Quantifying Partisan Media Bias: Gatekeeping, Presentation, and Speed of Coverage on Congressional Misconduct

A Thesis Submitted in Partial Fulfillment of the Honors Bachelor's Degree

Jack Andolina

Abstract

The increasing prevalence of partisan media outlets and political polarization in the United States has prompted concerns about the role of media bias in shaping public discourse (Groeling 2013; Iyengar et al. 2019). This study investigates partisan bias in the selection and presentation of congressional misconduct news by comparing coverage from the Republican-leaning outlet Newsmax to a centrist source, the Wall Street Journal. Drawing on theories from psychology and communications, I hypothesize gatekeeping bias, speed of coverage bias, and presentation bias will be evident in Newsmax's reporting through increased coverage of Democrat misconduct and greater rate of mentioning Republican partisans.

Utilizing a computational multi-stage Term Frequency-Inverse Document Frequency based approach, we match news articles to events to quantify potential biases. Our findings reveal systemic biases opposite to expectations, that instead Newsmax covered Republican misconduct with greater frequency than Democrat misconduct. This research challenges the generalized understanding of partisan media bias and prompt the need for more specialized study into Republican partisan media.

Under the Advisement of Dr. Umberto Mignozzetti

April 1, 2024
Acknowledgments

I would like to express my sincere gratitude to my advisor Dr. Umberto Mignozzetti for his invaluable guidance, advice, and support throughout the duration of this project. Special thanks also go to Dr. Davi Moriera for his informal support and insights. I am deeply thankful to Dr. Sean Ingham and Dr. Scott Desposato for leading the senior honors seminar and providing valuable feedback and encouragement at every stage of the research process.

I am grateful to Dr. Maureen Feeley, Dr. Neil Smith, and Dr. Feiyang Sun for their assistance and mentorship in developing my independent research skills. I would also like to extend my appreciation to Anthony Anderson and Linh Le for their exceptional support as teaching assistants in the senior honors seminar.

Finally, I am grateful to my peers in the senior thesis program for their constructive feedback and encouragement throughout this journey. Their insights and collaboration have been instrumental in shaping this research endeavor.
Quantifying Partisan Media Bias: Gatekeeping, Presentation, and Speed of Coverage on Congressional Misconduct

A Thesis Submitted in Partial Fulfillment of the Honors Bachelor's Degree

Abstract

The increasing prevalence of partisan media outlets and political polarization in the United States has prompted concerns about the role of media bias in shaping public discourse (Groeling 2013; Iyengar et al. 2019). This study investigates partisan bias in the selection and presentation of congressional misconduct news by comparing coverage from the Republican-leaning outlet Newsmax to a centrist source, the Wall Street Journal. Drawing on theories from psychology and communications, I hypothesize gatekeeping bias, speed of coverage bias, and presentation bias will be evident in Newsmax's reporting through increased coverage of Democrat misconduct and greater rate of mentioning Republican partisans.

Utilizing a computational multi-stage Term Frequency-Inverse Document Frequency based approach, we match news articles to events to quantify potential biases. Our findings reveal systemic biases opposite to expectations, that instead Newsmax covered Republican misconduct with greater frequency than Democrat misconduct. This research challenges the generalized understanding of partisan media bias and prompt the need for more specialized study into Republican partisan media.

April 1, 2024
Contents

1 Introduction

Partisan Media

Contemporary Partisan News

Perceptions of Media Bias

Misinformation and Partisan Epistemology

2 Measuring Media Bias

3 Methodology

3.1 Data

Legislator Misconduct

Newsmax and Wall Street Journal

3.2 Event-Article Classification

Constructing the Feature Datasets

3.3 Operationalizing Bias

Gatekeeping Dataset

Presentation Dataset

4 Results

Gatekeeping Bias
5 Discussion

5.1 Limitations

5.2 Further Research

5.3 Conclusion
1 Introduction

The rise of partisan media outlets and political polarization in the United States over the past two decades has raised concerns about the media's role in shaping public discourse (Groeling 2013; Iyengar et al. 2019). While short-term exposure to partisan news has been shown to exacerbate polarization and outgroup hostility (Garrett et al. 2014, 2019; Stroud 2010; Levendusky 2013), there remains limited research analyzing the true extent and manifestation of partisan bias in real-world news coverage. Considering the growth of partisan news in the United States (Iyengar et al. 2019) and the greater susceptibility for Republican voters to consume partisan media (Stroud 2010), republican partisan news outlets in the United States are especially important to study. Understanding the nature of any systemic biases is crucial for validating existing experimental studies (which largely rely on theoretical definitions of media bias (Wojcieszak et al. 2023)) on the impact partisan media may have on rising polarization in America.

This study aims to conduct a large-scale empirical analysis of partisan bias in the selection and presentation of congressional misconduct news by the Republican-leaning partisan outlet Newsmax, through a comparison of coverage to a centrist source, the Wall Street Journal. Reflecting theories from psychology of cognitive dissonance (Jean Tsang 2019) and social judgment (Giner-Sorolla and Chaiken 1994; Sherif and Hovland 1961), which suggest individuals' tendencies to seek out attitude-consistent information, I hypothesize that Newsmax will exhibit gatekeeping bias, publishing fewer instances of misconduct by Republican congresspeople than Democratic counterparts (H1). Additionally, we expect Newsmax to exhibit speed of coverage bias, covering Democratic misconduct faster than Republican misconduct.
(H2), and presentation bias, including more references to Republican political actors in coverage (H3).

Utilizing a three-level term frequency-inverse document frequency (TF-IDF) approach to map news articles to a comprehensive list of congressional misconduct events, this study quantifies potential biases in both story selection and framing. Regression analyses assess the influence of the news outlet's partisan leaning on various bias metrics. By rigorously documenting partisan bias patterns in a leading conservative news source, this research makes several important contributions. First, it provides empirical grounding for experimental studies on partisan media effects specific to American Republican-leaning outlets. Moreover, the findings offer insights into how partisan outlets may prioritize and present news in systematically divergent ways that could reinforce existing audiences' beliefs and sentiments. From a broader perspective, this study illuminates an understudied yet increasingly influential component of America's politically fractured media landscape.

**Partisan Media**

Decades of research in media studies have led to a widely accepted definition of partisan media bias, being the extent to which news willfully and systemically favors one party over another (Groeling 2013; Stevenson et al., 1973). This study seeks to analyze two types of partisan media bias - the selection and presentation of information. Selection bias is the purposeful selection of stories which positively reflect on a preferred party or negatively reflect on an opposing party. This is also known as gatekeeper bias, as news organizations are considered the gatekeepers of information to the public, as people obtain the majority of their political information through news, and news organizations make their own deliberations as to what to publish (Shoemaker & Reese, 1996). Presentation bias is the most commonly covered
type of bias by the extant literature, measured in a variety of methods (Groeling 2013; There is No Liberal Media Bias). If gatekeeping is bias in what news organizations cover, presentation is bias in how news organizations cover (D’Allessio & Allen, 2000). This can include bias in framing, tone, sources present, or the issues covered (Groeling 2013). Existing research has measured partisan media bias mainly by either analysis of the actual content of published articles or information about the news organization itself. This study will critically examine biases in both the selection and presentation of news.

**Contemporary Partisan News**

While quantifying media bias itself has been a central focus of research, scholars have also explored the factors that may contribute to or exacerbate biases in news coverage. This line of inquiry distinguishes between demand-side and supply-side factors (Puglisi, 2015).

Demand-side factors relate to the preferences and biases of news consumers, which can influence the incentives and decisions of media organizations. The phenomenon of selective exposure, where individuals exhibit a preference for news sources that align with their ideological beliefs, has been well-documented (Stroud, 2011; Iyengar, 2007). This selective exposure contributes to the creation of "echo chambers," where like-minded news reinforces existing beliefs and attitudes, limiting exposure to opposing viewpoints (Wojcieszak, 2016; Arceneaux, 2012).

These demand-side factors, driven by audience preferences and psychological biases, can create incentives for media organizations to cater to specific ideological niches, potentially amplifying partisan distortions in coverage.

On the supply side, factors related to the competitive dynamics and economic pressures within the media industry may also contribute to biased reporting. Some researchers have
examined the influence of market competition on the extent and nature of bias exhibited by news outlets (Puglisi, 2015). In highly competitive markets, organizations may be incentivized to produce more sensationalized or polarizing content to attract and retain viewers, potentially exacerbating partisan biases. In the United States, partisan news outlets have increasingly taken viewership from centrist networks (Groseclose and Milyo 2005), a trend expected to continue (Kelly 2019)

Furthermore, the spread of misinformation and its acceptance among partisan audiences has emerged as a growing concern. Research by Rini (2017) and Pereira (2018) indicates that partisan identity significantly influences the belief in and dissemination of misinformation, with individuals more likely to accept news that portrays their political ingroup positively, even when the information is false. Clayton (2019) highlighted the role of partisan motivated reasoning in processing news content, further compounding the challenge of combating misinformation in a polarized media landscape.

As the internet is freely available and a source for many’s news, the spread of polarization may be encouraged by the availability of partisan news sources that cater to specific ideological niches, providing audiences with a constant stream of attitude-consistent information and narratives. As a result, partisan media consumption can lead to a hardening of attitudes and a decreased openness to alternative viewpoints (Garrett et al., 2014; Stroud, 2010; Levendusky, 2013). In turn, such effects on political opinion may create a cycle for individuals who reaffirm partisan views and information from alternative sources as biased if the coverage differs from the preferred source.

Both demand-side factors, such as audience preferences and psychological biases, and supply-side factors, like market competition and economic pressures, can contribute to the
manifestation and amplification of media biases. These factors create complex incentives and dynamics that shape the production and consumption of news content, potentially exacerbating partisan distortions and polarization. As competition for viewers in the news media market and appetite for partisan news grows in the United States appears to be increasing (Kelly 2019), research measuring the specific content of US partisan news is urgently necessary.

**Perceptions of Media Bias**

Studies have shown that perceptions of media bias are often shaped by individuals' personal political beliefs and identities. Patterson (1996) highlighted that journalists' own partisan leanings can influence their news decisions, suggesting that biases may stem not only from editorial stances but also from the ideological backgrounds of reporters themselves. Gunther (2012) explored the "hostile media phenomenon," where individuals perceive even balanced reporting as biased against their own views, indicating that bias perceptions are highly subjective, and readers may not be well enabled to identify bias in coverage.

This phenomenon underscores the subjective nature of bias perceptions and how they are often filtered through the lens of pre-existing beliefs and identities. Even when media outlets strive for objectivity, their coverage may be perceived as biased by audiences with contrasting worldviews, leading to accusations of media bias from various ideological camps.

The concept of the "hostile media effect" is further supported by studies examining the influence of cognitive biases and motivated reasoning on news consumption and information processing. Individuals tend to seek out and interpret information in ways that align with their pre-existing beliefs, leading to a reinforcement of existing attitudes and a dismissal of contrary evidence (Lord et al., 1979; Taber & Lodge, 2006).
These psychological factors contribute to the polarization of perceptions regarding media bias, with different ideological groups harboring divergent views on the trustworthiness and impartiality of various news sources. As a result, efforts to combat perceived biases through increased transparency or fact-checking initiatives may be met with skepticism or outright rejection by audiences whose beliefs are deeply entrenched.

Because perception of bias is often contingent on one’s own political beliefs, this study avoids any human content analysis or room for subjective judgements. Instead, the feature variables are true/false statements that are assigned computationally based on the objective reality of the article itself (publish date and whether a political actor is mentioned).

**Misinformation and Partisan Epistemology**

The proliferation of misinformation and its acceptance among partisan audiences has emerged as a significant challenge in the modern media landscape, with profound implications for public discourse and the functioning of democratic societies.

Research has consistently demonstrated that partisan identity plays a crucial role in shaping individuals' beliefs and their susceptibility to misinformation. Studies by Rini (2017) and Pereira (2018) indicate that individuals are more likely to believe and disseminate news and information that aligns with their partisan identity, even when the content is factually inaccurate or misleading.

This phenomenon is rooted in the cognitive biases and motivated reasoning processes that influence how individuals process and interpret information. People tend to evaluate the credibility of information through the lens of their pre-existing beliefs and group identities, leading to a greater acceptance of claims that portray their political ingroup in a positive light (Lord et al., 1979; Kahan, 2017).
The role of partisan motivated reasoning in the spread of misinformation is further highlighted by Clayton (2019), who examined how individuals' partisan identities influence their beliefs and the dissemination of false information. This research underscores the challenge of combating misinformation in a highly polarized media environment, where partisan allegiances can override objective facts and evidence.

Moreover, the phenomenon of selective exposure to partisan news sources, discussed in the previous section, can exacerbate the acceptance and spread of misinformation. When individuals are primarily exposed to information and narratives that reinforce their existing beliefs, they may become more susceptible to believing and sharing misinformation that aligns with those beliefs (Garrett et al., 2019; Vosoughi et al., 2018). Research into those most vulnerable to misinformation have found that some groups are more prone than others.

Historical patterns (Domke et al 1999; Watts et al. 1999) and experimental research (Kelly 2019) have indicated that Republicans are more prone to consume partisan news without recognizing bias. Republican politicians have labeled mainstream news of favoring Democrats for decades as biased against their party for decades (Domke et al 1999; Watts et al. 1999), while most of the accusations of news bias by Democrat politicians are directed toward a singular source—Fox News (Groeling 2013). A continuation of the historical trend of Republicans and mainstream media labeling mainstream news as biased and endorsement of alternative complementary news sources, alongside contemporary research finding Republicans increased preference for partisan news (Kelly 2019) makes Republican partisan news especially relevant to study. Due to these considerations, this study focuses on analyzing Republican partisan news.

2 Measuring Media Bias
Quantifying media bias, particularly partisan and ideological biases in news coverage, has been a longstanding challenge for researchers. Early attempts often relied on simple measures of balance, assuming that deviations from equal coverage of political parties or candidates indicated bias (Schiffer, 2006). However, this approach has been criticized for failing to account for legitimate, non-partisan factors that influence editorial decisions and news judgments, such as the inherent newsworthiness of certain events or issues (Groeling, 2013). To address these limitations, researchers have developed more nuanced and sophisticated methodologies to detect and measure media bias.

Manual content analysis is the traditionally dominant method for detecting bias in media content (Goreling 2013). However, research has consistently shown that identification of bias is highly contingent on the reader’s preexisting beliefs (Vallone et al. 1985; Gunter 1992; Gunther et al. 2001; Feldman 2011; Gunther et al. 2012; Stroud et al. 2014; Kelly 2019), and hand coding may leave the study open to this bias in the identification of bias. In addition, content analysis on a large scale is time-exhaustive, limiting studies utilizing this technique to studying small sets of articles or necessitating large budgets. Due to advantages in timeliness and replicability, automated content and organizational analyses have been implemented to measure and identify biases in news articles (Groeling 2013).

One prominent strand of research has focused on comparing media coverage to benchmarks beyond just political parties, such as public opinion polls or the stated positions of elected officials. This approach aims to identify biases in the intensity, tone, or framing of coverage relative to these external reference points. For instance, Puglisi (2015) provides a framework for quantifying bias by contrasting media spin with audience beliefs. Lott (2004) employed an econometric technique that controls for the underlying nature of news events,
finding evidence that U.S. newspapers tend to provide more positive economic news coverage when Democrats are in the White House compared to Republican administrations.

Another line of inquiry has involved content analysis to examine the selection of news stories and the framing of political issues as potential indicators of bias. Elejalde (2018) utilized social media data, specifically tweets, to compute the political and socioeconomic orientations of news outlets, revealing measurable biases in their coverage choices and narratives. Coffey (1975) highlighted the importance of developing objective techniques for determining bias in political reporting, emphasizing the ethical obligation of journalists to provide balanced coverage.

Recent years have witnessed a surge in the use of computational methods to measure bias, misinformation, and the presence of fake news. Latent Dirichlet Allocation (LDA) and Structural Topic Modeling (STM) are popular techniques employed by researchers to effectively identify bias in framing of events over large corpuses (Akcakir et al., 2023; Hamborg et al., 2020). Lexicon-based approaches are used to measure the sentiment of articles, but lack context sensitivity (Spinde et al., 2021). Some advanced techniques, such as neural networks and deep learning models, have emerged as powerful tools for automated bias detection in news articles (Spinde et al., 2021). Language models like BERT, offer sophisticated means of identifying biases trained by bias experts (Spinde 2022).

With the advent of computational methods and machine learning, recent studies have leveraged these tools to systematically quantify media bias at scale. Budak (2016) combined machine learning algorithms and crowdsourcing techniques to investigate how major U.S. news outlets select and frame political issues. Surprisingly, their findings suggested that these organizations are more similar than often believed, with little evidence of systematic partisan differences in story selection, except in the context of political scandals.
Building on these approaches, Chen (2020) developed a neural model to assess political bias and unfairness in news articles at various levels of granularity, from individual words to overarching discourse patterns. Their analysis identified insightful bias patterns, demonstrating the potential of advanced computational techniques to uncover nuanced manifestations of media bias.

The choice of methodological approach to measure media bias hinges on the specific type of bias and context under examination. While traditional methods like hand-coding persist, computational techniques offer scalability and efficiency, particularly when validated by qualitative assessments. Human and machine learning-based content analysis attempting to identify subjective traits of articles can be prone to biases. To avoid biases in measuring bias and preserve computational efficiency, this study will utilizes a computational TF-IDF based approach not observed yet in the extant literature to provide a comprehensive representation of media bias.

3 Methodology

3.1 Data

Legislator Misconduct

The data on legislator misconduct and congresspeople (name, gender, party, etc) is supplied from GovTrack
d. Before subsetting the GovTrack legislator misconduct dataset, it contains descriptions (consequences, dates, descriptions of allegations, etc) of 491 “alleged and actual misconduct” by United States congress people from 1789 to the present. The data includes “public information about congressional investigations, criminal convictions, censures by and  

\[1\) https://www.govtrack.us/misconduct
expulsions from Congress” sourced by contemporary news reports, the US Senate Historical Office dataset of Senate Election, Expulsion, and Censure cases, the Washington Post’s list of congressional indictments, and Wikipedia’s list of Convictions of American Politicians as of January 23rd 2018. A note to acknowledge is that other datasets were used for misconduct which occurred previous to the timeframe of this study which are not incorporated into the misconduct dataset. Specifically, the dataset of our timeframe includes “all letters of reproval, censures, and expulsions from congress”, “all investigations by the House Office of Congressional Ethics (OCE)... , the Huse Committee on Ethics (HCE)... , the Senate Selecte Committee on Ethics... , all senate votes on ‘exclusion’ related to personal misconduct”, “other investigations by a body of congress and monetary settlements that involved alleged personal misconduct” including “settlements administered by Congress’s office of compliance regarding sexual harassment claims”, resignations deemed relevant to misconduct allegations, and felony convictions and cases of misconduct of national significance before and after the legislators terms. GovTrack makes subjective judgements in terms of what constitutes national significance. Objectivity in the measurement of the list of newsworthy events is a central concern of media bias studies due to the population of events problem, but GovTrack’s subjective judgements constitute a slim minority of the cases.

This dataset is subsetted to only include instances of misconduct which occurred during the timeframe of the news data (January 1st, 2010 to November 1st, 2023), resulting in 189 observations. Then, the dataset is merged with GovTrack’s congressional dataset, the main variable of interest being the congressperson’s party affiliation. The following visualizations depict the merged dataset.

2 https://www.govtrack.us/misconduct
Graph 1: Total number of allegation accusations per party

Graph 2: Allegation accusations per party over time
Notably, Republicans are subject to slightly more cases of misconduct (Democrats: 92, Republican 97) over the same time frame. In addition, there is a large spike in Democrat misconduct in July 2022. ~86% of these allegations concern arrests of over a dozen Democrat legislators at a protest at the US Capitol over abortion rights.

**Newsmax and Wall Street Journal**

This study refers to the AllSides Media Bias Chart in choosing the strongly partisan (Newsmax) and centrist (Wall Street Journal) media. Because there exists no available dataset of partisan news articles and no out-of-the-box news crawler was adapted for our purposes, it was necessary to create a specialized news crawler built to navigate Newsmax’s online archive. This crawler collected the article titles, date of publication, main text, and link of each article available on Newsmax’s website from January 1st, 2010 to November 1st, 2023. The Wall Street Journal (WSJ) corpus was provided through ProQuest’s TDM Studio, licensed through UC San Diego. Both datasets are subsetted to only include articles which mentioned any congressperson who were subject to a misconduct filing’s last name. Preceding event classification, the WSJ and Newsmax corpuses included 834,924 and 107,240 articles, respectively.
3.2 Event-Article Classification

The goal of our event-article classifier is to find articles which are most likely covering any one from a list of events. Possibly due to the specificity of each sought-after event relative to the large extent of articles and misconduct allegations, topic models (STM, LDA) were inaccurate in identifying the events of legislative misconduct. Large language models (BERT and ChatGPT) were also tested but found to be computationally expensive and relatively inaccurate. Because our events of interest - legislator misconduct - are described in objective detail in the GovTrack dataset, TF-IDF similarity scores were found to be a highly accurate method of automating the selection of all articles concerning the events. After a series of testing with various NLP techniques, it was determined that a TF-IDF similarity score validated with a set of logical rules resulted in the most accurate classification pipeline.

TF-IDF operates by assessing the significance of specific words within a corpus of documents. In our context, it calculates the frequency of terms associated with legislator misconduct within individual articles. The Term Frequency (TF) component measures how often words related to these events appear, while the Inverse Document Frequency (IDF) assesses the rarity of these terms across the entire document collection. This dual measurement allows for the identification of terms both prevalent within an article and distinctive to the sought-after event.

Table 1: Newsmax TF-IDF Example

<table>
<thead>
<tr>
<th>allegation</th>
<th>article_text</th>
<th>similarity_score</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘In 2015 Honda was investigated for using official resources for campaign purposes.’ ...</td>
<td>‘The House Ethics Committee says it is investigating Democratic Rep. Mike Honda of California for a possible ethics violation.’ ...</td>
<td>~0.65</td>
</tr>
<tr>
<td>‘In 2009 Waters was investigated for a conflict of interest with respect to meetings with a bank in which she had a financial interest.’...</td>
<td>&quot;The House Ethics Committee concluded Wednesday that Rep. Maxine Waters' rights weren't violated in an ongoing investigation, clearing the way for the panel to determine whether the California Democrat improperly tried to steer federal money to a bank where her husband is a shareholder.’...</td>
<td></td>
</tr>
<tr>
<td>&quot;The heads of 47 government watchdog agencies have written to ‘Congress claiming that federal departments are deliberately delaying or denying access to documents they are allowed by law to review, Politico reported.”...</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| “In 2014 Chu received a letter of reproval for using House staff to perform campaign activities and then obstructing the investigation.”... |

To ensure accuracy in matching articles covering specific events of congressional misconduct, I implemented two rules: a limit of 300 days on the time difference between the event of misconduct and possible articles, and a minimum similarity score of .3. These rules were selected based on qualitative evaluation of the collection of news articles.

By computing TF-IDF similarity scores between article terms and the objective descriptors of legislator misconduct in the GovTrack dataset, our study achieves a robust method of identifying articles closely related to these events. Furthermore, the integration of logical rules validated against these scores enhances the accuracy of our classification pipeline, enabling the automated selection of articles specifically concerning instances of legislator misconduct.

This amalgamation of TF-IDF methodology with a set of logical rules presents a promising framework for event-article classification, particularly in discerning nuanced and specific events within a voluminous array of articles.
Constructing the Feature Dataset

To collect the Newsmax article data, our study utilized Python's BeautifulSoup library for web scraping, compiling a comprehensive dataset comprising all article texts and their publication dates from Newsmax, covering January 2010 to November 2023. Additionally, the Wall Street Journal data was obtained through UC San Diego’s license with ProQuest.

Using TF-IDF scores as a basis, the first step consisted of establishing similarity metrics between articles and descriptions of relevant events. Despite initial setbacks such as misidentifying similarly named entities and events, the study implements a similarity score cutoff of 0.3, a number leading to the highest accuracy in the sample set.

To create the most highly accurate classification system for matching articles to events, we devised a two-pronged approach. Firstly, leveraging named entity recognition machine learning algorithms, supplemented by rule-based validation mechanisms, we identified articles mentioning specific legislators who had been accused of misconduct with an ~95% accuracy rate in the test set. Secondly, we utilized TF-IDF similarity scores to determine the most significant and differentiating words within both the articles and the events. The resulting process lead to an ~85% accuracy rate in matching articles to corresponding events.

3.3 Operationalizing Bias

A key difficulty in measuring media bias is the thoroughly studied unobserved population of events problem (Groeling 2013). In order to determine bias in coverage, we must understand the actual nature of the event. An event may be covered more by a news organization not because of bias, but instead by the inherent newsworthiness of the event. In addition, because we
do not naturally observe the full population of potentially newsworthy events, it is another problem to identify bias in the selection of news. Scholars have employed various techniques to handle the unobserved population of events, which impacts the validity of measuring all forms of media bias. In large part, we handle the population of events problem by focusing on coverage of a set list of events: congressional misconduct brought before the House Ethics Committee. By comparing the coverage from partisan news organizations to a centrist news source, we can control for newsworthiness in part.

Structural factors may have a large impact in determining the newsworthiness of an event. A political party’s power may have some influence in the determining of the newsworthiness of an event concerning a partisan. However, because the US is a majority two party system, and uniparty rule is not a factor in our selected time frame, we anticipate that power within congress has little influence on the newsworthiness of the respective parties.

As news consumers have exhibited a lack of ability to identify subtle forms of bias, especially bias in the selection and presentation of news stories (Ribeiro et al., 2018), these latent forms of bias may be the most impactful in guiding political attitudes. Therefore, this study will measure and analyze bias in the selection of stories, extent and speed of coverage, and the rate of referencing of political actors.

Partisan gatekeeping bias is the selection of news which manifests in a systematic tilt toward or against a political party (Groeling 2013). In the context of congressional misconduct, we hypothesize that partisan news will cover instances of congressional misconduct by ideologically opposing politicians with lesser frequency. This type of bias is impactful in the formation of political opinions, as it may lead to readers forming a view that one party is more corrupt. Polls show that perceptions of corruption in the opposing parties are growing in the
United States, particularly among Republicans\textsuperscript{4}. Partisan gatekeeping bias is particularly prone to the unobserved population of events problem; if the, as if we do not know the array of events which journalists must choose from is unknown, then consumers’ our measurements of bias will be weakened (Groeling 2013). Traditional content analyses of biased news often fail to properly control for this factor (Groeling 2013).

Every event of congressional misconduct filing is available to the public allowing for a highly accurate representation of the population of events. While there may be events of congressional misconduct which are not filed, we assume that those events are likely not filed because of lesser severity. We also assume that all instances of congressional misconduct will be reported.

**Gatekeeping Dataset**

This study measures gatekeeping bias in three forms: the selection, extent, and speed of coverage of events of congressional misconduct. In measuring the first form, selection, this study follows these steps:

1. Link articles to events utilizing TF-IDF representation.
2. Construct a misconduct event level dataset which indicates whether the news organization covered each event.
3. Run OLS regressions of the party of the congressperson on the binary coverage variable for the partisan and centrist news to answer the first research question: Is partisan media more likely to cover events of congressional misconduct by ideologically aligned congresspeople?

In measuring the second form of gatekeeping bias, the extent of coverage of news, this study follows these steps:

(1) Link articles to events utilizing TF-IDF representation.

(2) Construct a misconduct-event level dataset which indicates how many articles were written covering each event.

(3) Run OLS regressions of the party of the congressperson on the integer time difference variable for the partisan and centrist news to answer the second research question: Will partisan media cover events of congressional misconduct by ideologically aligned congresspeople less?

In measuring the third form of gatekeeping bias, the speed of coverage, this study follows these steps:

(1) Link articles to events utilizing TF-IDF representation.

(2) Construct a misconduct event level dataset which indicates the smallest time difference between the event and the publishing of the covering article.

(3) Run OLS regressions of the party of the congressperson on the integer smallest time difference variable for the partisan and centrist news to answer the third research question: Will partisan media cover events of congressional misconduct by ideologically aligned congresspeople slower?

**Presentation Dataset**

In the case of partisan presentation bias, a partisan news organization may frame instances of misconduct by ideologically opposed politicians as greater offenses, and ideologically aligned as less. In addition, they may frame the event through the lens of affiliated
partisans by quoting them and referencing their opinions. This study measures the latter, by finding all mentions of partisans per article and creating cumulative counts per misconduct event. 

1) Link articles to events utilizing TF-IDF representation.

2) Using Named Entity Recognition, find the frequency that opposing partisans are mentioned in articles on the instance of congressional misconduct.

3) Estimate party affiliation with a series of rules (if sentence contains ‘R-’ before their name, for example) followed by qualitative validation.

4) Create cumulative count scores of Republican and Democrat partisans per all articles on events of misconduct.

5) Run OLS regressions of the party of the congressperson on the sentiment of unrelated articles for the partisan and centrist news to answer the second research question: When partisan media cover events of congressional misconduct, do they mentioned aligned partisans more than centrist media?

4 Results

The following results test the three hypotheses of this paper meant to provide a thorough and nuanced description of partisan bias in news. Notably, these results show my first hypothesis is incorrect. Instead, the conservative partisan news source covered instances of Republican misconduct more than instances of Democrat misconduct, both in terms of binary coverage and number of articles. My second hypothesis is proven incorrect as well. Instead, the conservative partisan news source covered instances of Republican misconduct essentially the same as Democrat misconduct. My last hypothesis, that the conservative partisan news organization will reference Republican partisan actors across all events of congressional misconduct more than the centrist source, was found to be correct.
Gatekeeping Bias

Graph 3: Coverage of Misconduct Events by Party

The graph above shows that Newsmax covered significantly more cases of Republican misconduct than Democrat misconduct. When compared to the Wall Street Journal to control for newsworthiness, the trend holds. As expected, the Wall Street Journal shows no selection bias in the choosing of stories to cover.
Table 1: Effect of Party on Coverage of Congressional Misconduct

<table>
<thead>
<tr>
<th></th>
<th>Dependent variable:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Member of Republican Party</td>
</tr>
<tr>
<td></td>
<td>Wall Street Journal</td>
</tr>
<tr>
<td></td>
<td>Newsmax</td>
</tr>
<tr>
<td>Covered</td>
<td>0.070</td>
</tr>
<tr>
<td></td>
<td>(0.076)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.487***</td>
</tr>
<tr>
<td></td>
<td>(0.046)</td>
</tr>
<tr>
<td></td>
<td>0.206***</td>
</tr>
<tr>
<td></td>
<td>(0.076)</td>
</tr>
<tr>
<td></td>
<td>0.444***</td>
</tr>
<tr>
<td></td>
<td>(0.044)</td>
</tr>
<tr>
<td>Observations</td>
<td>189</td>
</tr>
<tr>
<td>R²</td>
<td>0.005</td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>-0.001</td>
</tr>
<tr>
<td>Residual Std. Error (df = 187)</td>
<td>0.501</td>
</tr>
<tr>
<td>F Statistic (df = 1; 187)</td>
<td>0.853</td>
</tr>
</tbody>
</table>

| Note:          | *p<0.1; **p<0.05; ***p<0.01 |

These first results show the first hypothesis, that conservative partisan media will cover instances of misconduct by ideologically aligned congresspeople less, is untrue. Instead, these findings point to the opposite. The linear regression of party association of individuals accused of congressional misconduct on a binary metric of coverage show that the partisan news organization covers ideologically aligned misconduct with greater frequency. The findings are not only highly statistically significant at the p<0.01 level, but show a clear divergence in the selection of stories from the centrist source through a comparison of the coefficients (.070 and .206***). The repeated linear regression on the centrist source’s data show no statistically
significant relationship between party association of congresspeople accused of misconduct and coverage of said event.

**Extent of Coverage**

**Table 2: Effect of Party on Extent of Coverage of Congressional Misconduct**

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>Wall Street Journal</th>
<th>Newsmax</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Number of Articles Covering Event</td>
<td>0.005 (0.008)</td>
<td>0.037*** (0.013)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.502*** (0.041)</td>
<td>0.466*** (0.039)</td>
</tr>
</tbody>
</table>

| Observations | 189 | 189 |
| R²           | 0.002 | 0.044 |
| Adjusted R²  | -0.003 | 0.038 |
| Residual Std. Error (df = 187) | 0.502 | 0.491 |
| F Statistic (df = 1; 187) | 0.414 | 8.506*** |

*Note: *p<0.1; *p<0.05; ***p<0.01

The above results show the second hypothesis, that the extent to which conservative partisan media will cover instances of misconduct by ideologically aligned congresspeople less, is untrue. Instead, these findings point to the opposite. The linear regression of party association of individuals accused of congressional misconduct on a cumulative metric of coverage show that the partisan news organization covers ideologically aligned misconduct with greater frequency and extent. The findings are not only highly statistically significant at the p<0.01
level, but show a clear divergence in the selection of stories from the centrist source through a comparison of the coefficients (.005 and .037***). The repeated linear regression on the centrist source’s data show no statistically significant relationship between party association of congresspeople accused of misconduct and the extent of coverage of said event.

**Speed of Coverage**

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>Wall Street Journal</th>
<th>Newsmax</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speed of First Coverage</td>
<td>$-0.0002$ (0.001)</td>
<td>$-0.001$* (0.001)</td>
</tr>
<tr>
<td>Constant</td>
<td>$0.555$*** (0.074)</td>
<td>$0.701$*** (0.074)</td>
</tr>
</tbody>
</table>

| Observations | 74 | 69 |
| R² | 0.001 | 0.041 |
| Adjusted R² | $-0.013$ | 0.027 |
| Residual Std. Error | $0.505$ (df = 72) | $0.482$ (df = 67) |
| F Statistic | $0.096$ (df = 1; 72) | $2.886$* (df = 1; 67) |

*Note: $^*$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01

Finally, the empirical tests for the last form of gatekeeping bias: the speed of coverage, show our hypothesis is incorrect. There appears to be no significant difference between the speed of which a news organization covers an event of congressional misconduct dependent on party. The coefficients are very close to zero, with very little difference between news organizations.

**4.2 Presentation**
Political Actor References

Table 5: Republican Affiliation of Alleged Congressperson Effect on Frequency of Mention of Republican Actors

<table>
<thead>
<tr>
<th></th>
<th>Dependent variable:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Congressional Misconduct</td>
</tr>
<tr>
<td></td>
<td>WSJ</td>
</tr>
<tr>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Mentioning Republican Partisans More</td>
<td>0.530***</td>
</tr>
<tr>
<td></td>
<td>(0.103)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.214**</td>
</tr>
<tr>
<td></td>
<td>(0.082)</td>
</tr>
<tr>
<td>Observations</td>
<td>75</td>
</tr>
<tr>
<td>R^2</td>
<td>0.266</td>
</tr>
<tr>
<td>Adjusted R^2</td>
<td>0.256</td>
</tr>
<tr>
<td>Residual Std. Error</td>
<td>0.432 (df = 73)</td>
</tr>
<tr>
<td>F Statistic</td>
<td>26.398*** (df = 1; 73)</td>
</tr>
</tbody>
</table>

Note: *p<0.1; **p<0.05; ***p<0.01

Table 4: Effect of Party on Frequency of Mention of Republican Partisans

The regression shown in Table 4 points to my final hypothesis being correct.

Conservative partisan media will mention Republican partisans more in articles on congressional misconduct. The trend is statistically significant across both models, but the coefficient for Newsmax is noticeably higher. The R^2 values are also relatively high for both models, at 0.266 for the Wall Street Journal and 0.371 for Newsmax.

5 Discussion

The results of this study offer significant insights into the manifestation of partisan bias in media coverage, particularly in the context of congressional misconduct. While the initial
hypotheses presumed a certain direction of bias in conservative partisan media, the empirical findings paint a more nuanced picture, challenging conventional assumptions and prompting further exploration into the intricacies of media bias.

First and foremost, the unexpected finding that conservative partisan media covers instances of Republican misconduct more than Democratic misconduct introduces a critical divergence from anticipated patterns. This counterintuitive outcome may reflect the complex interplay of factors shaping media coverage, including political dynamics, audience preferences, and editorial decisions. One plausible explanation could be the strategic positioning of conservative media outlets vis-à-vis the broader political landscape. Given the heightened scrutiny faced by Republican politicians in mainstream media, conservative outlets might seek to counterbalance this narrative by providing extensive coverage of misconduct within their own ideological camp. Such strategic framing aligns with broader partisan objectives aimed at mitigating negative portrayals and reinforcing ideological solidarity among supporters.

Moreover, the findings concerning the extent of coverage underscore the nuanced nature of gatekeeping bias in media discourse. While traditional assumptions might suggest a bias towards downplaying misconduct within ideologically aligned factions, the empirical evidence suggests a more complex pattern characterized by varying degrees of coverage intensity. This nuanced approach to gatekeeping bias highlights the multifaceted nature of media agendas, which may prioritize certain narratives or events based on a combination of ideological alignment, audience engagement, and editorial discretion. In this context, the divergent coverage patterns observed between conservative partisan media and centrist sources underscore the dynamic interplay between partisan agendas and editorial decisions, shaping the overall media landscape.
Furthermore, the analysis of speed of coverage reveals a surprising lack of significant differences between news organizations in reporting on congressional misconduct based on party affiliation. This finding challenges conventional assumptions regarding the timeliness and agenda-setting power of partisan media outlets. While prior research has suggested a potential bias in the speed and frequency of coverage, the empirical evidence presented here suggests a more nuanced reality, where factors beyond partisan affiliation may influence the timing and prominence of media coverage. Future research could delve deeper into the underlying mechanisms driving the temporal dynamics of media coverage, exploring factors such as newsroom routines, editorial priorities, and audience engagement dynamics.

The analysis of presentation bias offers further insights into the nuanced strategies employed by partisan media outlets in framing narratives around congressional misconduct. The finding that conservative partisan media tends to mention Republican partisans more frequently in coverage of misconduct events underscores the strategic framing tactics employed by media outlets to reinforce ideological narratives and cultivate partisan identities among their audience. This strategic alignment with partisan agendas reflects the symbiotic relationship between media outlets and political actors, whereby media coverage serves as a vehicle for amplifying partisan messaging and shaping public perceptions.

In conclusion, the findings of this study shed light on the complex interplay of factors shaping partisan media bias in coverage of congressional misconduct. By challenging conventional assumptions and highlighting nuanced patterns of bias, this research contributes to a deeper understanding of media dynamics in contemporary political landscapes. Moving forward, further research is needed to explore the underlying mechanisms driving media bias, including the role of editorial decisions, audience preferences, and broader socio-political
dynamics. Only through a comprehensive understanding of these factors can we develop
effective strategies to promote balanced and informed news consumption in an era of increasing
partisan polarization.

5.1 Limitations

Firstly, the reliance on a specific set of events, namely congressional misconduct filings,
introduces inherent limitations in the scope and generalizability of the findings. While this
approach enables a focused analysis of media bias within a specific context, it may overlook
other forms of misconduct or political events that could provide valuable insights into broader
patterns of media coverage. Future research could explore alternative event datasets or adopt a
more comprehensive approach to capture a wider range of political phenomena, thus enhancing
the breadth and depth of the analysis.

Secondly, the study's focus on conservative partisan media outlets, specifically Newsmax,
may limit the applicability of the findings to other ideological contexts. Given the diversity of
partisan media landscape and the unique strategies employed by different outlets, the findings
may not fully capture the dynamics of media bias across ideological spectra. Future research
could adopt a comparative approach, examining bias across multiple partisan outlets and
ideological orientations to provide a more comprehensive understanding of media dynamics.

Additionally, while the utilization of computational methods offers scalability and
efficiency in analyzing large datasets, it also poses certain limitations in terms of interpretability
and contextual sensitivity. The reliance on automated algorithms for event-article classification
and bias detection may overlook nuanced patterns of bias or miss contextual cues that could
inform more nuanced interpretations. Future research could complement computational
approaches with qualitative analyses or expert assessments to validate findings and provide deeper insights into the underlying mechanisms driving media bias.

Moreover, the study's focus on quantifying bias in terms of selection, extent, and speed of coverage, while informative, may overlook other dimensions of bias such as framing, tone, or emphasis. By adopting a more holistic approach to bias measurement, future research could provide a more nuanced understanding of media dynamics and their impact on audience perceptions and political attitudes.

Lastly, the study's reliance on retrospective data and cross-sectional analysis limits its ability to establish causal relationships or capture dynamic shifts in media bias over time. Longitudinal studies tracking media coverage patterns over extended periods could offer valuable insights into temporal trends and the evolving nature of media bias in response to changing political dynamics, audience preferences, and technological advancements.

In conclusion, while this study provides valuable insights into the manifestation of partisan bias in media coverage of congressional misconduct, it is essential to acknowledge the inherent limitations and complexities involved in studying media dynamics. By addressing these limitations and adopting a multidimensional approach to bias measurement, future research can advance our understanding of media dynamics and contribute to the development of more effective strategies for promoting balanced and informed news consumption.

5.2 Further Research

Moving forward, several avenues of research could further advance our understanding of media bias and its implications for democratic processes and public discourse. Firstly, future research could explore the interplay between media bias and audience perceptions, investigating how individuals interpret and respond to biased media coverage. By integrating experimental
designs or survey methodologies, researchers could examine the impact of biased media narratives on attitudes, beliefs, and political behaviors, shedding light on the mechanisms through which media bias influences public opinion.

Secondly, longitudinal studies tracking media coverage patterns over extended periods could provide valuable insights into temporal trends and dynamics of media bias. By examining shifts in coverage patterns, editorial priorities, and partisan agendas over time, researchers could uncover underlying drivers of bias and identify potential interventions to mitigate its negative consequences.

Moreover, future research could adopt a comparative approach, examining bias across multiple partisan outlets and ideological orientations. By analyzing differences in framing, emphasis, and editorial decisions among various media sources, researchers could elucidate the nuanced dynamics of media bias and its implications for political polarization and democratic governance.

Additionally, there is a need for research exploring the role of social media platforms in shaping media bias and disseminating partisan narratives. Given the increasing reliance on social media for news consumption and political engagement, understanding how algorithms, echo chambers, and online communities influence the spread of biased information is essential for addressing the challenges posed by media polarization and disinformation.

Furthermore, interdisciplinary research collaborations between political scientists, communication scholars, computer scientists, and ethicists could enrich our understanding of media bias and its broader societal implications. By integrating insights from diverse disciplines, researchers can develop innovative methodologies, ethical frameworks, and policy recommendations to address the challenges posed by biased media coverage in the digital age.
In conclusion, future research endeavors should aim to deepen our understanding of media bias, its underlying mechanisms, and its impact on democratic processes and public discourse. By adopting multidisciplinary approaches, leveraging new technologies, and embracing innovative methodologies, researchers can contribute to the development of evidence-based strategies for promoting balanced and informed news consumption in an era of increasing media polarization and disinformation.

5.3 Conclusion

Contrary to initial hypotheses, the study found that conservative partisan media covers instances of Republican misconduct more extensively than Democratic misconduct. This unexpected finding underscores the complexity of media bias and highlights the strategic framing tactics employed by partisan outlets to shape political narratives and reinforce ideological identities among their audience.

Moreover, the study revealed nuanced patterns of gatekeeping bias, with conservative partisan media exhibiting differential patterns of coverage based on ideological alignment. While traditional assumptions might suggest a bias towards downplaying misconduct within ideologically aligned factions, the empirical evidence suggests a more complex reality characterized by varying degrees of coverage intensity.

Additionally, the analysis of presentation bias demonstrated that conservative partisan media tends to mention Republican partisans more frequently in coverage of misconduct events, further highlighting the strategic alignment of media outlets with partisan agendas.

Overall, this study contributes to a deeper understanding of media bias and its implications for democratic processes and public discourse. By challenging conventional
assumptions and uncovering nuanced patterns of bias, the study underscores the importance of adopting a multidimensional approach to bias measurement and analysis.

Moving forward, further research is needed to explore the underlying mechanisms driving media bias, including the role of editorial decisions, audience preferences, and broader socio-political dynamics. By addressing these gaps in knowledge, researchers can develop evidence-based strategies for promoting balanced and informed news consumption in an era of increasing media polarization and disinformation.

In conclusion, this study underscores the importance of critically examining media bias and its implications for democratic governance, highlighting the need for interdisciplinary collaboration and evidence-based interventions to address the challenges posed by biased media coverage in contemporary societies.
References


*Media Bias Chart | AllSides.* (n.d.). Retrieved February 11, 2024, from [https://www.allsides.com/media-bias/media-bias-chart](https://www.allsides.com/media-bias/media-bias-chart)

Mitchell, A., Simmons, K., Matsa, K. E., & Silver, L. (2018). *Publics globally want unbiased news coverage, but are divided on whether their news media deliver.*


