PNAS BRIEF REPORT | POLITICAL SCIENCES
OPEN ACCESS

Measuring dynamic media bias

Eunji Kim[a], Yphtach Lelkes[b,1], and Joshua McCrain[c]

Edited by Mary Waters, Harvard University, Cambridge, MA; received February 8, 2022; accepted June 1, 2022

Ideological media bias is increasingly central to the study of politics. Yet, past literature often assumes that the ideological bias of any outlet, at least in the short term, is static and exogenous to the political process. We challenge this assumption. We use longitudinal data from the Stanford Cable News Analyzer (2010 to 2021), which reports the screen time of various political actors on cable news, and quantify the partisan leaning of those actors using their past campaign donation behavior. Using one instantiation of media bias—the mean ideology of political actors on a channel, i.e., visibility bias—we examine weekly, within-day, and program-level estimates of media bias. We find that media bias is highly dynamic even in the short term and that the heightened polarization between TV channels over time was mostly driven by the prime-time shows.

media bias | political communication | news

With the transformation of the media environment in the post broadcast democracy era, many media outlets now actively filter information or unequivocally distort facts. This ideological bias, often linked to political behaviors as well as public health outcomes (1, 2), has become indispensable to the study of American politics. But quantifying the media bias across outlets, and more importantly, over time, has immense methodological challenges.

We focus on measuring the media bias in cable TV news coverage over the last decade. Leveraging outlets' explicit endorsements, issue coverage, or linguistic patterns to measure media bias works well when we assume both the drivers of media bias—ranging from the ideological leanings of the media firms to the market structure—and the bias itself are fixed attributes. Yet these approaches are ill-suited to fully capture the dynamic nature of media bias in this evolving media environment. In a competitive, multiplatform market, the audience composition is no longer stable. Instead of waiting for reports on yearly ratings and circulation numbers, editors and producers can react quickly to audience demand. In other words, despite all the evidence that suggests that media bias is dynamic even in a short term, there have been few systematic attempts to quantify it using a scalable method.

Text-scaling approaches are often used in the literature to measure slant. These approaches face two shortcomings if they are to be used to measure slant dynamically and at scale. First, scaling approaches require collection, cleaning, and analysis of both the television transcripts and the reference text that are computation- and time-intensive tasks (and may be impractical if the source text is difficult to acquire). To accurately assess time-varying slant, text scaling would require repeated collection and analyses of both sources of text, especially as the reference text may quickly be out of date and no longer salient. Second, the n-grams identified in scaling methods often are obscure phrases, with no obvious latent ideological content, and are subject to substantial researcher degrees of freedom based on preprocessing methods (3).

To address this challenge, we develop a highly scalable, longitudinal measure of media bias, assembled by bridging data generated by machine learning and ideal point estimation. Our measure exploits the visibility of political actors, relying on the intuition that if an outlet is more or less likely to feature actors from the left than from the right, then that outlet is considered more liberal, which is sometimes called visibility bias (4). By examining actual appearances of all political actors on cable news outlets over the last decade, we sidestep the problems of indirect measurement inherent to existing studies and sparsity of data that prevented researchers from exploring the dynamics of changing media bias. Of course, because some actors appear frequently across channels because of their position in government, e.g., Donald Trump and Joe Biden, we use several different specifications of visibility, including one that does not include any politician.*

*Additional specifications and models appear in the public repository.

Author affiliations: [a]Department of Political Science, Columbia University, New York, NY 10027; [b]Annenberg School for Communication, University of Pennsylvania, Philadelphia, PA 19104; and [c]Department of Political Science, University of Utah, Salt Lake City, UT 84112

Author contributions: E.K., Y.L., and J.M. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no competing interest.

[1]To whom correspondence may be addressed. Email: ylelkes@upenn.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2202197119/-/DCSupplemental.

Published August 1, 2022.

Materials and Methods

We quantify a measure of media bias that relies on visibility of partisan actors (4–6). We use the Stanford Cable TV News Analyzer,[†] a comprehensive dataset that includes a decade worth of video—over 280,000 h of content—of CNN, Fox News, and MSNBC (1 January 2010 to 4 March 2021). Per-video metadata include the program names, dates/times they aired, and most importantly, the names of individuals who appeared on screen with how much time they were on screen. The screen time of individuals is calculated by first processing the video using a facial recognition algorithm, using faces in the Amazon Rekognition Celebrity Recognition Application Programming Interface (API). The sample is filtered only to include individuals with at least 10 h of screen time by August 2020 (cumulative).[‡] In total, there are 1,584 people for whom the TV News Analyzer counts screen time. Among these individuals, we first removed television anchors from the dataset, since they account for the bulk of screen time per show. We omit anchor ideology as it would overwhelm other actors, and most anchors do not donate to political campaigns and so do not have donation-based ideological scores. We then created a name-matching algorithm to match these names to the Database on Ideology, Money in Politics, and Elections (DIME) (updated to 2018) (7).[§]

Next, we matched the contributor list from DIME to the names in the Stanford database. This process first entailed a naive match of each person by first name, last name, and gender. With a team of research assistants, we then manually examined all 19,000 raw matches, including using multiple coders per name to rule out multiply matched names and increase confidence in the matches. We used additional information in the DIME data, such as donor's occupation, employer, and place of residence, to examine matches. We conducted a similar exercise for unmatched names using only the last name and the first initial of the first name, accounting for common nicknames, and manually checked the matches again. This process resulted in 704 total matches (of 1,225 nonanchors), 327 of which were politicians matched to the donation recipient data. Finally, we extracted data from the Stanford API at the daily-channel-person level, creating a channel-person-day dataset that includes information on each actor's screen time.

To check the robustness of the measure, we also create different versions of the measure—ones that are weighted by screen time, that exclude Trump, that exclude all politicians, or that exclude all individuals who were not likely to have been actual guests on a program (i.e., Robert Mueller). We also take advantage of the structure of the Stanford data and estimate within-day media bias by specific time blocks by channel—the morning block (8 to 10 AM), the afternoon block (1 to 3 PM), and the primetime block (7 to 9 PM).

Measurement Validation. Our primary measure of media bias is created by weighting a channel-day's DIME campaign finance (CF) score based on how much time an individual is on the screen.[¶] For example, if an individual is on the screen for 100 s and has a CF score of 0.5, this results in a weighted score of 50. We aggregate up these weighted scores per channel day and then take the weekly average per channel. To interrogate the validity of this measure, we provide basic descriptive patterns below. First, the measure appears to have face validity: the raw densities of the weighted CF score per channel, displaying remarkable similarity between CNN and MSNBC (weighted CF score of −9.7 and −14.1, respectively), and a much more right-leaning Fox News (weighted CF score of 49.8). Fig. 1 displays the mean level of the weighted CF scores for the top five programs per channel.[||] Programs on Fox News are on average more conservative than programs on CNN and MSNBC. Our program-level measures reveal, at least based on the guests that appear on the show, popular primetime news shows on CNN—such as *Anderson Cooper 360* or *CNN Tonight*—are more left leaning than well-known evening shows on MSNBC such as *The Rachel Maddow Show* or *The 11th Hour with Brian Williams*.

Because our measure of media slant relies heavily on the DIME estimation process, interrogating the validity of this measure requires demonstrating that such media bias is "visible"; we need to show whether the utterances of a speaker with a high CF score, when read with plain eyes, are deemed more conservative than those of a speaker with a low CF score. We employed two research assistants to sample 978 news transcripts and broke them up by speaker into 19,582 blocks of text. We then assessed whether a block of text demonstrates a liberal, neutral, or conservative slant. To prevent the possibility that research assistants' prior knowledge about partisan leaning of certain actors biases the coding results, we anonymized all actors from the transcripts. The resulting exercise generated moderately high intercoder reliability, with weighted Cohen's Kappa values of 0.664, 0.55, and 0.66 for CNN, MSNBC, and Fox News, respectively. The correlation between our primary measure of media bias and the human coded estimate was 0.72 (see *SI Appendix* for details).[**] Based on these correlations, we argue that our main measure is qualitatively better, or at least comparable, to the measures based on text-scaling–based methods adopted in previous literature based on face validity and construct validity. For instance, Gentzkow and Shapiro (8) report a correlation of 0.40 between their text-scaling approach and a hand-coded measure of newspaper ideology. The correlation between our measure and a text-scaling measure (9) was 0.59. As an additional validation, we compare the weighted ideology scores of members of congress that appear both in Padgett et al. (10) and in our data.

Descriptives. The granularity of our measure allows us to offer a richer set of descriptive patterns. First, we can examine the dynamics of polarization by calculating the distance between the weekly means of media bias, comparing each

[†]https://tvnews.stanford.edu/.

[‡]Additional details on the methodology for extracting screen time are available on the Stanford Cable TV News Analyzer website: https://tvnews.stanford.edu/methodology. Importantly, the Stanford team removes commercials from the sample and conducted manual checks of the accuracy of the facial recognition algorithm, finding 98.5% precision. Our own random checks of individuals found similarly high accuracy.

[§]DIME uses campaign donations to scale individuals (everyone who makes a donation) on an ideological space similar to dynamic weighted nominal three-step estimation (DW-NOMINATE) and has shown, within office holders, to correlate highly with such measures. See ref. 7 for further details.

[¶]The average number of weekly minutes used per each channel to create the weighted ideology measure is 1,192 for CNN, 1,391 for Fox News, and 1,538 for MSNBC.

[||]We used the Nielsen ratings from 2012 to 2020. We selected the top five most popular programs per channel that have been aired more than 3 consecutive years in this time period.

[**]Granted, we note that the correlations between the alternative measures of media bias and the human coding were slightly lower. A measure of media bias that includes only nonpoliticians and, separately, one that excludes Trump had a correlation of 0.65 and 0.68, respectively.

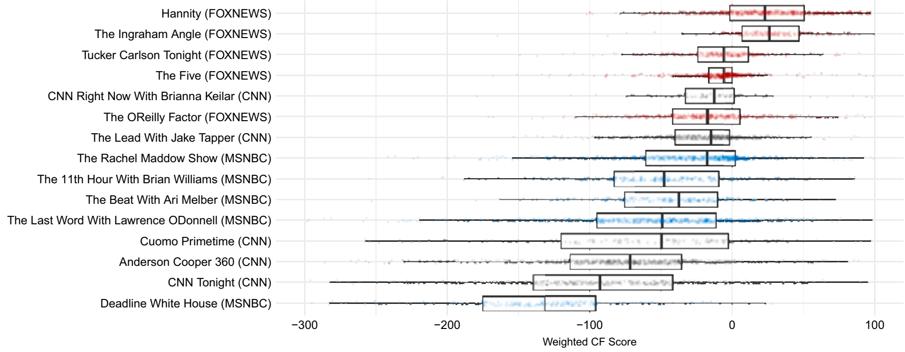

Fig. 1. Program-level media bias. Shown is the distribution of ideology score for the top five most popular programs from CNN, Fox News, and MSNBC (2010 to 2021).

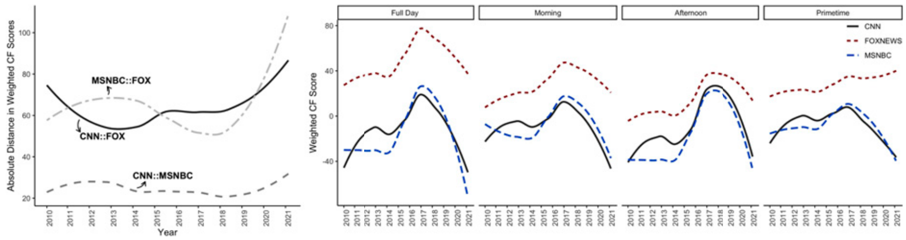

Fig. 2. (*A*) Distance between channels. (*B*) Within-day cable TV media bias by channel.

channel with the other channels. As Fig. 2*A* shows, while the distance between CNN and MSNBC has remained mostly unchanged over time, the relationship between Fox News and the other channels has polarized—growing increasingly distant from one another. This pattern of polarization started a bit earlier between Fox News and CNN than between MSNBC and Fox News, although the polarization accelerated for MSNBC and Fox News starting in 2018. In Fig. 2*B*, we first plot the distribution of weighted CF scores by channel and over time. While Fox News is always to the right of CNN and MSNBC, CNN and MSNBC become more in sync over time. Between 2011 and 2015, CNN was consistently to the right of MSNBC, which is no longer the case after 2015. Additionally, the channels seem to mirror the public mood—the channels shift to the right, in parallel when Republicans gain power in 2016, but shift back to the left following the Republican defeat in the midterm elections. The correlations between these trends are fairly high ($r_{Fox,CNN} = 0.68$; $r_{Fox,MSNBC} = 0.72$; $r_{MSNBC,CNN} = 0.68$). These results are similar even if we examine only scores that do not contain politicians (and, therefore, are not driven by prevalence of administration officials on television).

We also see interesting patterns by timeslot (afternoon, when most hard news is aired; morning, when talk shows appear, such as *Morning Joe* or *Fox & Friends*; and primetime, when popular shows air such as *The Rachel Maddow Show*, *Hannity*, and *Tucker Carlson Tonight*). Scores in the afternoon slot—when hard news shows air—are far more correlated between channels than in other time slots. While news shows across all time blocks have polarized since 2017, we find that this pattern is the most pronounced among primetime shows, with Fox News steadily becoming more conservative and MSNBC and CNN sharply becoming more liberal during the Trump presidency. The fact that morning shows on Fox News leaned more conservative than prime-time shows is consistent with the evidence that then President Trump was a regular and avid consumer of *Fox & Friends*. The granularity of our measure also allows us to see the variations in ideology in each of these programs over time. We find, for instance, that *The Rachel Maddow* Show became more liberal through the Trump presidency and that *The O'Reilly Factor* quickly turned farther right until the Fox News Channel expunged him from the channel (Fig. 3).

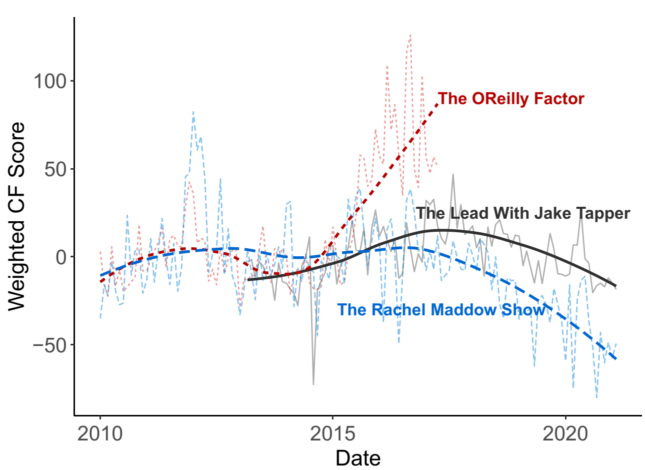

Fig. 3. Over-time variations in cable TV media bias by program.

Conclusion

We believe that this measure offers a highly scalable method for estimating bias in television outlets and programs. By constructing the most granular measure of media bias to date, we show that, unlike the conventional assumptions built in existing theories of media bias, the media bias is highly dynamic even in a short term. Importantly, our measure opens the door to answering key questions in the social sciences. We envision that this measure can be particularly useful for studying the supply side of media environment—ranging from the dynamics of market competition to the ownership change to business pressures or consumer boycotts—as well as understanding the demand side of of media bias (11, 12). For instance, to what extent is media bias driven by audience ratings? Finding such a link would not only help us understand the political economy of partisan media, but also validate, in and of itself, the dynamic nature of our media bias measure. This measure can also be used to better understand the influence of the media by combining it with longitudinal public opinion and elite behavior data. The temporal variations in our media bias measure offer a unique opportunity to probe why, when, and under what conditions media bias is amplified. Finally, this measure will be publicly accessible and straightforward to extend once additional data (screen time or donations) become available, facilitating rapid analysis of timely research questions.

Data Availability. Replication data and code files are available in the Harvard Dataverse. (The code is an R file, and the dataset is an RData. There is also a codebook and online appendix.) (https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/I2K54B) (13).

ACKNOWLEDGMENTS. We are grateful to Mike Barber, Josh Clinton, Josh Gubler, Gregory Martin, Matt Miles, Kevin Reuning, John Sides, James Snyder, Christopher Warshaw, participants at the American politics workshop at Brigham Young University, and the anonymous reviewers. We thank Suiyang Mai and Emily Tompkins for excellent research assistance.

1. G. J. Martin, A. Yurukoglu, Bias in cable news: Persuasion and polarization. *Am. Econ. Rev.* **107**, 2565–2599 (2017).
2. S. DellaVigna, E. Kaplan, The Fox News effect: Media bias and voting. *Q. J. Econ.* **122**, 1187–1234 (2007).
3. M. J. Denny, A. Spirling, Text preprocessing for unsupervised learning: Why it matters, when it misleads, and what to do about it. *Polit. Anal.* **26**, 168–189 (2018).
4. J. M. Eberl, H. G. Boomgaarden, M. Wagner, One bias fits all? Three types of media bias and their effects on party preferences. *Communic. Res.* **44**, 1125–1148 (2017).
5. T. Groseclose, J. Milyo, A measure of media bias. *Q. J. Econ.* **120**, 1191–1237 (2005).
6. R. Durante, B. Knight, Partisan control, media bias, and viewer responses: Evidence from Berlusconi's Italy. *J. Eur. Econ. Assoc.* **10**, 451–481 (2012).
7. A. Bonica, Mapping the ideological marketplace. *Am. J. Pol. Sci.* **58**, 367–386 (2014).
8. M. Gentzkow, J. M. Shapiro, What drives media slant? Evidence from US daily newspapers. *Econometrica* **78**, 35–71 (2010).
9. G. J. Martin, J. McCrain, Local news and national politics. *Am. Polit. Sci. Rev.* **113**, 372–384 (2019).
10. J. Padgett, J. L. Dunaway, J. P. Darr, As seen on TV? How gatekeeping makes the US House seem more extreme. *J. Commun.* **69**, 696–719 (2019).
11. H. J. Hassell, M. R. Miles, K. Reuning, Does the ideology of the newsroom affect the provision of media slant? *Polit. Commun.* **39**, 184–201 (2021).
12. M. Gentzkow, J. M. Shapiro, D. F. Stone, "Media bias in the marketplace: Theory" in *Handbook of Media Economics*, S. Anderson, J. Waldfogel, D. Stromberg, Eds. (Elsevier, 2015), vol. 1, pp. 623–645.
13. J. McCrain, E. Kim, Y. Lelkes, "Replication Data for: Measuring Dynamic Media Bias." Harvard Dataverse. https://doi.org/10.7910/DVN/I2K54B. Deposited 21 June 2022.