

Media Slant is Contagious

Working Paper**Author(s):**

Widmer, Philine; Galletta, Sergio; Ash, Elliott

Publication date:

2020-10

Permanent link:

<https://doi.org/10.3929/ethz-b-000454192>

Rights / license:

In Copyright - Non-Commercial Use Permitted

Originally published in:

Center for Law & Economics Working Paper Series 14/2020

Center for Law & Economics Working Paper Series

Number 14/2020

Media Slant is Contagious

**Philine Widmer
Sergio Galletta
Elliott Ash**

October 2020

Media Slant is Contagious^{*}

Philine Widmer¹, Sergio Galletta^{2,3}, Elliott Ash²

¹*University of St. Gallen*

²*ETH Zürich*

³*University of Bergamo*

Abstract

This paper provides causal evidence on how partisan news messaging from cable television influences the content published by newspapers in U.S. localities. We introduce a new parallel corpus of newspaper articles (24M articles in 600+ local newspapers) and transcribed television news shows (40K cable news episodes from Fox News Channel, CNN, and MSNBC) for the years 2005-2008. We measure media influence using a supervised learning model that predicts, for a given piece of text, the probability that it comes from a Fox News transcript, rather than from CNN or MSNBC. After validating the measure, we apply it to the local newspaper article texts. Exogenous variation in news viewership across localities comes from relative channel numbering, which we use as instruments. We find that an exogenous increase in local viewership of a cable news network shifts the textual content of local newspapers toward that network's content. Televised media slant works not just through persuading viewers, but through influencing other media outlets.

Preliminary working paper. Comments welcome.

*We thank for helpful comments and suggestions Ruben Durante, Claudio Ferraz, Roland Hodler, Eunji Kim, Ro'ee Lee, Josh McCrain, Giovanni Prarolo, Pia Raffler, Jim Snyder, David Yang, and seminar and conference participants at the Harvard Political Economy Seminar, Econometric Society World Congress, Bologna Workshop on Elections, Democracy, and Populism, University of St. Gallen Economics Seminar, and Zurich Text as Data Workshop. Romina Jafarian and Matteo Pinna provided excellent research assistance.

Corresponding author: Philine Widmer.

Email addresses: philine.widmer@unisg.ch (Philine Widmer), sergio.galletta@unibg.it (Sergio Galletta) ashe@ethz.ch (Elliott Ash).

1. Introduction

Does exposure to politically biased news messaging distort the news output of other media organizations? This is an important question for social scientists and for policymakers making decisions about media regulation. While the literature shows that partisan media affect voting by citizens and politicians, there is no previous evidence on whether and how national (slanted) media influence media content at the local level.

We seek to provide evidence on this issue using a new corpus of 24 million article snippets from 600+ local newspapers in the United States for the years 2005 through 2008. We combine these texts with a parallel corpus of the full-text transcripts from news shows on Fox News Channel (FNC), CNN, and MSNBC. We add rich metadata on newspaper circulation, television channel positioning and ratings, and political and demographic covariates.

The text data are used to measure the linguistic similarity between cable news and local newspapers. We introduce a novel measure of news influence using a supervised learning model, which produces a prediction about whether a given body of text is from a relatively conservative network (FNC), rather than from a relatively liberal network (CNN or MSNBC). We validate the model and associated predictions.

The model is applied to the corpus of local newspaper articles. For each article, we then have a text-based metric that indexes similarity to language from FNC shows (relative to shows on the other networks). We then aggregate that measure to the level of the newspaper, to get a novel measure of partisan news slant for each newspaper.

We then ask if relative similarity to language in a cable news network increases in response to higher viewership in a newspaper's market. Cross-sectional or panel data estimates of this relationship would likely be biased. For example, estimates might be confounded if more conservative(-trending) counties have higher Fox News viewership and more conservative(-trending) local news reporting. Hence, to obtain causal estimates, we exploit exogenous variation in news exposure across counties coming from random variation in the relative channel numbering of the three cable networks ([Martin and Yurukoglu, 2017](#)). We provide a number of checks to validate first stage relevance and exogeneity of the instrument. In particular, the instrument (channel position) is uncorrelated with other local characteristics that are predictive of viewership or are predictive of the relevant dimensions of local newspaper content.

According to our estimates, higher cable news network viewership increases the in-

fluence of a network’s content on local newspaper articles. A one standard deviation increase in viewership increases similarity to a network’s content by about one-third of a standard deviation. This result is robust to a number of specification checks, including controls for local demographics, for local cable television market characteristics, and for text readability metrics (e.g. word length).

Recent discussions in media politics have emphasized the potential for cable news to polarize discourse, given that Fox News is relatively right-wing while CNN and (especially) MSNBC are relatively left-wing (e.g. [Martin and Yurukoglu, 2017](#)). We explore this issue in our context by comparing heterogeneous effects on newspapers based on their history of candidate endorsements. That is, we split newspapers into three groups: those who have historically endorsed Democrats, those who have historically endorsed Republicans, and those without or with mixed endorsements. We find that exposure to Fox News (relative to CNN/MSNBC) tends to polarize language: Historically Republican newspapers become more conservative (FNC-like), while historically Democrat newspapers become more liberal (CNN/MSNBC-like). Media slant from cable news is selectively contagious, in a way that increases political polarization in local news discourse.

The findings add to the literature in political science and political economy on biased media (e.g. [Ashworth and Shotts, 2010](#); [Prat, 2018](#)).¹ This literature has produced good evidence that mass media have an impact on election outcomes. [Gentzkow et al. \(2011\)](#) report that the opening of local newspapers boosts voter turnout, while [Chiang and Knight \(2011\)](#) show that a newspaper endorsement for a presidential candidate shifts voting intentions in favor of this candidate. For Fox News in particular, [DellaVigna and Kaplan \(2007\)](#) and [Martin and Yurukoglu \(2017\)](#) document that a quasi-experimental increase in exposure leads to higher Republican vote shares. Cable news can affect voter knowledge about political issues ([Hopkins and Ladd, 2014](#); [Schroeder and Stone, 2015](#)) as well as fiscal policy decisions ([Galletta and Ash, 2019](#)). Outside the US, [Enikolopov et al. \(2011\)](#) find that Russian voters with access to an independent television station are more supportive of anti-Putin parties.²

We contribute to these strands of the literature by studying a potential channel

¹For surveys on the empirical and theoretical literature, see [Puglisi and Snyder \(2015\)](#) and [Gentzkow et al. \(2015\)](#), respectively.

²See [Prat and Stroemberg \(2013\)](#) and [Stroemberg \(2015\)](#) for surveys on the political effects of mass media.

through which partisan media may affect electoral outputs, policies, and political knowledge: local newspapers. Local news have been shown to be important in increasing citizen engagement and allowing citizens to hold their political representatives accountable (e.g. [Snyder and Strömberg, 2010](#)). [George and Waldfogel \(2006\)](#) find that the market entry of a national media outlet (in their case the New York Times) causes local outlets to focus more on local coverage. [Martin and McCrain \(2019\)](#) show that the acquisition of US local TV stations by the national conglomerate owner Sinclair led to an increased share of national as opposed to local content. Similarly, [Mastrorocco and Ornaghi \(2020\)](#) document that these acquisitions by Sinclair lead to lower coverage of local crime and subsequently to lower violent crime clearance rates. We contribute to these debates by analyzing if higher exposure to national cable news changes the slant of local news content.

Methodologically, this research is innovative in its combination of natural language processing, machine learning, and causal inference, extending the use of text data to understand partisan influences in the media. This literature includes [Gentzkow and Shapiro \(2010\)](#), [Gentzkow et al. \(2019\)](#), and [Ash et al. \(2017\)](#), who analyze divisiveness in congressional language. Our paper is the first to combine text-based methods with an instrumental-variables framework to analyze the diffusion of political messaging. The methods could be useful for economists seeking to use text in a causal framework. As detailed below, we address a number of issues in terms of high dimensionality, lack of interpretability, and omitted variable bias.

More broadly, our work contributes to the long-lasting debate on the importance of (un)biased media in democratic politics – a topic that has become especially important in the current era of extreme polarization and increasing inequality in American society.

2. Data

This section enumerates our data sources. The data come from cable news channels and from local newspapers. Our resulting panel is from 2005 through 2008, the years for which we can construct cable news viewership by locality. Summary statistics for all variables are reported in Appendix Table [A2](#).

Figure 1: Example of a local newspaper article snippet

Alameda Times-Star

County outlines ways to lower shelter hostility

8 March 2005

Can Alameda County blunt opposition to current plans to permit emergency homeless shelters at hundreds of residential locations in unincorporated communities? That appeared likely Monday as county planners suggested ways in which shelters - such as in the land-use game Monopoly - would not automatically pass go and neighbors could voice their approval or opposition. [...]

2.1. Local newspaper article excerpts

The first ingredient in our analysis is a corpus of local newspaper articles. Our data source is the news aggregation site NewsLibrary, from which we scraped the headlines and first 80 words of all published articles for various local U.S. newspapers for 2005-2008. With a set of scripts, we read through the snippets and extract the newspaper name, the title and the plain text of the article, as well as the date. An example of an article snippet is shown in Figure 1. Our main dataset contains 16 million article snippets of on average 80 words each (starting from the beginning of the article) from 305 unique newspaper titles. In the robustness checks, we use a larger dataset of 24 million articles from 682 titles. More information is contained in Appendix C.1.

2.2. News show transcripts for FNC, CNN, and MSNBC

Our second corpus is from the cable news networks. We gather the news show transcripts for FNC, CNN, and MSNBC from LexisNexis. The corpus includes transcripts from around 40,000 episodes of prime time shows for the three networks for 2005-2008. We have a series of scripts that reads through the transcripts and filters out metadata and other non-speech content.

As mentioned, the newspaper article snippets contain approximately the first 80 words of the article. The transcripts tend to be much longer. To make the corpora more comparable, we segment the transcripts into shorter 80-word snippets to match the length of the newspaper article snippets.

2.3. Newspaper-level circulation data

Next, we match each local newspaper outlet to one or more counties. We use audited county-level circulation data from the Alliance for Audited Media, which is available for around 305 unique newspapers (that also appear in the NewsLibrary and the Nielsen channel ratings data). We thus have 3,781 observation units at the newspaper-county level (see Section 4) for our main analyses. Appendix C.2 describes an alternative matching method that results in 682 observation units (that we use for robustness checks).

2.4. Channel positions and viewership

From Nielsen we have data on channel positions and ratings for Fox News Channel, CNN, and MSNBC. These are the same as the data used by [Martin and Yurukoglu \(2017\)](#). First, we have the channel lineup for all the U.S. broadcast operators and the respective zip code areas served. Second, we avail of viewership information representing the share of individuals tuned in to each channel by zip code. This value is proportional to the average number of minutes spent watching a channel per household. As the original data are at the zip code level we follow [Galletta and Ash \(2019\)](#) and aggregate both the ratings and the channel positions at the county level. Specifically, we create county average channel positions weighting the observations by population size in the zip code, while we weight ratings by the number of survey individuals in the zip code according to Nielsen.

2.5. Other demographic covariates

Finally, we have a rich set of demographic covariates from the 2000 census. These variables are measured at the zip code level. To get the aggregate value for the county we weight them by zip code population. Appendix Table A2 lists these variables along with summary statistics.

3. Measuring Media Slant

This section describes how we construct the language measures to be used as outcome variables in our regression analysis. We aim at capturing textual similarity between the newspaper article snippets on the one hand and TV show transcripts on the other. We therefore implement a supervised machine learning approach to predict if newspaper article language resembles that from a particular TV station (FNC or CNN and MSNBC).

The approach is related to [Gentzkow et al. \(2019\)](#), who also use a regularized linear model with n-gram inputs.³

3.1. Text pre-processing and featurization

First, we preprocess all texts (newspaper articles and TV station transcripts). We convert them to lower case and remove non-meaningful stopwords (like *and* or *or*), all non-letter characters, and extra white-spaces. Second, for each word, we perform stemming (employing the Porter stemming algorithm). Finally, we form bigrams (two-word phrases) from the word stems.

Let M be the set of documents (snippets indexed by m) from the transcripts corpus. For now, we group CNN and MSNBC together (for notational simplicity, we refer to the CNN/MSNBC label as CNN). Thus the label we will predict is FNC_m : For each transcript snippet m , $FNC_m = 1$ if it comes from a Fox News transcript and $FNC_m = 0$ otherwise (if it comes from CNN/MSNBC). We produce a balanced sample of documents, with half from Fox, a quarter from CNN, and a quarter from MSNBC.

Let V_k give the vocabulary of bigrams used by a given channel $k \in \{FNC, CNN\}$. Let F_k^b be the frequency of bigram b on channel k . We construct V_{FNC} and V_{CNN} and then intersect the two, imposing the condition that any bigram b must appear more than 20 times in *both* corpora. The resulting set of bigrams is denoted as

$$V = \{b \in V_{FNC} \ \& \ b \in V_{CNN} \mid F_{FNC}^b > 20 \ \& \ F_{CNN}^b > 20\}$$

The frequency threshold serves to exclude infrequent bigrams which are highly distinctive for a given channel (such as a moderator’s name or a show name), but carry little substantive political or topical information. This procedure produces a vocabulary V with 65,000 bigrams.

3.2. Classifying transcripts by TV source

Now we will train a machine classifier to predict whether a transcript snippet m comes from FNC or CNN/MSNBC. For evaluating the performance of the classifier, we

³Our different approach reflects a different scientific objective. [Gentzkow et al. \(2019\)](#) are interested in measuring the level of polarization between groups in language. We are interested in forming a predicted probability of the source of a document. Our penalized logistic regression cannot be parallelized by phrase, but it gives better test-sample fit based on our own experiments ([Gentzkow et al. \(2019\)](#) do not report test-set accuracy).

shuffle the corpus and split into 80% training data and 20% test data. We build the classifier in the training set and evaluate its performance in the test set.

We take two steps to further pre-process the features. This is done using the training set to ensure a clean evaluation in the test set. First, we do supervised feature selection to reduce the dimensionality of the predictor matrix. Out of the 65,000-bigram dictionary, we select the 2,000 most predictive features based on their χ^2 score for the true label FNC . Second, we scale all predictors in S to variance one (we do not take out the mean, however, as then we would lose sparsity). Let S be the vector of selected and scaled features, indexed by s . Let B_m^s be the frequency of bigram s in transcript m (and B_m the vector of frequencies for transcript m , of length $|S| = 2000$).

Our principal classification method is penalized logistic regression (Hastie et al., 2009). We parametrize the probability that a transcript is from Fox News as

$$\widehat{FNC}_m = \Pr[FNC_m = 1|B_m] = \frac{1}{1 + \exp(-\psi' B_m)}$$

where ψ is a 2000-dimensional vector of coefficients on each feature. The L1-penalized logistic regression model chooses ψ to minimize the cost objective

$$J(\psi) = -\frac{1}{M^*} \sum_{m=1}^{M^*} \left(FNC_m \log(\widehat{FNC}_m) + (1 - FNC_m) \log(1 - \widehat{FNC}_m) \right) + \lambda |\psi| \quad (1)$$

where M^* gives the number of documents in the training sample.

The rightmost term in Equation (1) is the regularization penalty. We use L1 (Lasso) regularization, as indicated by the L1 norm $|\cdot|$.⁴ Lasso mitigates over-fitting of the training set by shrinking coefficients (and setting small ones to zero).⁵ Regularization strength is calibrated by the hyperparameter $\lambda \geq 0$, selected using five-fold cross-validated grid search in the training set. The optimal penalty in our data is $\lambda^* = 2$, although we got almost identical performance with larger or smaller penalties.

We evaluated classifier performance in the test set, obtaining an accuracy of 0.73 (with a standard deviation of 0.02 across five folds). This performance is much better than guessing (which would produce accuracy = 0.5 in the balanced sample) and comparable

⁴The L1 norm of a vector is the sum of the absolute values of the vector's elements.

⁵We obtained similar prediction accuracy when using an L2 (Ridge) penalty.

Table 1: Test-Set Prediction Performance for Identifying Cable News Source

	Predicted CNN	Predicted FNC
Actual CNN	38.3% (235K)	11.7% (72K)
Actual FNC	15.0% (92K)	35.0% (215K)

Notes: Confusion matrix for test-set predictions. Top left gives true positives for the CNN/MSNBC class; bottom right gives true positives for the FNC class; top right gives false negatives for CNN/MSNBC; bottom left gives false negatives for FNC.

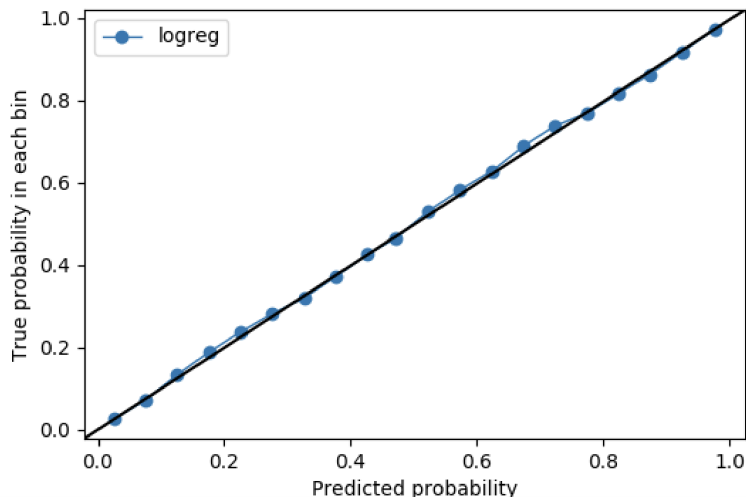
with other works in this literature.⁶ Table 1’s confusion matrix demonstrates the good performance in terms of precision and recall across the two categories. The on-diagonal cells have most of the mass and are quite balanced. Looking at the off-diagonals, we can see that FNC transcripts are somewhat more likely to be classified as CNN (compared to the converse, CNN transcripts being classified as FNC). This error tendency could mean that a subset of FNC transcripts take a straight-news approach (and are classified as CNN).

Figure 2 reports the calibration plot for our predictions. The figure shows – for the test set – the binned means (rates) of coming from the Fox News transcripts, conditional on the predicted probability from our model. The 45 degree line indicates how the line would look if the model replicated the distribution in the data – for example, for the set of observations with about 30% predicted probability, we see that about 30% of them are truly from Fox News transcripts. Similarly, this holds for all twenty bins (of 5% increments). As can be seen in the figure, the fit is remarkably good. The conditional predicted rates are almost perfectly on top of the 45 degree line.

To understand better how the model is making these predictions, we examine the bigram features that are most important for its classification choices. An advantage of logistic regression in this respect is its interpretability: The trained model’s estimated coefficients $\hat{\psi}$ provide a ranking across the 2,000 predictive bigrams in terms of their relative contribution to the predictions. Because the predictors are standardized to the same variance, the coefficients are comparable and roughly interpretable as the relative marginal effect of the associated bigram on the predicted probability that a document

⁶For example, the prediction accuracy for partisan affiliation in U.K. parliament obtained by [Peterson and Spirling \(2018\)](#): They obtain an accuracy of between 0.6 and 0.8, depending on the time period in the data. [Kleinberg et al. \(2017\)](#) obtain an AUC of .71 in predicting recidivism from criminal defendant characteristics.

Figure 2: Model Predicted Probabilities Accurately Reproduce Test-Set Distribution



Notes: Calibration plot for our predictions. The figure shows – for the test set – the binned means (rates) of coming from the Fox News transcripts, conditional on the predicted probability from our model. The 45 degree line indicates how the line would look if the model replicated the distribution in the data perfectly – for example, for the set of observations with a 30% predicted probability, exactly 30% of them are truly from Fox News transcripts. Similarly, this holds for all twenty bins (of 5% increments). As can be seen in the figure, the fit is remarkably good. The conditional predicted rates are almost perfectly on top of the 45 degree line.

Table 2: Distinctive phrases associated with Fox News and CNN/MSNBC

FNC	CNN/MSNBC
american troop	gay communiti
crime punish	crime alleg
immigr author	politico report
border mexican	usmexico border
nation economi	economi world
support gun	woman black

Notes: Examples of bigrams with positive (predictive for FNC transcripts) or negative (predictive for CNN/MSNBC) coefficient values in the penalized logistic regression of a label equaling one for FNC snippets (and zero for CNN/MSNBC snippets) on the bigrams used in a snippet.

is from Fox News.

Table 2 shows examples of bigrams with positive (predictive for FNC transcripts) or negative (predictive for CNN/MSNBC) values of $\hat{\psi}_s$.⁷ Fox bigrams allude to intuitively conservative priorities such as the troops, crime, immigration, and guns. CNN/MSNBC bigrams have a more liberal flavor, including mentions of the gay community, women, and minorities. Interestingly, both sets of features include mentions to crime, immigration, and economy. This suggests that the differences in cable-news rhetoric are due in part to framing of topics, rather than just choice of topics.

3.3. Text similarity between newspapers and TV stations

Having validated that our model captures useful information, we can now take it to the newspaper snippets to score their relative similarity to each cable news network. Let N be the set of newspaper article snippets (indexed by n) and A_n^s the frequency of predictive bigram s in snippet n . A_n is the vector of frequencies (of length S) for article n . Our prediction of FNC , \widehat{FNC}_n , for snippet n is hence:

$$\widehat{FNC}_n = \Pr[FNC = 1|A_n] = \frac{1}{1 + \exp(-\hat{\psi}'_s A_n)}$$

which gives a predicted probability (between zero and one) for how likely each newspaper snippet was generated by Fox News.

Note that, as newspaper article snippets do not come with any label, we cannot evaluate accuracy in predicting newspaper article language. However, we provide some interpretive validation in Appendix C, where we list the news article snippets with the highest and lowest \widehat{FNC}_n . We find that the topical and rhetorical content of the article snippets intuitively reflect intuitions about the ideological commitments of the networks. In Appendix Table A4, we see that FNC-related articles include defenses of U.S. military involvement in Africa, crime, Bush’s opposition to troop withdrawals, and a Supreme Court case about the Second Amendment (right to bear arms). Articles that are closest to CNN/MSNBC (Appendix Table A5) are about campus groups supporting gay rights, the AIDS crisis in Africa, President Bush’s responsibility for the financial crisis, and HIV in the gay community.

We now have \widehat{FNC}_n as a similarity measure between TV channel language and the

⁷Appendix Table A1 includes a longer list, with 30 bigrams per class.

language in newspaper article n . To link the article-level data to the other datasets at the newspaper level j or the county level i , we aggregate by taking the mean values of the contained news articles. Hence, we define \widehat{FNC}_{jik} to be the probability (or share) of newspaper j 's snippets to be FNC-like, and likewise \widehat{FNC}_{ik} for articles in county i , where k indicates state.

3.4. Linked dataset

For the main analysis, we combine our text similarity measures with data on cable news viewership. In our main dataset, there are 305 newspapers circulating in 12.4 counties on average (the median is six counties), resulting in the aforementioned 3,781 observations.

4. Econometric framework

Our main hypothesis is that higher viewership of a cable channel in a county will cause the local newspapers to feature content similar to that channel's. This section outlines our method to test for this causal relationship.

4.1. Instrumental variables specification

Our main results are at the newspaper-county level (indexed by jik , with j for newspapers, i for county, and k for state). The problem can be formulated as a standard linear regression:

$$\widehat{FNC}_{jik} = \alpha_k + \theta \text{FNCViewership}_{ik} + K_{ik}\beta + L_{jik}\gamma + \epsilon_{jik} \quad (2)$$

where our dependent variable, \widehat{FNC}_{jik} , is the relative similarity to Fox News for newspaper j issued in county i in a state k . The treatment variable, $\text{FNCViewership}_{ik}$, is the county-level Fox News viewership (averaged over the period 2005 to 2008). In our main analyses, it is specified in relative terms, compared to averaged CNN and MSNBC viewership.⁸ We also look at absolute FNC viewership. The regression includes state fixed

⁸Formally, $\text{FNCViewership}_{ik}$ is given by $(\text{FNC Viewership} - 0.5 * (\text{MSNBC viewership} + \text{CNN viewership}))$.

effects (α_k), a vector of demographic and cable system controls (K_{ik}), and newspaper-county specific characteristics (L_{jik}). Finally, ϵ_{jik} is the error term. θ is the causal parameter of interest.

There are many political and economic factors that may correlate with both Fox News viewership and the use of Fox News language by local newspapers – in particular, any pre-existing ideological preferences of the county. The potential correlation between $\text{FNCViewership}_{ik}$ and ϵ_{jik} in (2) would be confounded by these endogenous factors. OLS estimation of this equation is therefore likely to provide biased estimates of θ .

To address this problem, we take an instrumental-variables approach. We require an instrument Z_{ik} that is correlated with $\text{FNCViewership}_{ik}$ but not with other factors affecting $\widehat{\text{FNC}}_{jik}$. Inspired by [Martin and Yurukoglu \(2017\)](#), we use cable network channel positioning to construct Z_{ik} . In the absolute-position specification, we instrument $\text{FNCViewership}_{ik}$ by the ordinal position of FNC in the channel system. In the relative-position specification, viewership is instrumented by the position of Fox News minus the average position of CNN and MSNBC.

We define the first stage regression:

$$\text{FNCViewership}_{ik} = \alpha_k + \delta Z_{ik} + K_{ik}\beta + L_{jik}\gamma + \eta_{jik} \quad (3)$$

which, combined with Equation (2), can be estimated with two-stage least squares (2SLS) to procure causal estimates for θ .

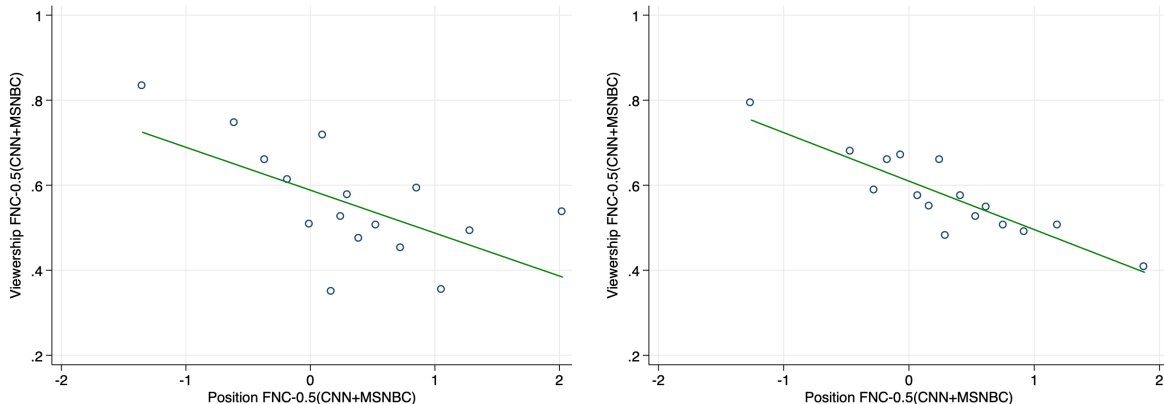
To facilitate the interpretation of the coefficients, we standardize the instrument, endogenous regressor, and outcome by dividing the original values by the standard deviations. All observations are weighted by their circulation. We multiway-cluster standard errors at the county and at the newspaper level.⁹

4.2. Instrument Validity

2SLS requires relevance in the first stage. [Figure 3](#) visualizes the first-stage relation between the FNC channel position (relative to the averaged position of CNN and MSNBC) and FNC viewership (also relative to CNN/MSNBC). The relationship is strongly negative and similar with (right panel) and without controls (left panel) In the tables below, we report Kleinbergen-Paap cluster-robust first-stage F-statistics.

⁹We obtain similar results when standard errors are clustered at the state level.

Figure 3: First Stage: Cable Channel Position and Cable News Viewership



Binned scatterplots (16 bins) of standardized viewership of FNC-0.5(CNN+MSNBC) against standardized position of FNC-0.5(CNN+MSNBC). Cross-section with newspaper-county-level observations weighted by newspaper circulation in each county. On the left, state fixed effects are included. On the right, state fixed effects, as well as demographic controls (see Appendix Table A2), channel controls (population share with access to each of the three TV channels), and generic newspaper language controls (vocabulary size, avg. word length, avg. sentence length, avg. article length) are included.

Beyond relevance, 2SLS imposes three requirements for consistent estimates. The first two, exclusion and monotonicity, are not problematic in our context. Exclusion requires that the channel position affects local news reporting only through its effect on cable news viewership. We believe this is a reasonable assumption in our context. The monotonicity assumption is that the channel position influences news viewership in the same direction for all counties. We believe that it is reasonable to assume that increasing the channel position would not systematically increase viewership.

The third assumption, exogeneity, is that Z_{ik} is uncorrelated with ϵ_{jik} . More concretely, we need that the channel position is not endogenously selected with county-specific preferences for conservative or liberal news reporting. The main identification problem is that channel positions could be allocated strategically in response to local factors correlated with conservative news messaging.

Martin and Yurukoglu (2017) provide a detailed discussion and a set of checks supporting the exogeneity assumption. Based on qualitative research, they highlight that channel positions have an important arbitrary, historical component, with significant inertia and path dependence. Quantitatively, the instrument is not correlated with Republican vote shares before the introduction of Fox News Channel. Galletta and Ash (2019) report a number of additional checks at the county level showing the instrument to be unrelated to demographic characteristics that predict policy preferences or news

Table 3: Identification checks: Instrument is Uncorrelated with Relevant Covariates

	Reduced form			
	<i>Viewership</i> FNC^*	\widehat{FNC}_{jik}^*	<i>Viewership</i> $(FNC - 0.5$ $(MSNBC+CNN))^*$	\widehat{FNC}_{jik}^*
	(1)	(2)	(3)	(4)
Channel position FNC	-0.022 (0.043)	0.017 (0.013)		
Channel position FNC - 0.5(MSNBC+CNN)			0.006 (0.025)	-0.006 (0.006)
N observations	3781	3781	3781	3781
State FE	X	X	X	X

Notes: Estimates are based on OLS with newspaper-county-level observations weighted by newspaper circulation in the respective county. The dependent variable is the predicted viewership of FNC in column (1), the predicted newspaper language similarity with FNC in columns (2) and (4), and the predicted viewership of FNC relative to averaged MSNBC and CNN viewership in (3). The predictions are derived from regressions that include the full set of demographic controls and state fixed effects. Standard errors are multiway-clustered at the county and at the newspaper level (in parenthesis): * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

channel viewership.

We apply the same identification checks to the counties in our sample (newspaper-county-level data). As in [Galletta and Ash \(2019\)](#), we use linear regressions with demographic characteristics and state fixed effects as covariates to predict viewership and newspaper content. First, we obtain predictions related to viewership: *Viewership* $(FNC - 0.5(MSNBC+CNN))^*$ (FNC viewership minus averaged MSNBC and CNN viewership) and *Viewership* FNC^* (Fox News viewership). Second, we form predictions for the probability of newspaper content to be Fox-like: \widehat{FNC}_{jik}^* . In both cases, the asterisk denotes that these variables are linear predictions based on demographics and state fixed effects only. These predictions summarize the variation in viewership and news content that is due to pre-existing cultural, economic, and political characteristics of these counties.

We then regress these predictions on different definitions of our instrument Z_{ik} . Table 3 shows the results of this identification check. Columns 1 and 2 document that there is no significant relationship between the absolute position of Fox News and the predicted values for viewership or newspaper. Columns 3 and 4 show that there is no significant relationship between relative FNC channel position and the predictions.

As an alternative check for the exogeneity assumption, we produce placebo estimates using newspaper content from before the introduction of Fox News Channel. Specifically,

we estimate 2SLS for our main specifications, but with a different outcome variable: the predicted text similarity to cable news using local newspaper articles from 1995 and 1996.¹⁰ The placebo check estimates, reported in Appendix Table A6, show no significant effects. Reassuringly, there was not a pre-existing Fox-like language dimension in locations that later had a lower Fox channel position.¹¹

Overall, these results support the view that channel positions are not selected or adapted to county characteristics that are otherwise important for our endogenous regressor or our outcome. The placebo nulls provide additional support for instrument validity. We therefore have support for the exogeneity assumption and for our 2SLS empirical strategy.

5. Results

This section reports the results. We begin with the main results, which include reduced form and two-stage-least-squares estimates.¹² Second, we report robustness checks.

5.1. Main Results

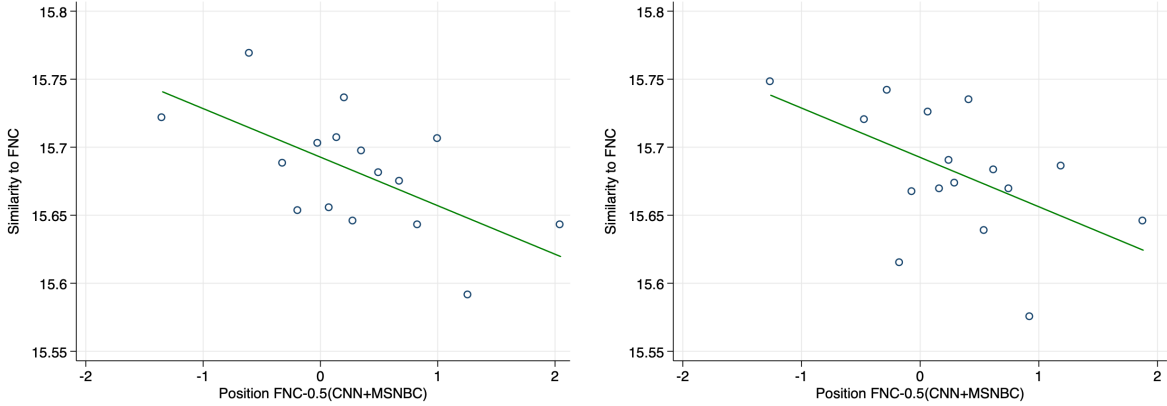
Figure 4 visualizes the reduced-form relationship between the FNC channel position (relative to the averaged MSNBC and CNN position) and local newspaper content similarity to FNC. In the left part of the Figure, the outcome and the instrument are residualized on state fixed effects. On the right-hand side, we additionally include demographic controls (as listed in Appendix Table A2), channel controls (share of households with access to each of the three channels), and generic newspaper language features (vocabulary size, average word length, average sentence length, and average article length). There is a clear downward relationship, suggesting that easier access to FNC is associated with more FNC-like language in the county’s local newspapers.

¹⁰The text similarity scores use the 2005-2008 transcripts (same as above) because FNC and MSNBC did not yet exist in 1995-1996.

¹¹The placebo regressions are based on fewer observations than the main results because some newspaper titles are not yet available in NewsLibrary in 1995 and 1996. Our main results remain qualitatively similar and are significant at the $p < 0.01$ level if we only use the observations entering the placebo regression.

¹²As discussed in Section 4, OLS estimates of the relationship between FNC viewership and the presence of FNC-related language in local newspapers do not have a causal interpretation. These are reported in Appendix F.

Figure 4: Reduced Form: Cable News Channel Position and Local Newspaper Content Similarity



Notes: Binned scatterplots (16 bins) of standardized textual similarity with Fox News against standardized position of $FNC-0.5(CNN+MSNBC)$. Cross-section with newspaper-county-level observations weighted by newspaper circulation in each county. On the left, state fixed effects are included. On the right, state fixed effects, as well as demographic controls (see Appendix Table A2), channel controls (population share with access to each of the three TV channels), and generic newspaper language controls (vocabulary size, avg. word length, avg. sentence length, avg. article length) are included.

Table 4 shows two-stage-least-squares estimates of the effect of higher FNC viewership on newspaper language similarity with FNC (the average probability that a snippet from a newspaper is predicted to be from FNC based on the bigrams it contains). The right-hand side variable of interest is instrumented FNC viewership relative to averaged CNN and MSNBC viewership, denoted as $Viewership(FNC - MSNBC - CNN)$.

All columns include state fixed effects and demographic controls. Column 2 also includes controls for the share of households with potential access to each of the three TV channels. Column 3 additionally controls for generic newspaper language features.

In all three columns, the estimated treatment effects are positive and statistically significant. The magnitudes across columns are highly similar, they range from 0.31 in columns 1 and 2 to 0.32 in column 3. This means that the channel and language controls do not change the estimates relative to the baseline in column 1 where only state fixed effects and demographic controls are included. All variables are standardized, so the interpretation is as follows: if Fox News viewership (relative to averaged CNN and MSNBC viewership) increases by one standard deviation in county ik where newspaper jik circulates, the similarity of its content with FNC, \widehat{FNC}_{jik} , increases by 0.31 (0.32) standard deviations.

These results support the hypothesis that increased exposure to Fox News increases the similarity of language in local newspapers to the language of Fox News. Next, we

Table 4: Cable News Effects on Newspaper Content (2SLS)

<i>Dep. variable: $\widehat{FNC}_{jik}=P(FNC Text_{jik})$</i>	(1)	(2)	(3)
<i>Viewership (FNC - MSNBC - CNN)</i>	0.314*** (0.114)	0.311*** (0.113)	0.318** (0.126)
F-test	36.553	36.298	34.147
N observations	3781	3781	3781
State FE	X	X	X
Demographic controls	X	X	X
Channel controls		X	X
Newspaper language controls			X

Notes: 2SLS estimates. Cross-section with newspaper-county-level observations weighted by newspaper circulation in each county. The dependent variable is newspaper language similarity with FNC (the average probability that a snippet from a newspaper is predicted to be from FNC): $\widehat{FNC}_{jik}=P(FNC|Text_{jik})$. The right-hand side variable of interest is instrumented FNC viewership relative to averaged CNN and MSNBC viewership: *Viewership (FNC - MSNBC - CNN)*. All columns include state fixed effects and demographic controls as listed in Appendix Table A2. Column 2 also includes channel controls (population shares with access to each of the three TV channels). Column 3 controls for generic newspaper language features (vocabulary size, avg. word length, avg. sentence length, avg. article length). Standard errors are multiway-clustered at the county and at the newspaper level (in parenthesis): * p < 0.1, ** p < 0.05, *** p < 0.01.

will run some specification checks and shed light on the mechanisms.

5.2. Specification and Robustness Checks

This section provides a number of specification and robustness checks for the main results on the effect of cable news viewership on newspaper content.

Absolute and Relative FNC Viewership. Appendix Tables A11 to A13 replicate the main results, but using different specifications for the instrument. Instead of FNC viewership relative to CNN and MSNBC viewership combined, we look at FNC viewership relative to CNN and MSNBC separately (Appendix Table A11 for CNN and Appendix Table A17 for MSNBC). Appendix Table A13 looks at absolute FNC viewership. Qualitatively, the results are consistent with our main results. All effect estimates enter with positive coefficients, but not all of them are precisely estimated. The instrumented relative viewership measures enter with larger and more precisely estimated coefficients than instrumented absolute viewership. In particular, our main results seem to be driven by higher exposure to FNC relative to MSNBC (more strongly so than relative to CNN).

Table 5: Cable News Effects on Newspaper Content (2SLS) – Alternative Matching Procedure

<i>Dep. variable: $\widehat{FNC}_{jik}=P(FNC Text_{jik})$</i>	(1)	(2)	(3)
<i>Viewership (FNC - MSNBC - CNN)</i>	0.859** (0.338)	1.051** (0.426)	1.164** (0.455)
F-test	19.704	14.828	13.869
N observations	682	682	682
State FE	X	X	X
Demographic controls	X	X	X
Channel controls		X	X
Newspaper language controls			X

Notes: 2SLS estimates. Cross-section with newspaper-level observations weighted their total circulation. The dependent variable is newspaper language similarity with FNC (the average probability that a snippet from a newspaper is predicted to be from FNC): $\widehat{FNC}_{jik}=P(FNC|Text_{jik})$. The right-hand side variable of interest is instrumented FNC viewership relative to averaged CNN and MSNBC viewership: *Viewership (FNC - MSNBC - CNN)*. All columns include state fixed effects and demographic controls as listed in Appendix Table A2. Column 2 also includes channel controls (population shares with access to each of the three TV channels). Column 3 controls for generic newspaper language features (vocabulary size, avg. word length, avg. sentence length, avg. article length). Standard errors are clustered at the state level (in parenthesis): * p < 0.1, ** p < 0.05, *** p < 0.01.

Alternative Matching Procedure. Table 5 shows our main specifications, but relying on a different matching of newspapers to counties. We assign each newspaper to a main county based on its name and other metadata. For the circulation, we use more broad-based, but less granular data (Appendix Section C.2). Hence, each newspaper is now only assigned to one county, where its total circulation is assumed to accrue. As Table 5 shows, with the alternative matching, the coefficients are again highly significant and around three to four times larger than in the main Table 4. Appendix Tables A16 to A18 confirm the previous finding that the effects are particularly strong for FNC viewership relative to MSNBC viewership (more so than FNC viewership relative to CNN or absolute FNC viewership).¹³

Alternative Weighting and Clustering. Appendix Table A14 reproduces Table 4, but with different circulation weights: each county is assigned its respective share of the total circulation of the newspaper across counties. Hence, each newspaper – ranging

¹³Appendix Tables A16 to A18 in Appendix J are the alternative matching procedure-based counterparts to Tables A11 to A13 in Appendix G.

from the “Butler Bulletin” with a total circulation of around 400 copies to the “Houston chronicle” with around 1.5 million copies – has now a total weight of one in the regression. With relative weights, the coefficients are still positive, but smaller by a factor of around 2.5 and not significant. We thus conclude that relative circulation weights support our main results in terms of coefficient sign. However, the main results seem to be driven – to a large extent – by giving more weight to more-read newspapers (i.e., adopting a strategy where individual readers are approximated).

Appendix Table A15 I shows our main results again, but with standard errors clustered at the state level (instead of multi-way clustered at the county and at the newspaper level). The state-level clustered errors are slightly larger, but the main effects are still significant at the level of $p < 0.05$.

Effects of FNC on Generic Language Features. The politicized nature of the bigrams predictive of FNC versus MSNBC/CNN (see Section 3.2) suggests that it is indeed media slant that spreads from FNC to local newspapers. Alternatively, it could be that more generic (and apolitical) features of the newspaper language change in response to higher FNC exposure. In Appendix Table A19, we do not uncover any statistically significant effect of higher exposure to FNC on generic language features (vocabulary size, average word length, average sentence length, and average article length).¹⁴

6. Mechanisms

6.1. Cable News Effects are Larger in Headquarter Counties

Both demand- or supply-driven mechanisms could explain the effects observed in Subsection 5. Newspaper readers, influenced by their cable news consumption, might demand more slanted news content. At the same time, editors or journalists exposed to certain channels may promote similarly slanted content. Without additional assumptions on how news are produced (e.g., where and how editorial decisions are taken or where the journalists are based), it is not possible to conclusively disentangle demand and supply effects. What we can show, however, is that the effects are driven, in particular,

¹⁴This null effect is why we include the generic text features as controls in some of our regression specifications. If they had been affected, they would be a “bad control” or “collider”.

Table 6: Cable News Effect on Local Newspapers is Concentrated in Newspaper Headquarter County

<i>Dep. variable: $\widehat{FNC}_{jik}=P(FNC Text_{jik})$</i>	(1)	(2)	(3)	(4)
<i>Viewership (FNC - MSNBC - CNN)</i>	0.136 (0.084)	0.112 (0.109)	0.367** (0.149)	0.353** (0.148)
F-test	47.698	46.439	11.630	11.982
N observations	3507	3507	263	263
Non-Headquarters	X	X		
Headquarters			X	X
State FE	X	X	X	X
Demographic controls	X	X	X	X
Channel controls		X		X
Newspaper language controls		X		X

Notes: 2SLS estimates. Cross-section with newspaper-county-level observations weighted by newspaper circulation in each county. The dependent variable is newspaper language similarity with FNC (the average probability that a snippet from a newspaper is predicted to be from FNC): $\widehat{FNC}_{jik}=P(FNC|Text_{jik})$. The right-hand side variable of interest is instrumented FNC viewership relative to averaged CNN and MSNBC viewership: *Viewership (FNC - MSNBC - CNN)*. Columns (1) and (3) include state fixed effects and demographic controls as listed in Appendix Table A2. Columns (2) and (4) additionally include channel controls (population shares with access to each of the three TV channels), and controls for generic newspaper language features (vocabulary size, avg. word length, avg. sentence length, avg. article length). Standard errors are multiway-clustered at the county and at the newspaper level (in parenthesis): * p < 0.1, ** p < 0.05, *** p < 0.01.

by the newspaper-county observations where the county coincides with the newspaper headquarters.

Table 6 reports the 2SLS coefficients based on reduced samples where we either only consider newspaper-county observations where the county does *not* coincide with where the immediate owner of the newspaper is based (columns 1 and 2) or only headquarter newspaper-county observations (columns 3 and 4). For example, consider “The York Dispatch”, which serves the counties York, Lancaster, Dauphin, and Adams. It is based in York County.¹⁵ In columns 1 and 2 of Table 6, we exclude headquarter counties such as York County in this example (but keep Lancaster, Dauphin, and Adams). In columns 1 and 3, we only include demographic controls, while we also add channel and language controls in columns 2 and 4.

The results suggest that our observed effects are concentrated in headquarter counties. The coefficient signs for non-headquarter counties (columns 1 and 2) are positive

¹⁵We assign the city where the *immediate* owner of the local newspaper is based – according to the Alliance for Audited Media – to a U.S. county. That is, we do not consider the location of the parent company for newspapers that are owned by a conglomerate.

but smaller (by a factor of two to three) and not statistically significant. In columns 3 and 4, where we look at headquarter counties, the coefficients are larger than the main results (around 11% larger) and statistically significant with $p < 0.05$.

6.2. Cable News Media Slant Polarizes Local Newspapers

Next, we investigate effect heterogeneity with respect to Presidential endorsements of newspapers in the pre-FNC era. To that end, we distinguish three groups of newspapers based on the 1996 U.S. presidential election endorsements: (1) those that endorsed the Republican candidate Bob Dole, (2) those that endorsed the Democrat candidate Bill Clinton, and (3) those that did not endorse either candidate.¹⁶ We think of endorsements as a signal for whether the pre-FNC political leaning of a newspaper was relatively conservative or liberal. The non-endorsers can be seen as politically neutral.

Table 7 shows the heterogeneity analysis. As with the main results, we focus on instrumented FNC viewership relative to averaged CNN and MSNBC viewership: *Viewership (FNC - MSNBC - CNN)*. All columns include our demographic, channel, and language controls. The difference is subsetting by groups of newspapers according to pre-FNC endorsements.

In column 1, we subset on newspapers that endorsed the Democrat candidate. The negative coefficient suggests that newspapers with a more liberal leaning (according to our measure) become *less* FNC-like in their reporting, the more exposed they are to FNC (relative to MSNBC and CNN). Conversely, in column 4, for newspapers with a conservative leaning, we see a statistically significant positive effect. Meanwhile, when looking at newspapers with no endorsements (column 3) or combining the Democrat- and null-endorsers (column 2), there is no effect of cable media slant on newspaper content.

Taken together, these estimates provide some evidence that higher exposure to FNC relative to CNN/MSNBC (or vice versa) leads to more polarized local newspaper language. Newspapers with a right-wing partisan leaning in the pre-FNC era become more right-wing in response to FNC viewership increases. Instead, left-wing papers become more left-wing. This could be due to a market positioning effect, where the conservative papers situate themselves to accommodate FNC-viewer news preferences. In turn, liberal papers respond and situate to accommodate non-FNC-viewer news preferences.

¹⁶In this latter group we include newspapers where we could not find an explicit endorsement.

Table 7: Cable News Exposure Tends to Polarize Local Newspaper Content

<i>Dep. variable: $\widehat{FNC}_{jik}=P(FNC Text_{jik})$</i>	(1)	(2)	(3)	(4)
<i>Viewership (FNC - MSNBC - CNN)</i>	-0.463*	0.060	0.062	0.259**
	(0.246)	(0.125)	(0.160)	(0.116)
F-test	10.586	30.195	19.708	16.070
N observations	872	2735	1858	1040
Endorsed Democrat	X	X		
No Endorsement		X	X	
Endorsed Republican				X
State FE	X	X	X	X
Demographic controls	X	X	X	X
Channel controls	X	X	X	X
Newspaper language controls	X	X	X	X

Notes: 2SLS estimates. Cross-section with newspaper-county-level observations weighted by newspaper circulation in each county. The dependent variable is newspaper language similarity with FNC (the average probability that a snippet from a newspaper is predicted to be from FNC): $\widehat{FNC}_{jik}=P(FNC|Text_{jik})$. The right-hand side variable of interest is instrumented FNC viewership relative to averaged CNN and MSNBC viewership: *Viewership (FNC - MSNBC - CNN)*. In column 1 we only consider newspapers that endorsed the Democratic Presidential candidate in 1996 (the pre-FNC era). Column 2 includes newspapers that endorsed the Democratic candidate or did not endorse either candidate (or for which endorsement data is not available). In column 3, we focus on newspapers that did not endorse either candidate (or for which endorsement data is not available). Column 4 considers only newspapers that endorsed the Republican candidate. All columns include state fixed effects, demographic controls as listed in Appendix Table A2, channel controls (population shares with access to each of the three TV channels), and generic newspaper language features (vocabulary size, avg. word length, avg. sentence length, avg. article length). Standard errors are multiway-clustered at the county and at the newspaper level (in parenthesis): * p < 0.1, ** p < 0.05, *** p < 0.01.

7. Conclusion

This research analyzes the causal effect of partisan cable television news content on local printed newspaper content. Using a regularized prediction of whether text comes from Fox News or from competing networks (CNN/MSNBC), we show that exposure to Fox News Channel in the market of a local newspaper increases the linguistic similarity of the newspaper content to Fox News shows, relative to shows on other news networks. Cable news content influences local news content. Our results add to the literature on the political effects of biased news. We provide new evidence on how partisan media influence not just voting and policy, but also discourse. This means that previous estimates of effects of media could be measuring not just direct effects on primary media consumers, but also indirect effects through diffusion into secondary media sources.

References

- Ash, E., Morelli, M., and Van Weelden, R. (2017). Elections and divisiveness: Theory and evidence. *Journal of Politics*.
- Ashworth, S. and Shotts, K. W. (2010). Does informative media commentary reduce politicians' incentives to pander? *Journal of Public Economics*, 94(11-12):838–847.
- Chiang, C.-F. and Knight, B. G. (2011). Media bias and influence: Evidence from newspaper endorsements. *Review of Economic Studies*, 78(3):795–820.
- DellaVigna, S. and Kaplan, E. (2007). The fox news effect: Media bias and voting. *The Quarterly Journal of Economics*, 122(3):1187–1234.
- Enikolopov, R., Petrova, M., and Zhuravskaya, E. (2011). Media and political persuasion: Evidence from Russia. *American Economic Review*, 101(7):3253–3285.
- Galletta, S. and Ash, E. (2019). How cable news reshaped local government. Technical report, SSRN.
- Gentzkow, M. and Shapiro, J. M. (2010). What drives media slant? evidence from us daily newspapers. *Econometrica*, 78(1):35–71.
- Gentzkow, M., Shapiro, J. M., and Sinkinson, M. (2011). The effect of newspaper entry and exit on electoral politics. *American Economic Review*, 101(7):2980–3018.
- Gentzkow, M., Shapiro, J. M., and Stone, D. F. (2015). Media bias in the marketplace: Theory. In Anderson, S. P., Waldfogel, J., and Stroemberg, D., editors, *Handbook of Media Economics*, volume 1, pages 623–645. North-Holland.
- Gentzkow, M., Shapiro, J. M., and Taddy, M. (2019). Measuring group differences in high-dimensional choices: Method and application to congressional speech. *Econometrica*, 87(4):1307–1340.
- George, L. M. and Waldfogel, J. (2006). The new york times and the market for local newspapers. *American Economic Review*, 96(1):435–447.
- Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning*. Springer, New York, NY.
- Hopkins, D. J. and Ladd, J. M. (2014). The consequences of broader media choice: Evidence from the expansion of Fox News. *Quarterly Journal of Political Science*, 9(1):115–135.
- Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J., and Mullainathan, S. (2017). Human decisions and machine predictions. *The Quarterly Journal of Economics*, 133(1):237–293.

- Martin, G. and Yurukoglu, A. (2017). Bias in cable news: Persuasion and polarization. *American Economic Review*.
- Martin, G. J. and McCrain, J. (2019). Local news and national politics. *American Political Science Review*, 113(2):372–384.
- Mastrorocco, N. and Ornaghi, A. (2020). Who Watches the Watchmen? Local News and Police Behaviour in the United States. Technical report, Trinity Economics Papers.
- Peterson, A. and Spirling, A. (2018). Classification Accuracy as a Substantive Quantity of Interest: Measuring Polarization in Westminster Systems. *Political Analysis*, forthcoming.
- Prat, A. (2018). Media power. *Journal of Political Economy*, 126(4):1747 – 1783.
- Prat, A. and Stroemberg, D. (2013). *The Political Economy of Mass Media*, volume 2 of *Econometric Society Monographs*, page 135–187. Cambridge University Press, Cambridge, UK.
- Puglisi, R. and Snyder, J. M. (2015). Empirical studies of media bias. In Anderson, S. P., Waldfogel, J., and Stroemberg, D., editors, *Handbook of Media Economics*, volume 1, pages 647–667. North-Holland.
- Schroeder, E. and Stone, D. F. (2015). Fox news and political knowledge. *Journal of Public Economics*, 126:52–63.
- Snyder, J. M. and Strömberg, D. (2010). Press coverage and political accountability. *Journal of Political Economy*, 118(2):355–408.
- Stroemberg, D. (2015). Media and politics. *Annual Review of Economics*, 7:173–205.

A. Bigrams most predictive for FNC or CNN/MSNBC

Table A1 lists the 30 bigrams which are most predictive for a transcript being from FNC or CNN/MSNBC, respectively. Specifically, we list the bigrams with the largest absolute coefficients in the logistic regression described in Section 3. Coefficients are positive for FNC and negative for CNN/MSNBC.

Table A1: Top 30 of bigrams predictive for FNC or CNN/MSNBC transcripts

Fox	CNN/MSNBC
unusu case	way pass
open mouth	iraq immedi
american troop	reason fact
vp choic	someth mani
read thing	place democrat
home tri	conspiraci theorist
sever offici	peopl quit
sort right	campaign ran
allow govern	work went
allow right	harri counti
import life	car start
answer phone	break thing
crimin convict	north vietnames
obvious hes	discuss today
arab alli	bailout money
truth iraq	view peopl
work campaign	think robert
peopl sinc	kidnap case
came help	mexican flag
read lot	dont involv
new hous	vote tax
talk al	posit power
time read	know fine
want extend	like set
say night	requir troop
work direct	presidenti nomine
dont start	parti congress
immigr author	come kid
door shut	launch anoth
term hes	today high

B. Summary Statistics

Table A2: Summary Statistics

Variable	Mean	Std. Dev	Min	Max	N
<i>Newspapers (and newspaper language)</i>					
Probability FNC	0.434	0.029	0.38	0.655	3781
Circulation	5067.311	18487.863	1	390687.8	3781
Vocabulary size	0.026	0.023	0.007	0.146	3781
Word length	7.221	0.288	6.292	8.321	3781
Sentence length	35.846	6.826	18.858	69.036	3781
Article length	447.782	112.426	193.957	918.617	3781
# collected articles	83289.858	57642.203	1067	286027	3781
<i>News channels</i>					
FNC channel position 2005-2008	42.647	11.502	5	74.625	3781
CNN channel position 2005-2008	30.299	10.408	1	66.306	3781
MSNBC channel position 2005-2008	45.138	13.006	4	128.5	3781
Position (Fox News-0.5(MSNBC+CNN))	4.928	11.219	-48.042	54.75	3781
Position (Fox News - MSNBC)	-2.491	13.641	-81.417	55	3781
Position (Fox News - CNN)	12.347	12.769	-39	59.618	3781
Ratings % Fox News 2005-2008	0.539	0.354	0	5.475	3781
Ratings % MSNBC 2005-2008	0.14	0.384	0	13	3781
Ratings % CNN 2005-2008	0.303	0.229	0	3.7	3781
Ratings (%Fox News - 0.5(%MSNBC - %CNN))	0.318	0.411	-7.600	5.412	3781
Ratings (%Fox News - %MSNBC)	0.399	0.512	-12.55	5.45	3781
Ratings (%Fox News - %CNN)	0.236	0.377	-2.65	5.375	3781
Share pop. access to FNC	0.934	0.106	0.039	1	3781
Share pop. access to MSNBC	0.891	0.174	0.004	1	3781
Share pop. access to CNN	0.962	0.067	0.055	1	3781
<i>Demographic</i>					
Population 2000	210572.061	509458.494	712	9818535	3781
Republican vote share 1996	0.426	0.098	0.093	0.79	3781
White	0.858	0.139	0.057	0.993	3781
Black	0.077	0.116	0	0.795	3781
Asian	0.014	0.028	0	0.359	3781
Hispanic	0.064	0.114	0.002	0.973	3781
Male	0.493	0.014	0.433	0.627	3781
Age 10-19	0.16	0.016	0.088	0.277	3781
Age 20-29	0.12	0.03	0.047	0.327	3781
Age 30-39	0.144	0.017	0.082	0.23	3781
Age 40-49	0.153	0.013	0.096	0.213	3781
Age 50-59	0.116	0.012	0.063	0.177	3781
Age 60-69	0.082	0.017	0.035	0.173	3781
Age 70-79	0.066	0.018	0.013	0.172	3781
Age 80-89	0.039	0.013	0.003	0.121	3781
Urban	0.538	0.284	0	1	3781
High school	0.342	0.072	0.111	0.527	3781
Some college	0.263	0.049	0.108	0.424	3781
Bachelor	0.125	0.053	0.028	0.397	3781
Postgraduate	0.068	0.04	0.013	0.31	3781
Land area	309.047	405.331	1.768	6812.404	3781
Population density	4.454	1.706	-1.08	10.307	3781
Mean log. income	10.794	0.231	10.105	11.597	3781
Gini index	0.429	0.036	0.335	0.604	3781
Occ. management and professional	29.63	6.791	16.6	61.3	3781
Occ. service	15.466	2.809	8.1	31.9	3781
Occ. sales and office	24.449	2.978	13.6	32.6	3781
Occ. farming, fishing, and forestry	1.43	1.722	0	24.9	3781
Occ. construction, extraction, and maintenance	10.818	2.644	2.3	24.5	3781

C. Data Appendix

This section provides some more detail on the data.

C.1. Newspaper Articles

First, some more info on NewsLibrary. Figure [_](#) shows a screenshot of the site. These are the search results, which for each article contain the newspaper, headline, date, byline (if any), and first 80 words (approximately) of the body text.

In principle, NewsLibrary encompasses around 4,000 unique titles for 2005-2008. However, for many titles, there are only a handful of articles available: around 1,500 titles contain less than 1,000 snippets (for all four years combined). In all our analyses, we only consider titles with more than 1000 articles. Also, many titles are not local newspapers in the sense that they cannot be assigned to a county (e.g., the “Army Communicator” or the “Air & Space” magazine). Furthermore, NewsLibrary often lists different editions of the same title separately. For instance, “Augusta Chronicle, the (GA)”, “Augusta Chronicle, the: Web Edition Articles (GA)”, and “Augusta Chronicle, the: Blogs (GA)” are listed separately. While our initial corpus covers all 2,618 titles with $>1,000$ articles (amounting to almost 50 million article snippets, see Section [C.3](#)), our main analyses focus on 305 titles for which county-level circulation data is available (see Section [2.3](#)). We also collapse different editions of the same outlet (as in the Augusta Chronicle example) to one observation because the Alliance for Audited Media circulation data is typically not available separately for different editions of the same title. The 16 million articles mentioned in the main text above refer to the outlets used in our main regression analyses.

C.2. Alternative County Matching of Newspapers

For robustness, we also apply a matching procedure which covers more newspaper titles, but which only provides total circulation instead of county-specific circulation. First, we obtain the main county for each newspaper title based on the newspaper name and geographical information provided by NewsLibrary (e.g., *The Call (Woonsocket, RI)* or the *Albany Democrat-Herald (OR)*), the U.S. Newspaper Directory, or a manual web search. We then assign total circulation (as of 2004) according to the Inter-University Consortium for Political and Social Research (IPCSR) to this main county. With this

matching approach, we have a dataset of 682 unique newspaper titles and 24 million article snippets.

C.3. Filtering and Number of Article Snippets

Table A3 gives an overview of the number of articles collected and how we obtain the number of articles used in our main analyses and robustness checks.

Table A3: Number of Articles Collected and Filtering

Filtering	# Articles	Results
No filtering: raw scrapes	49,891,120	None (not possible: no county assignment)
County assignment as in C.2 and total circulation data available (ICPSR)	23,979,516	Tables 5 and A16 to A18
County assignment as in 2.3 and county-level circulation data available (AAM)	16,098,537	Main results: Tables 4 and A11 to A13

D. Example Articles by Fox News Similarity

Table A4: Newspaper Articles that are most similar to Fox News Channel Shows

The Free Lance-Star (Fredericksburg, VA), 2 January 2008

98% similarity to FNC

Regarding their recent op-ed ["The Pentagon should stay out of Africa," Dec. 14], I am afraid Danny Glover and Nicole Lee are victims of misinformation about U.S. Africa Command. AFRICOM is not part of a "U.S. military expansion," nor will it involve placing many "American troops on foreign soil." Rather, AFRICOM marks recognition of the growing importance of Africa and reallocates responsibility for U.S. security interests accordingly. The U.S. Department of Defense assigns [...]

The Sacramento Bee (CA), 19 May 2007

97% similarity to FNC

Don Kercell thinks he's earned a second chance. The Contractors State License Board does not agree. And therein lies a tale of choices and consequences; crime and punishment; addiction and rehabilitation; public protection and personal redemption – and second chances. Kercell is a 48-year-old resident of Rio Linda. In his youth, he discovered two things. One was that he had a talent for working with concrete. The other was methamphetamine. [...]

Joplin Globe, the (MO), 28 April 2007

95% similarity to FNC

Bush vows to veto any attempts by Dems to force troop pullout CAMP DAVID, Md. - President Bush warned Congress Friday that he will continue vetoing war spending bills as long as they contain a timetable for the withdrawal of American troops from Iraq. Speaking a day after the Democratic-controlled Congress approved legislation that requires that a troop drawdown begin by Oct. 1, Bush said - as he has before - he will veto it because of that demand. He [...]

The Commercial Appeal (Memphis, TN), 16 June 2008

83% similarity to FNC

WASHINGTON - One momentous case down, another equally historic decision to go. The Supreme Court returns to the bench Monday with 17 cases still unresolved, including its first-ever comprehensive look at the Second Amendments right to bear arms. The guns case - including Washington, D.C.'s ban on handguns - is widely expected to be a victory for supporters of gun rights. Top officials of a national gun control organization said this week that they expect the handgun ban to be [...]

Table A5: Newspaper Articles that are most similar to CNN and MSNBC Shows

The Sun (San Bernardino, CA), 21 March 2005

3% similarity to FNC

REDLANDS - A week after a state judge ruled that banning gay marriage is unconstitutional, students at University of Redlands will celebrate the milestone along with continued efforts to raise awareness of the gay community. The PRIDE Alliance, a campus group devoted to promoting tolerance on campus for gay, lesbian, bisexual and transgender students, will celebrate PRIDE Week at the university through Friday. A series of events is scheduled to raise awareness on campus and in the [...]

Robesonian, the (Lumberton, NC), 12 October 2007

4% similarity to FNC

About \$18 billion a year has been drained from Africa by nearly two dozen wars in recent decades, a new report states, a price some officials say could've helped solve the AIDS crisis and created stronger economies in the world's poorest region. "This is money Africa can ill afford to lose," Liberian President Ellen Johnson Sirleaf wrote in an introduction to the report by the British charity Oxfam and two groups that seek tougher controls on small arms, Saferworld [...]

Denver Examiner (CO), 26 September 2008

5% similarity to FNC

John McCain and Barack Obama will indeed debate tonight at 7 p.m at the University of Mississippi, moderated by Jim Lehrer. The debate is scheduled to focus on issues of foreign policy, but given the economic meltdown of the last two weeks, and the Bush administration's proposed \$700 billion bailout plan, Politico is reporting that Lehrer might add in some questions on the economy. Also, Rich Lowry from National Review is reporting that everyone at Ole Miss "hates" McCain for [...]

Long Beach Press-Telegram (CA), 22 June 2006

6% similarity to FNC

There is finally some good news about the most sinister drug on the black market: crystal methamphetamine. Nationwide demand and production is down, according to federal drug cops. Meth, which has been linked to the spread of HIV in Long Beach's gay community, is still out there, but law enforcement officials say plenty of busts are reducing supplies. We hope that treatment is part of the equation nationwide as it is California, where voters agreed to put more users in treatment than in [...]

E. Placebo: Similarity with FNC in 1995/96

Table A6: Placebo Cable News Effects on Newspaper Content (2SLS)

<i>Dep. variable: $\widehat{FNC}_{jik}=P(FNC Text_{jik})$</i>	(1)	(2)	(3)
<i>Viewership (FNC - MSNBC - CNN)</i>	0.232 (0.301)	0.216 (0.290)	0.193 (0.321)
F-test	34.462	35.399	27.281
N observations	1892	1892	1892
State FE	X	X	X
Demographic controls	X	X	X
Channel controls		X	X
Newspaper language controls			X

Notes: 2SLS estimates. Cross-section with newspaper-county-level observations weighted by newspaper circulation in each county. The dependent variable is newspaper language similarity with FNC (the average probability that a snippet from a newspaper is predicted to be from FNC) **in 1995/1996** (pre-FNC era): $\widehat{FNC}_{jik}=P(FNC|Text_{jik})$. The right-hand side variable of interest is instrumented FNC viewership relative to averaged CNN and MSNBC viewership: *Viewership (FNC - MSNBC - CNN)*. All columns include state fixed effects and demographic controls as listed in Appendix Table A2. Column 2 also includes channel controls (population shares with access to each of the three TV channels). Column 3 controls for generic newspaper language features (vocabulary size, avg. word length, avg. sentence length, avg. article length). Standard errors are multiway-clustered at the county and at the newspaper level (in parenthesis): * p < 0.1, ** p < 0.05, *** p < 0.01.

F. OLS Results

Tables A7 to A10 show OLS results for regressions of predicted Fox News similarity for newspaper jik , \widehat{FNC}_{jik} , on TV channel viewership. Table A7 looks at FNC viewership relative to averaged MSNBC and CNN viewership. It hence shows the OLS estimates that mirror the 2SLS results in the main Table 4. Then, we look at FNC viewership relative to CNN and MSNBC separately (Table A8 for CNN and Table A9 for MSNBC). Table A10 looks at absolute FNC viewership. In each table, all columns include state fixed effects and demographic controls as listed in Appendix Table A2. Column 2 also includes channel controls (population shares with access to each of the three TV channels). Column 3 further controls for generic newspaper language features (vocabulary size, avg. word length, avg. sentence length, avg. article length). All four instruments enter with smaller (compared to the 2SLS estimates) coefficients that are significant only for the absolute FNC viewership in Table A10.

Table A7: Cable News Effects on Newspaper Content (OLS)

<i>Dep. variable: $\widehat{FNC}_{jik}=P(FNC Text_{jik})$</i>	(1)	(2)	(3)
<i>Viewership (FNC - MSNBC - CNN)</i>	0.016 (0.011)	0.017 (0.011)	0.015 (0.011)
N observations	3781	3781	3781
State FE	X	X	X
Demographic controls	X	X	X
Channel controls		X	X
Newspaper language controls			X

Notes: OLS estimates. Cross-section with newspaper-county-level observations weighted by newspaper circulation in each county. The dependent variable is newspaper language similarity with FNC (the average probability that a snippet from a newspaper is predicted to be from FNC): $\widehat{FNC}_{jik}=P(FNC|Text_{jik})$. The right-hand side variable of interest is FNC viewership relative to averaged CNN and MSNBC viewership: *Viewership (FNC - MSNBC - CNN)*. All columns include state fixed effects and demographic controls as listed in Appendix Table A2. Column 2 also includes channel controls (population shares with access to each of the three TV channels). Column 3 controls for generic newspaper language features (vocabulary size, avg. word length, avg. sentence length, avg. article length). Standard errors are multiway-clustered at the county and at the newspaper level (in parenthesis): * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A8: Cable News Effects on Newspaper Content (OLS)

<i>Dep. variable: $\widehat{FNC}_{jik}=P(FNC Text_{jik})$</i>	(1)	(2)	(3)
<i>Viewership (FNC - CNN)</i>	0.022 (0.013)	0.022 (0.013)	0.019 (0.014)
N observations	3781	3781	3781
State FE	X	X	X
Demographic controls	X	X	X
Channel controls		X	X
Newspaper language controls			X

Notes: OLS estimates. Cross-section with newspaper-county-level observations weighted by newspaper circulation in each county. The dependent variable is newspaper language similarity with FNC (the average probability that a snippet from a newspaper is predicted to be from FNC): $\widehat{FNC}_{jik}=P(FNC|Text_{jik})$. The right-hand side variable of interest is FNC viewership relative to CNN viewership: *Viewership (FNC - CNN)*. All columns include state fixed effects and demographic controls as listed in Appendix Table A2. Column 2 also includes channel controls (population shares with access to each of the three TV channels). Column 3 controls for generic newspaper language features (vocabulary size, avg. word length, avg. sentence length, avg. article length). Standard errors are multiway-clustered at the county and at the newspaper level (in parenthesis): * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A9: Cable News Effects on Newspaper Content (OLS)

<i>Dep. variable: $\widehat{FNC}_{jik}=P(FNC Text_{jik})$</i>	(1)	(2)	(3)
<i>Viewership (FNC - MSNBC)</i>	0.012 (0.010)	0.013 (0.010)	0.012 (0.010)
N observations	3781	3781	3781
State FE	X	X	X
Demographic controls	X	X	X
Channel controls		X	X
Newspaper language controls			X

Notes: OLS estimates. Cross-section with newspaper-county-level observations weighted by newspaper circulation in each county. The dependent variable is newspaper language similarity with FNC (the average probability that a snippet from a newspaper is predicted to be from FNC): $\widehat{FNC}_{jik}=P(FNC|Text_{jik})$. The right-hand side variable of interest is FNC viewership relative to CNN viewership: *Viewership (FNC - MSNBC)*. All columns include state fixed effects and demographic controls as listed in Appendix Table A2. Column 2 also includes channel controls (population shares with access to each of the three TV channels). Column 3 controls for generic newspaper language features (vocabulary size, avg. word length, avg. sentence length, avg. article length). Standard errors are multiway-clustered at the county and at the newspaper level (in parenthesis): * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A10: Cable News Effects on Newspaper Content (OLS)

<i>Dep. variable: $\widehat{FNC}_{jik}=P(FNC Text_{jik})$</i>	(1)	(2)	(3)
<i>Viewership FNC</i>	0.028** (0.013)	0.029** (0.013)	0.028** (0.013)
N observations	3781	3781	3781
State FE	X	X	X
Demographic controls	X	X	X
Channel controls		X	X
Newspaper language controls			X

Notes: OLS estimates. Cross-section with newspaper-county-level observations weighted by newspaper circulation in each county. The dependent variable is newspaper language similarity with FNC (the average probability that a snippet from a newspaper is predicted to be from FNC): $\widehat{FNC}_{jik}=P(FNC|Text_{jik})$. The right-hand side variable of interest is FNC viewership: *Viewership FNC*. All columns include state fixed effects and demographic controls as listed in Appendix Table A2. Column 2 also includes channel controls (population shares with access to each of the three TV channels). Column 3 controls for generic newspaper language features (vocabulary size, avg. word length, avg. sentence length, avg. article length). Standard errors are multiway-clustered at the county and at the newspaper level (in parenthesis): * p < 0.1, ** p < 0.05, *** p < 0.01.

G. Robustness: Absolute and Relative FNC Viewership

Table A11: Cable News Effects on Newspaper Content (2SLS)

<i>Dep. variable: $\widehat{FNC}_{jik}=P(FNC Text_{jik})$</i>	(1)	(2)	(3)
<i>Viewership (FNC - CNN)</i>	0.186 (0.116)	0.182 (0.117)	0.157 (0.143)
F-test	42.208	40.254	39.173
N observations	3781	3781	3781
State FE	X	X	X
Demographic controls	X	X	X
Channel controls		X	X
Newspaper language controls			X

Notes: 2SLS estimates. Cross-section with newspaper-county-level observations weighted by newspaper circulation in each county. The dependent variable is newspaper language similarity with FNC (the average probability that a snippet from a newspaper is predicted to be from FNC): $\widehat{FNC}_{jik}=P(FNC|Text_{jik})$. The right-hand side variable of interest is instrumented FNC viewership relative to CNN viewership: *Viewership (FNC - CNN)*. All columns include state fixed effects and demographic controls as listed in Appendix Table A2. Column 2 also includes channel controls (population shares with access to each of the three TV channels). Column 3 controls for generic newspaper language features (vocabulary size, avg. word length, avg. sentence length, avg. article length). Standard errors are multiway-clustered at the county and at the newspaper level (in parenthesis): * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A12: Cable News Effects on Newspaper Content (2SLS)

<i>Dep. variable: $\widehat{FNC}_{jik}=P(FNC Text_{jik})$</i>	(1)	(2)	(3)
<i>Viewership (FNC - MSNBC)</i>	0.370*** (0.111)	0.373*** (0.112)	0.409*** (0.126)
F-test	34.396	33.815	32.120
N observations	3781	3781	3781
State FE	X	X	X
Demographic controls	X	X	X
Channel controls		X	X
Newspaper language controls			X

Notes: 2SLS estimates. Cross-section with newspaper-county-level observations weighted by newspaper circulation in each county. The dependent variable is newspaper language similarity with FNC (the average probability that a snippet from a newspaper is predicted to be from FNC): $\widehat{FNC}_{jik}=P(FNC|Text_{jik})$. The right-hand side variable of interest is instrumented FNC viewership relative to CNN viewership: *Viewership (FNC - MSNBC)*. All columns include state fixed effects and demographic controls as listed in Appendix Table A2. Column 2 also includes channel controls (population shares with access to each of the three TV channels). Column 3 controls for generic newspaper language features (vocabulary size, avg. word length, avg. sentence length, avg. article length). Standard errors are multiway-clustered at the county and at the newspaper level (in parenthesis): * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A13: Cable News Effects on Newspaper Content (2SLS)

<i>Dep. variable: $\widehat{FNC}_{jik}=P(FNC Text_{jik})$</i>	(1)	(2)	(3)
<i>Viewership FNC</i>	0.139 (0.184)	0.136 (0.178)	0.075 (0.197)
F-test	12.405	13.039	15.011
N observations	3781	3781	3781
State FE	X	X	X
Demographic controls	X	X	X
Channel controls		X	X
Newspaper language controls			X

Notes: 2SLS estimates. Cross-section with newspaper-county-level observations weighted by newspaper circulation in each county. The dependent variable is newspaper language similarity with FNC (the average probability that a snippet from a newspaper is predicted to be from FNC): $\widehat{FNC}_{jik}=P(FNC|Text_{jik})$. The right-hand side variable of interest is instrumented FNC viewership: *Viewership FNC*. All columns include state fixed effects and demographic controls as listed in Appendix Table A2. Column 2 also includes channel controls (population shares with access to each of the three TV channels). Column 3 controls for generic newspaper language features (vocabulary size, avg. word length, avg. sentence length, avg. article length). Standard errors are multiway-clustered at the county and at the newspaper level (in parenthesis): * p < 0.1, ** p < 0.05, *** p < 0.01.

H. Robustness: Relative Circulation Weights

Table A14: Cable News Effects on Newspaper Content (2SLS)

<i>Dep. variable: $\widehat{FNC}_{jik}=P(FNC Text_{jik})$</i>	(1)	(2)	(3)
<i>Viewership (FNC - MSNBC - CNN)</i>	0.121 (0.137)	0.116 (0.140)	0.145 (0.133)
F-test	36.553	36.298	34.147
N observations	3781	3781	3781
State FE	X	X	X
Demographic controls	X	X	X
Channel controls		X	X
Newspaper language controls			X

Notes: 2SLS estimates. Cross-section with newspaper-county-level observations weighted by their circulation *share* in each county. The dependent variable is newspaper language similarity with FNC (the average probability that a snippet from a newspaper is predicted to be from FNC): $\widehat{FNC}_{jik}=P(FNC|Text_{jik})$. The right-hand side variable of interest is instrumented FNC viewership relative to averaged CNN and MSNBC viewership: *Viewership (FNC - MSNBC - CNN)*. All columns include state fixed effects and demographic controls as listed in Appendix Table A2. Column 2 also includes channel controls (population shares with access to each of the three TV channels). Column 3 controls for generic newspaper language features (vocabulary size, avg. word length, avg. sentence length, avg. article length). Standard errors are multiway-clustered at the county and at the newspaper level (in parenthesis): * p < 0.1, ** p < 0.05, *** p < 0.01.

I. Robustness: State-Level Clustered Errors

Table A15: Cable News Effects on Newspaper Content (2SLS)

<i>Dep. variable: $\widehat{FNC}_{jik}=P(FNC Text_{jik})$</i>	(1)	(2)	(3)
<i>Viewership (FNC - MSNBC - CNN)</i>	0.314** (0.120)	0.311** (0.117)	0.318** (0.135)
F-test	35.682	36.496	37.544
N observations	3781	3781	3781
State FE	X	X	X
Demographic controls	X	X	X
Channel controls		X	X
Newspaper language controls			X

Notes: 2SLS estimates. Cross-section with newspaper-county-level observations weighted by newspaper circulation in each county. The dependent variable is newspaper language similarity with FNC (the average probability that a snippet from a newspaper is predicted to be from FNC): $\widehat{FNC}_{jik}=P(FNC|Text_{jik})$. The right-hand side variable of interest is instrumented FNC viewership relative to averaged CNN and MSNBC viewership: *Viewership (FNC - MSNBC - CNN)*. All columns include state fixed effects and demographic controls as listed in Appendix Table A2. Column 2 also includes channel controls (population shares with access to each of the three TV channels). Column 3 controls for generic newspaper language features (vocabulary size, avg. word length, avg. sentence length, avg. article length). Standard errors are clustered at the state level (in parenthesis): * p < 0.1, ** p < 0.05, *** p < 0.01.

J. Robustness: Alternative County Matching

Table A16: Cable News Effects on Newspaper Content (2SLS) – Alternative Matching Procedure

<i>Dep. variable: $\widehat{FNC}_{jik}=P(FNC Text_{jik})$</i>	(1)	(2)	(3)
<i>Viewership (FNC - CNN)</i>	0.362 (0.224)	0.406* (0.237)	0.406* (0.238)
F-test	43.498	41.551	40.294
N observations	682	682	682
State FE	X	X	X
Demographic controls	X	X	X
Channel controls		X	X
Newspaper language controls			X

Notes: 2SLS estimates. Cross-section with newspaper-level observations weighted their total circulation. The dependent variable is newspaper language similarity with FNC (the average probability that a snippet from a newspaper is predicted to be from FNC): $\widehat{FNC}_{jik}=P(FNC|Text_{jik})$. The right-hand side variable of interest is instrumented FNC viewership relative to CNN viewership: *Viewership (FNC - CNN)*. All columns include state fixed effects and demographic controls as listed in Appendix Table A2. Column 2 also includes channel controls (population shares with access to each of the three TV channels). Column 3 controls for generic newspaper language features (vocabulary size, avg. word length, avg. sentence length, avg. article length). Standard errors are clustered at the state level (in parenthesis): * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A17: Cable News Effects on Newspaper Content (2SLS) – Alternative Matching Procedure

<i>Dep. variable: $\widehat{FNC}_{jik}=P(FNC Text_{jik})$</i>	(1)	(2)	(3)
<i>Viewership (FNC - MSNBC)</i>	0.929** (0.391)	1.159** (0.523)	1.344** (0.591)
F-test	13.393	9.357	8.513
N observations	682	682	682
State FE	X	X	X
Demographic controls	X	X	X
Channel controls		X	X
Newspaper language controls			X

Notes: 2SLS estimates. Cross-section with newspaper-level observations weighted by their total. The dependent variable is newspaper language similarity with FNC (the average probability that a snippet from a newspaper is predicted to be from FNC): $\widehat{FNC}_{jik}=P(FNC|Text_{jik})$. The right-hand side variable of interest is instrumented FNC viewership relative to CNN viewership: *Viewership (FNC - MSNBC)*. All columns include state fixed effects and demographic controls as listed in Appendix Table A2. Column 2 also includes channel controls (population shares with access to each of the three TV channels). Column 3 controls for generic newspaper language features (vocabulary size, avg. word length, avg. sentence length, avg. article length). Standard errors are clustered at the state level (in parenthesis): * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A18: Cable News Effects on Newspaper Content (2SLS) – Alternative Matching Procedure

<i>Dep. variable: $\widehat{FNC}_{jik}=P(FNC Text_{jik})$</i>	(1)	(2)	(3)
<i>Viewership FNC</i>	0.320 (0.324)	0.472 (0.404)	0.357 (0.388)
F-test	21.163	14.762	15.094
N observations	682	682	682
State FE	X	X	X
Demographic controls	X	X	X
Channel controls		X	X
Newspaper language controls			X

Notes: 2SLS estimates. Cross-section with newspaper-level observations weighted by their total circulation. The dependent variable is newspaper language similarity with FNC (the average probability that a snippet from a newspaper is predicted to be from FNC): $\widehat{FNC}_{jik}=P(FNC|Text_{jik})$. The right-hand side variable of interest is instrumented FNC viewership: *Viewership FNC*. All columns include state fixed effects and demographic controls as listed in Appendix Table A2. Column 2 also includes channel controls (population shares with access to each of the three TV channels). Column 3 controls for generic newspaper language features (vocabulary size, avg. word length, avg. sentence length, avg. article length). Standard errors are multiway-clustered at the state level (in parenthesis): * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A19: 2SLS: Cable News Effects on Text Readability Metrics

<i>Dep. variable</i>	<i>Vocab. size</i>	<i>Len. words</i>	<i>Len. sent</i>	<i>Len. art</i>
<i>Viewership (FNC - MSNBC - CNN)</i>	-0.227 (0.543)	0.885 (0.740)	0.154 (0.393)	0.863 (0.554)
F-test	36.380	36.380	36.380	36.380
N observations	3781	3781	3781	3781
State FE	X	X	X	X
Demographic controls	X	X	X	X
Channel controls	X	X	X	X
Corpus size control	X	X	X	X

Notes: 2SLS estimates. Cross-section with newspaper-county-level observations weighted by newspaper circulation in each county. The dependent variable is vocabulary size in column 1, average word length in column 2, average sentence length in column 3, and average total article length in column 4. The right-hand side variable of interest is instrumented FNC viewership relative to averaged CNN and MSNBC viewership: (*Viewership (FNC - MSNBC - CNN)*). All columns include state fixed effects, demographic controls as listed in Appendix Table A2, channel controls (population shares with access to each of the three TV channels), and a control for the size of the newspaper-specific corpus. Standard errors, multiway-clustered at the county and at the newspaper level, in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

K. Robustness: Effects of Cable News on Generic Language Features

In Table A19, re-run our main specification, but instead of bigram-based similarity with FNC, we regress vocabulary size (normalized by the total size of the corpus, column 1), average word length (column 2), average sentence length (column 3), and average article length (column 4) on instrumented FNC viewership relative to MSNBC and CNN. As before, we include demographic and channel controls. We also account for the size of the newspaper-specific corpus.¹⁷ None of the coefficients are significant or close to significant. These results are consistent with the interpretation that our main results are driven by FNC-specific bigrams that diffuse into local newspaper language.¹⁸

¹⁷The number of articles scraped is given by the availability on NewsLibrary. It does not seem to follow a pattern: correlation between corpus size and circulation by newspaper is rather small, around 0.3. The correlation between similarity with FNC and corpus size is, if anything, negative (around -0.21).

¹⁸The insignificance of the coefficients in Table A19 should not come as a surprise given that the main results in Table 4 barely change when we move from column 2 to column 3 (where generic newspaper language controls are introduced).