing the divergence in speech patterns between Fox News and CNN; we build a similar index to measure political polarization using speech data captured on the floor of Congress. We next use an Instrumental Variables approach to demonstrate that changes in media polarization causally effect changes in political polarization.
1 Introduction and Literature Review

That political polarization has risen in the United State over the last thirty years is undeniable. A host of theories exist to explain the phenomena, but there has been no comprehensive, quantitative analysis of the role media plays in driving political division. In this analysis we assess the causal impact of media on political polarization. We find a strong causal effect of media polarization on political polarization. Indeed, the relationship is highly significant and robust to the inclusion of economic and political controls, suggesting that changes in media coverage profoundly affect the political environment.

In the subsequent analysis, we apply a novel textual data set to the model developed in Gentzkow, Shapiro, and Taddy (2019). Their model allows us to quantify media polarization and political polarization in each month between September 2013 and September 2021 using text data. Our text data is divided into two components. The first is drawn from articles published by Fox News and CNN. The second is drawn from the proceedings of Congress. We pivot these raw text data into two matrices whose columns represent counts of phrases used by media personalities and members of Congress. We then use this count information to construction a measure of dispersion in language use between each group pair – Fox and CNN and Democrats and Republicans. This metric is our measure of polarization.

From there, we conduct a causal analysis between the two polarization series. We use a two stage estimation approach, instrumenting media polarization with both television viewing hours and measures of coverage satisfaction reported by views of each media outlet. This analysis yields a positive, highly significant effect of media polarization on political polarization. The result is robust to the addition of economic and political controls.

The remainder of the paper is structured as follows: we first review the existing literature on political and media polarization, specifically examining papers that utilize language processing techniques. We then elaborate on our data, and offer an expansive discussion of our collection and refinement processes. In section three, we introduce our model of speech and our measure of partisanship. Finally, we assess causality between our two series using an instrumental variables estimator. We discuss the validity of our instruments and present results.
1.1 Literature Review

There is a wealth of literature concerning political polarization – and I’ll write about it later. A few papers to consider:

1. Gentzkow, Shaprio, Taddy (i.e the big one)

2. Prior (2013) poly-sci paper with the same research question – there’s no model. It’s all bullshit

3. Waller and Anderson (2021) measure political polarization on Reddit

2 Data

Our raw data set consists of text drawn from 50,000 thousand documents published by Fox News and CNN between 2013 and 2021. From CNN we extract 16 million words written in opinion pieces and political reporting. From Fox News we collect another 13 million words, also from political and opinion pieces. For both sets of documents, we identify the month of publication and the author.

In addition, we text mine the Congressional Record, a collection of 1,200 documents containing full text of House and Senate proceedings between 2013 and 2021. From these documents we extract roughly 80 million words. Below we discuss the intricacies of our collection methodology. The proceeding section is split into two subsections, one for media speech data and one for political speech data.

2.1 Media Speech Data

To obtain raw text inputs required for our analysis, we begin by extracting a set of URLs from CNN. These URLs will allow us to extract full text from their associated articles. CNN operates a comprehensive archive of all published articles running back to 2011, making this extraction process straightforward. We simply use python’s web-scraping tools to obtain a list of all URLs in the archive and then filter down for political and opinion pieces. In total this process yields 97,054 links to CNN political news and opinion pieces.
The Fox News website is possessed of no such archive. As such, URL extraction requires us to manipulate the site’s internal search engine. We use a webdriver interface program to automate this process, but search results of the Fox News website are still limited to 100 results. We therefore cannot simply find all articles categorized as ‘political’ or ‘Op-Ed’ by searching for these terms across the 2011-2021 time horizon. Instead we feed a series of search terms into the engine, filter for articles, and then extract all displayed URLs. We repeat this process for every month between July of 2011 and September of 2021. The result is 23,255 web addresses.

Once we have obtained lists of web addresses from both media websites, we use python’s text extraction tools to obtain full text documents from each webpage. The end result is a collection of roughly 23,000 articles from Fox News and 25,000 documents sampled from our CNN addresses. Each text document is then filtered to remove webpage formatting, to extract a publication date and an author. The filtered documents are then split into individual words, bigrams, and trigrams, which in turn are filtered for ‘stop words’, remaining webpage formatting, photograph captions, and author names.

The resulting dataset is far too large to read into the memory of a conventional computer. We accordingly write each word, bigram, and trigram – along with the associated date, author, and news outlet – into a SQLite database. The database allows us to manage this staggering amount of information from a hard disk, rather than access memory. Memory management has proved a challenge throughout this analysis – and for both datasets we will need to restrict the total number of phrases to accommodate the constraints of a decidedly un-super computer.

We next pivot these data into a matrix of counts. Each row represents a media personality at a particular month. Each column represents counts of a particular phrase. For example, we have columns representing counts of the phrases “Gun Control Myths” and “President Barack Obama”. The rows represent the count of key phrases used by each author in a particular month. The matrix itself consists of 4,200 rows and 18,000 columns – 4,200 personality-month pairs, 18,000 phrases.
2.2 Political Speech Data

Our political speech data is obtained directly from both the Congressional Record and transcripts of Congressional Committee hearings. As above, we construct a dataset in which the columns represent particular phrases, and the rows represent counts of those phrases by speaker and month. Here, the speakers are members of the United States Congress (i.e. fucking idiots), divided into Democratic and Republican parties.

We use python’s pdf reader tools to extract full text information from 1,200 daily records of the US Congress. We extract speakers by exploiting the formatting of these documents. Throughout the Congressional Record, speakers are delineated in all capital letters and prefaced by their title. We thus capture the speaker associated with each phrase by splitting each document into sections between title-speaker pairs. Title-speaker pairs that do not correspond to members of congress, such as those of associated with committee witnesses, are discarded.

The result of this process is a collection of 80 million words, spoken by roughly 800 members of congress between 2013 and 2021. From there, it is a simple matter of pivoting the data into counts arranged by speaker and date. To conserve computer memory, we restrict ourselves to phrases spoken at least thirty times and consider only the 6,000 most commonly spoken phrases. The end result is a data set with 25,000 rows and 6,000 columns.

3 Estimation

Our estimation strategy consists of two stages. First we quantify both media and political polarization following the methodology proposed by Gentzkow, Shapiro, and Taddy (2019).

3.1 Quantifying Polarization

Let $i$ index authors and $t$ index time. Each row of data is a $J$ dimensional vector, $c_{it}$, of phrase counts, which we assume follow a multinomial distribution. For each observation, the multinomial density is parametrized by two objects:

1. A verbosity measure $m_{it} = \sum_j c_{jit}$, which measures the total number of
phrases utilized by speaker i in time t

2. A J-dimensional vector of phrase probabilities, which is taken to depend on group affiliation. We denote this vector \( q^{P(i)}(x_{it}) \), where \( P(i) \) represents group affiliation (Fox or CNN/Republican or Democrat) and \( x_{it} \) is a vector of speaker and time specific covariates.

Our subsequent formulation is not substantially different from a standard multinomial logistic model. We assume that the elements of \( q^{P(i)}(x_{it}) \) are determined by a standard softmax functional form. That is, the individual choice probabilities are given by:

\[
q^{P(i)}_{jt} = \frac{e^{\mu_{ijt}}}{\sum_l e^{\mu_{ilt}}}
\]  

(1)

Naturally, the \( \mu \) terms are assumed to have a linear specification:

\[
\mu_{ijt} = \alpha_{jt} + x'_{it} \gamma_{jt} + \varphi_{jt} I(Fox = 1)
\]  

(2)

The above formulation closely resembles a standard multinomial logit problem. The central difficulty in our analysis is the recovery of parameters \( \alpha, \gamma \) and \( \varphi \) given that the number of choices is in the tens of thousands. Below we describe an estimator that allows for the recovery of these parameters in a high-dimensional setting. First, however, we will define our measure of polarization. Now, define

\[
\rho_t(x) = \frac{q^{Fox}(x)}{q^{Fox}(x) + q^{CNN}(x)}
\]  

(3)

We measure partisanship as,

\[
\pi_t(x) = \frac{1}{2} q^{Fox}(x) \rho_t(x) + \frac{1}{2} q^{CNN}(x)(1 - \rho_t(x))
\]  

(4)

This measure has a natural interpretation: it is the posterior probability that an observer with an even prior expects to assign to a speaker’s correct affiliation after hearing the speaker utter a single phrase. We then define average partisanship as,
\[
\hat{\pi}_t = \frac{1}{N} \sum_i \pi_t(x_{it})
\]

Notice here that our goal is to recover estimates of \( \pi_t(x) \), not to precisely estimate the causal terms \( \alpha, \gamma, \) and \( \varphi \). The conceptually simplest approach to form estimates of \( \pi \) would utilize maximum likelihood estimation. There are three limitations associated with direct maximum likelihood estimation. First, the multinomial formulation ignores important speech patterns in assuming independence between phrase counts. This means that the multinomial distribution is most certainly not the true distribution of phrase counts – and hence parameter estimates obtained via MLE will be biased.

Second, bias exists in MLE estimates due to sampling error. Our phrase count data is incredibly sparse, and so certain phrases may be spoken more by Fox News or CNN simply by chance. This will tend to increase our values of even prior \( \hat{\varphi} \), relative to the true value of \( \varphi \). As we’ll see the penalized estimator of Gentzkow, Shapiro, and Taddy (2019) allows us to account for this difficulty by assigning values of \( \varphi_j = 0 \) to least relevant phrases.

Third, and perhaps most important, the objective function associated with the maximum likelihood approach requires substantial computational power to evaluate. In our two applications below, the number of choice outcomes numbers well into the thousands. This will require the evaluation of denominator sums with thousands of terms, for each observation, at each iteration of our optimization routine. Computation would require several months of run time on a conventional computer.

To account for this difficulty, we utilize the penalized estimator of Gentzkow, Shapiro, and Taddy (2019). The estimator uses the Poisson approximation to a multinomial likelihood function first posited by Palmgren (1981) and extended by Baker (1994). This approximation saves us a considerable amount of computing power. The penalty terms also mitigate the sampling error caused by our data’s sparsity.

We minimize the following expression by our choice of \( \alpha, \gamma \) and \( \varphi \):
\[
\sum_j \left[ \sum_t \sum_i \left[ m_{it} \exp(\alpha_{jt} + x'_{it} \gamma_{jt} + \varphi_{jt} I(P)) - c_{ijt}(\alpha_{jt} + x'_{it} \gamma_{jt} + \varphi_{jt} I(P)) \right] + \Psi(||\alpha_{jt}|| + ||\gamma_{jt}||_1) + \lambda_j ||\varphi_{jt}|| \right] \tag{6}
\]

Notice that this Poisson objective is separable across phrases and does not contain large sums of exponential terms. These two facts will make computation of optimal parameters substantially faster, as they allow for the utilization of distributed computing techniques. The penalty terms will ensure that our predicted probabilities have low errors. Since we are interested in prediction, and not the parameters themselves, the inclusion of these terms improves our measure of partisanship. Further, the penalty terms guarantee that many model parameters are set to zero. This will simplify the computation of the \( \hat{q} \) terms need to measure polarization.

The term \( \lambda_j \), which limits the magnitude of our party effect, \( \varphi \), is determined through an iterative process. We begin by setting \( \lambda_j \) to be large enough that the estimate of \( \varphi \) equals 0. For each phrase, we trace the value of lambda down to 0 in 100 steps, running the optimization routine over \( i \) and \( t \) each time. We then pick the value of lambda that minimizes a Bayesian Information Criterion.

For the penalty term \( \Psi \), we use a value of \( 10^{-5} \), as in GST (2019). This low penalty value ensures that solutions to the minimization problem always exist. The term also ensure that our matrix of estimates will be sparse, which proves computationally convenient.

### 3.2 Polarization Results

In our subsequent analysis we utilize trigram phrases. We compute the minimizing parameter values in equation (6) for both congressional speech data and media speech data. The resulting parameter vectors, \( \alpha \) and \( \gamma \) are much too large to present within this paper, and are not of particular interest. In table 1 below we present selected \( \varphi \) terms obtained from both estimation contexts.

Once we have obtained estimates of all model parameters, we calculate analogs of equation (1). For each individual \( i \) in time \( t \), we compute \( \hat{q}_i^{\text{Fox}} \) and \( \hat{q}_i^{\text{CNN}} \), then calculate \( \rho \) and \( \pi \) for all individuals. Notice that we are
MNO – Polarization

comparing the divergence in phrase probabilities within individuals induced by a change in group affiliation. We then simply average over each speaker in a time period to obtain an overall measure of divergence in language use patterns. The result is two time series, one measuring media polarization, the other political polarization. We plot both series below:

Figure 1: Media Polarization 2013-2021

Figure 2: Political Polarization 2013-2021

3.3 Causal Inference

The central purpose of this analysis is to determine the direction of causality between the two series recovered above. To assess causality between media polarization and political polarization we first employ an instrumental variables approach. To instrument for media polarization, we use two monthly series: surveys of viewer satisfaction for both media outlets and the average number
MNO – Polarization

of TV viewing hours in the US.

Coverage satisfaction is doubtless correlated with media polarization. Views. Similarly, Americans might consume more television media when media polarization is high. However, television viewing habits could only conceivably be correlated with political speech though its relationship to the overall media landscape.

Formally, we specify our first model as a standard two-stage least squares estimator, where the first stage is given by:

\[ Media_t = \omega_0 + \omega_1 Sat_t^{fox} + \omega_2 Sat_t^{cnn} + \omega_3 TV_t + \epsilon_{1t} \]  \hspace{1cm} (7)

Where the variable \( Sat \) represents the proportion of viewers of each outlet claiming satisfaction with network programming. The variable \( TV_t \) represents the number of hours of television media consumed by the average American household each day.

\[ Pol_t = \beta_0 + \beta_1 Media_t + \beta_P X_{tP} + \beta_E X_{tE} + \beta_T X_{tT} + \epsilon_{2t} \]  \hspace{1cm} (8)

Where \( pol \) and \( media \) represent indices of political and media polarization respectively. \( X_{tP}, X_{tE}, \) and \( X_{tT} \) represent the political, economic, and time covariates introduced above.

In estimating equation (8), we apply a basic Durbin-Watson test to detect the presence of autocorrelation in our error structure. The test yields convincing evidence that autocorrelation is indeed present, a fact we counter with robust standard errors. Throughout this section, all presented standard errors are inflated via White’s methodology.

The results of the two stage least squares analysis are given in table below. We report a number of specification results and include economic, political, and time controls. Our economic controls include the monthly unemployment rate, the quarterly GDP growth rate, the monthly CPI inflation rate, the quarterly share of output awarded to labor, and the monthly average of the S & P 500 index. Political controls include dummies representing the control of the House and Senate by the Republican party. Finally, we include time fixed effects.

In all cases, we find a positive, statistically significant coefficient on \( media, \)
suggesting a causal relationship between media polarization and political polarization. In the full specification, we obtain a media coefficient of roughly 0.40, suggesting a one point increase in media polarization increases the political polarization index by 0.40. That this relationship is far from one-to-one is hardly surprising, media always exhibits a greater degree of partisanship than politics.

3.4 Robustness Check

The first, and most obvious, robustness check is to repeat the entirety of the above analysis using counts of bigrams, rather than trigrams. Below we plot both polarization series obtained from this analysis, and show that the plots reveal similar trends to those in the trigram analysis. This implies that the polarization using less informative phrases mimics and supports the initial results uncovered using the trigrams as inputs.

Bigram IV Analysis is ongoing.
MNO – Polarization

Figure 3: Media Polarization 2013-2021 (Bigrams)

Figure 4: Political Polarization 2013-2021 (Bigrams)

4 Conclusion

In this analysis we have demonstrated that a causal effect exists between the American media environment and political polarization. Rising media polarization produces a positive, significant change in political polarization. This finding is in contrast to the results of Prior (2013), however there are a number of important distinctions between our analysis and his.

First, our paper examines a time period running from 2013-2021. This time horizon includes several important polarizing events – the Benghazi hearings, the election of Donald Trump, the Charlottesville rally – to name only a few. These events engendered a great deal of attention in both the media and in Congress. It is thus possible that the causal effect of media is itself a new phenomena.
MNO – Polarization

At this juncture, there remains considerable work to finish – robustness checks, further data refinement, the construction of a time series model. But these initial results are promising.