Conservative News Media and Criminal Justice: Evidence from Exposure to Fox News Channel*

Elliott Ash and Michael Poyker

October 30, 2019

Abstract

Exposure to conservative news causes judges to impose harsher criminal sentences. Our evidence comes from an instrumental variables analysis, where randomness in television channel positioning across localities induces exogenous variation in exposure to Fox News Channel. These treatment data on news viewership are taken to outcomes data on almost 7 million criminal sentencing decisions in the United States for the years 2005–2017. Higher Fox News viewership increases incarceration length, and the effect is stronger for black defendants and for drug-related crimes. We can rule out changes in the behavior of police, prosecutors, or potential offenders as significant drivers. Consistent with changes in voter attitudes as the key mechanism, the effect on sentencing harshness is observed for elected, but not appointed, judges.

Keywords: Partisan Media, Judge Elections, Incarceration, Racial Bias

JEL Codes: D72, H76, K41, L82

*Elliott Ash: ETH Zürich, ashe@ethz.ch; Michael Poyker (corresponding author): Columbia University, mp3780@columbia.edu. We are grateful to Sergio Galletta and Greg Martin for help with understanding the datasets. We thank Mike Mueller-Smith, Daniel Nagin, Arnaud Philippe, Ali Yurukoglu, conference participants at TWEC, and seminar participants at CMU Heinz and ETH Zürich for helpful comments. We thank David Cai, Romina Jafarian, and Matteo Pinna for helpful research assistance. All errors are ours.
1 Introduction

A recent literature has documented that greater exposure to partisan television news has an impact on voting in presidential elections (DellaVigna and Kaplan, 2007; Martin and Yurukoglu, 2017) and congressional position-taking (Clinton and Enamorado, 2014; Arceneaux et al., 2016). An unexamined question is whether partisan news would have an effect on judge decision-making. The goal of this paper is to provide the first evidence on this issue.

If judges are apolitical and make their decisions without regard to outside influences, partisan news exposure should have no effect (see, e.g., Posner, 2008; Epstein, Landes and Posner, 2013). But recent empirical work has documented that judges do respond to non-legal influences, political and otherwise (Berdejó and Yuchtman, 2013, Ash and MacLeod, 2015, 2017, Chen, Moskowitz and Shue, 2016, Berdejó and Chen, 2017, and Cohen and Yang, 2018). In addition, there is evidence suggesting that the judiciary has become more conservative over time (e.g., Ash, Chen and Naidu, 2017, Ash, Chen and Lu, 2017). This research asks whether we can attribute a causal influence to partisan news media in this trend.

The empirical context is criminal courts in U.S. states for the years 2005 through 2017. We use combined microdata on criminal sentencing decisions from the National Corrections Reporting Program (hereafter, NCRP) and a unique dataset with the universe of sentencing decisions linked to judge biographies from ten states (Poyker and Dippel, 2019), paired with data on cable news viewership at the county level. The measure of conservative news exposure is Fox News viewer-ship (relative to other cable news networks), where exogenous variation comes from the channel positioning of Fox News across counties. As demonstrated in Martin and Yurukoglu (2017), this channel-number variation can be used as an instrument for TV viewership across channels.

We replicate the strong first stage from Martin and Yurukoglu (2017) at the county level in our sample of states. We document that current Fox channel position is unrelated to preexisting markers for conservative policy, such as historical Republican vote shares, past crime rates, or past sentencing rates.

We use the first-stage prediction for Fox News viewership to estimate the impact on criminal sentencing outcomes in a two-stage-least-squares (2SLS) framework. We find that an exogenous increase in Fox News exposure is associated with an increase in criminal sentence length. We find no effect on the extensive margin; i.e., probability of being sent to prison. The result is robust to the inclusion of rich demographic controls and case controls, and to including controls for viewership of other cable news networks (CNN and MSNBC). The results hold using within-judge across-time variation in ratings, suggesting that conservative media works through treating incumbent judges rather than selection of conservative judges.

An immediate mechanism question is whether the effect is really going through judges, rather than other actors in the legal system. Media could affect legislators (either directly or through voter preferences), which would result in harsher laws. We can rule this out in our context because our empirical specification uses state-year interacted fixed effects, as well as state-specific crime and recidivism characteristics, which would absorb any state-specific changes in legislation. Another
possibility would be that police officers, after watching Fox News, might arrest more people or focus on more serious crimes, while prosecutors might become more aggressive in charging decisions. We do not find effects of Fox News on factors that can be affected by police and prosecutors but not judges: i.e., the number or types of charges that defendants face. Finally, we find no effect on crime rates, meaning that Fox News does not seem to have an effect on potential offenders (through changing local policies, for example).

Given that our effect goes through judges, it could be working by changing judge preferences or — for elected judges — by changing voter preferences. Voters might become more conservative due to Fox News exposure, and in particular due to media attention on felony cases. Meanwhile, lawyers/prosecutors put active pressure on judges threatening to find candidates to displace them (Berdejó and Yuchtman, 2013); that would increase electoral pressures on judges to be harsher in sentencing decisions. To distinguish direct judge effects from voter preferences effects, we run regressions separately for elected and appointed judges. The appointed judges have tenure, and therefore face minimal political pressures once in office. We find that Fox News increases sentencing only for elected judges, and not for appointed judges. We also use within-state variation in judge selection process in Kansas to show that the effect exists only in counties with elected judges. These results are consistent with voter attitudes providing a possible mechanism for our effects. However, we find evidence that Fox News has a smaller effect on judges that are closer to election, suggesting that electoral pressures are substitutable (rather than complementary) with partisan media pressures.

We further investigate the nature of the mechanism by using the transcripts from Fox News shows to measure mentions of crime. We interact the number of monthly crime mentions with Fox News consumption and show that our whole effect is explained by this interaction. It is not just Fox News driving sentencing harshness; it is Fox’s mentions of crime on the air. This effect is focused on the month of sentencing, rather than neighboring months (e.g., the month the crime was committed), adding further support for the judge’s decision-making (rather than, for example, the defendant’s) as the relevant mechanism.

In heterogeneity analysis, we find that the effect is larger for black defendants than for white or Hispanic defendants, and the effect is larger for drug-related crimes. Following up with additional text analysis of the show transcripts, we show that the Fox News effect is stronger when there are more mentions of drugs, or when the word “Black” is mentioned together with crime or drug words. Thus, by radicalizing voters with news about “Black crimes” Fox News appear to accentuate existing racial biases in sentencing decisions in states with elected judges.

These results will be of interest to scholars in empirical political economy, and in particular for those who study courts and the mass media (e.g., Lim, Snyder Jr and Strömberg, 2015 and Mastrorocco and Minale, 2018). Using French data, Philippe and Ouss, 2018 find that defendants whose cases received attention in the media receive longer sentences. Our contribution is that besides the direct behavioral effect on juries through viewing stories about a case, media have an additional effect by changing preferences of voters and exercising electoral pressures on judges. Building on the cross-sectional evidence for a partisan gap in racial sentencing disparities (Cohen
and Yang, 2018), our estimates have a causal interpretation for shifts in ideology. The findings are relevant to recent debates on how judges should be selected, retained, and compensated (Epstein, Landes and Posner, 2013; Ash and MacLeod, 2017), along with recent debates on polarization and media regulation (Boxell, Gentzkow and Shapiro, 2017; Allcott and Gentzkow, 2017). We provide causal evidence that incentives of elected judges are distorted as they cater to voters as politicians, rather than operating on facts and laws (Kessler and Piehl, 1998).

The rest of this paper is organized as follows. Section 2 provides background information about Fox News. Section 3 describes the data. Section 4 presents our identification strategy. Section 5 contains estimation results. Section 6 looks to mechanism by analyzing judicial elections and cable news show language. Section 7 concludes.

2 Background

This paper is motivated by previous evidence that Fox News is conservative, and the ongoing discourse on how conservative media impact social attitudes and policy outcomes. Figure 1 shows three pieces of evidence on this point. First, Panel A, from Martin and Yurukoglu (2017), shows that Fox News tends to use politicized phrases associated with Republican politicians. Second, in Panel B, we see that for the years 2005–2008, Fox speakers mention crime more often than speakers on CNN and MSNBC.¹ Third, in Panel C, we show in our data that places with higher Fox News ratings share tend to impose longer criminal sentences.²

An additional piece of cross-sectional evidence on how Fox News is related to criminal justice is reported in Panel D. To make this graph, we produced average sentence length metrics by court. We then plotted the trends in sentence length separately by quartiles in Fox News viewership (for the years 2005–2008). We can see that in the places with more Fox News viewership, there was a much larger jump in sentencing lengths starting in 2009. This is some descriptive evidence that places with more Fox News exposure had harsher criminal justice outcomes. The question, for this paper, is whether this correlation in the courts is due to a causal link.

To better understand the crime-related discourse of Fox News, we used natural language processing tools to understand the language associations in cable news shows. We trained word2vec, a popular word embedding model (Mikolov et al., 2013), on transcripts for Fox, CNN, and MSNBC, for the years 2001 through 2013. This model works by reading through sentences and locating words close to each other in a vector space if they tend to occur in similar contexts (that is, windows of neighboring words). Similarity between words can then be measured using the cosine of the angle between the vector representations of each word. In the transcripts data, the most sim-

¹These are counts of “crime,” “criminal,” “murder,” and “homicide,” divided by the number of spoken sentences, in transcripts for prime time shows for each network. These years were used because we had transcripts data for all three networks.

²We don’t take a position on whether Fox News policy advocacy is “biased” away from some optimum; we are only speaking relative to the CNN and MSNBC reference point. In addition, we don’t take a position on the motivations underlying this advocacy; it could be due to political motivations, due to trying to get more viewers, or for other reasons.
Figure 1: Fox News is Conservative and Correlates with Sentencing Length

Notes: Illustrations for Fox News conservatism. Panel A is predicted ideology based on political phrases used by Republicans and Democrats. Panel B is the number of references to crime per sentence spoken in cable news transcripts. Panel C is a binscatter for the OLS correlation between incarceration length and Fox Nielsen rating. Online Appendix Figure 2, Online Appendix Figure 3, and Online Appendix Figure 4 contain results for CNN and MSNBC. Panel D correlates Fox News consumption and average sentence length; the figure shows average sentence length in top 25%, bottom 25%, and middle 50% of the counties by Fox News consumption.
ilar words to “crime” were “crimes,” “murder,” “homicide,” “perpetrator,” “felonies,” and other synonyms or closely related terms.

What is most interesting for our purposes is the differences in word associations across the networks. To get the crime words most associated for Fox News, for example, we take the Fox cosine similarity and divide by the average of the similarities for CNN and MSNBC. We computed a symmetric measure for CNN and MSNBC. We then ranked the most associated words for each network.

Word clouds illustrating the most crime-associated words for each of the three networks are reported in Figure 2. The results are striking. One can see immediately in Panel A that at Fox, discourse on crime is racialized. The highest-associated term is “black-on-white,” and “white-on-black” is also highly ranked. Other words seem to personalize crime victimization: “victimize,” “muggings.” They also arguably demean the accused: “perps” and “priors.”

The word clouds for CNN (Panel B) and MSNBC (Panel C) have very different flavors. One can see that CNN focuses on criminal organizations and conspiratorial language: “terrorism,” “mobsters,” and “underworld.” The top terms from MSNBC, “[Daryll] Littlejohn” and “Imette [St. Guillen],” respectively refer to the defendant and victim of a particular sensational New York City murder case from 2006.

3 Data

3.1 Sentencing Data

The data on sentencing come from the National Corrections Reporting Program (ICPSR 36373, hereafter NCRP). This is a standard dataset for the literature and it contains data for all prison admissions in the United States from 2000 to 2014.

This dataset has several important characteristics that make it crucial for our study. As it spans all U.S. states, it gives us more variation in our main explanatory variable at the county level. In addition, we have eight years of overlap for our explanatory variable (2005 to 2008 and 2010 to 2014).

Our main outcome variable is the length of sentence imposed. We also use defendant and case characteristics. The seriousness of a crime is one of the main features of the judgment of a court, and the classification of offenses in the NCRP is standardized. Therefore we include in our regressions a matrix of 180 fixed effects for the offense with the longest sentence length. We also include criminal history (recidivism), education, military background, and demographic characteristics, including age at conviction, gender, and race (Asian, Black, Hispanic, Native American, White, and other).

We supplement our NCRP data with the sentencing data from Poyker and Dippel (2019). It is superior to the NCRP’s dataset in the sense that it has (i) case-level information on each sentencing

Panel A. Most Similar Words to “Crime”: Fox News

Panel B. Most Similar Words to “Crime”: CNN

Panel C. Most Similar Words to “Crime”: MSNBC

Figure 2: Crime Discourse in U.S. Cable News Channels

Notes: Most closely related terms to “crime,” in Fox, CNN, and MSNBC, respectively. Similarities computed from word2vec models trained separately on the transcript corpora for each network. Larger words mean the word has higher similarity for the indicated network and lower similarity for the other two networks.
decision; (ii) information on the defendants that were found not guilty or did not go to prison (e.g., went on probation); (iii) information on judges. Here, we only use data from the 10 states that have judges’ information: Alabama, Colorado, Georgia, Kentucky, Minnesota, North Carolina, Pennsylvania, Tennessee, Virginia, and Washington. The years covered by the convictions data vary from state to state but range from 1980 to 2017.

To construct the length of sentence imposed we assign zero for all cases in which the defendant is found not guilty or put on probation. In the case of consecutive sentences, those are summed. In the case of concurrent sentences, we take the max.

The classification of offenses varies across states and trying to harmonize them would be complex and require many subjective decisions. Therefore we include in our regressions a separate set of offense class fixed effects for each state. Data also contains information on recidivism and basic demographic characteristics, including age at conviction, gender, and race.

Finally, we have information on judicial elections for a subset of states. To construct electoral cycles we use judges’ biographies from www.ballotpedia.org linked to the sentencing dataset.4 In four states (Arizona, Indiana, Kansas, and Missouri), we have the list of counties where judges are elected and those that are appointed.

3.2 Media Data

The data on channel positions and ratings come from Nielsen. This is an expanded version of the dataset in Martin and Yurukoglu (2017). The data includes channel listings by system and year, with associated zip codes, for the years 1998 through 2017. It includes zip code level viewership for Fox, CNN, and MSNBC for the years 2005 through 2008. It includes Designated Market Area (hereafter, DMA) level data on viewership for the years 2010 through 2017. Because our observation is a sentence in state trial courts mapped to counties,5 we aggregate our treatment (viewership for Fox, CNN, and MSNBC) and instrument (channel positions) to the county level.6

The viewership data are for all shows on the networks, so they include a collection of “news” shows (which claim to report straight news) and “pundit” shows (which have an acknowledged political viewpoint).7 For the text illustrations (shown previously), we downloaded full-text transcripts of prime time shows on Fox News, CNN, and MSNBC from LexisNexis. For the Google Trends results, those are downloaded from the Google Trends web site and matched on DMA.

4 Empirical Specification and Identification

The identification strategy adopts an instrumental variables approach based on Martin and Yurukoglu (2017). The instrument relies on exogenous variation in where Fox News Channel appears

---

4See details of the construction of electoral cycles and linking in Dippel and Poyker (2019).
5Their label varies; in some states they are labeled circuit courts, district courts, or superior courts, but they are always identified as being above the courts of limited jurisdiction and below the state appellate courts.
6We use population weights; however, all results hold if we don’t weight by population.
7More precisely, the 2005–2008 ratings data is for all shows, while the 2010–2017 ratings data are for prime time shows. This is what Nielsen made available to us through the data purchase.
in the channel number lineup across counties relative to other cable news channels. Counties are the lowest-level geographical unit for the sentencing data (i.e., courts); in most states judicial districts are composed of multiple counties. Throughout the paper, an observation is a sentencing decision \(i\) that took place in county \(c\) of state \(s\) at year \(t\).

The first-stage estimating equation is:

\[
T_{ct} = \alpha_{st} + \gamma Z_{ct} + X_{i(c)t}\beta + \eta_{i(c)t},
\]

where \(T_{ct}\) is the time spent watching Fox News as a share of total television watched (share) for county \(c\) at time \(t\), \(\alpha_{st}\) includes state-year (interacted) fixed effects, \(Z_{ct}\) is the channel number for Fox News, and \(\eta_{i(c)t}\) is an error term. \(X_{i(c)t}\) includes other covariates describing demographics and cable system characteristics. Legislated trends, which are at the state level, will be absorbed by \(\alpha_{st}\) taking the form of state-year fixed effects in our preferred specification. \(^9\) From Martin and Yurukoglu (2017) we expect a negative and significant estimate for \(\gamma\).

The second-stage estimating equation models an outcome \(Y_{i(c)t}\) (e.g., criminal sentencing harshness for case \(i\) in courthouse/county \(c\) at time \(t\)) as:

\[
Y_{i(c)t} = \alpha_{st} + \rho T_{ct} + X_{i(c)t}\beta + \mu_{st} + \epsilon_{i(c)t},
\]

where the terms are the same as in equation (1) and \(\epsilon_{i(c)t}\) is the error term. The identification assumption is that conditional on the fixed effects and covariates, the channel position \(Z_{ct}\) affects the outcome \(Y_{i(c)t}\) only through Fox News viewership \(T_{ct}\). As treatment is on the county level, we cluster standard errors by county. \(^{10}\)

We prefer decision-level data (case \(i\) and year \(t\)) rather than aggregating to the court-year-level (court \(c\) and year \(t\)) for three reasons. First, we expect that Fox News may influence views on harshness toward particular types of defendants or types of criminal offenses (see Panel A of Figure 2). Second, there are different mixes of offense types across jurisdictions. Third, NCRP data are not representative across counties. \(^{11}\)

We require instrument relevance. Online Appendix Figure 1 Panel A shows graphical evidence of the first-stage variation we are using. There is a clear downward trend, with higher channel numbers having lower viewership. In the regression tables below, we report the F-statistic of excluded instruments for each regression, and they are consistently greater than 10. \(^{12}\)

Martin and Yurukoglu (2017) provide a lengthy discussion and set of checks for exogeneity of

\(^8\)Note that for the 2010-2017 data sample, ratings are at the DMA level. For some states (Connecticut, Hawaii, New Mexico, Rhode Island, Utah, Vermont, and Washington D.C.), DMAs coincide with states. So these are effectively dropped.

\(^9\)This means identification comes from within-year within-state between-county variation in Fox News consumption and Fox News channel position.

\(^{10}\)Our results hold if we cluster by state, DMA, or by county and year. These results are available upon request.

\(^{11}\)The latter is a lesser concern in Section 6.2 where we use the universe of sentencing decisions for 10 states.

\(^{12}\)Online Appendix Figure 1 Panel B shows a binscatter for the reduced form. The vertical axis is the outcome (log incarceration length) and the horizontal axis is Fox channel position. We can see a negative relation, reflecting that in counties with a lower Fox channel position judges tend to be harsher in sentencing.
the instrument (Fox channel position). In our data, as in theirs, the instrument is not related to past Republican vote shares. In our data, moreover, channel position is not related to past crime rates, nor is it related to past sentencing harshness. To further support the exogeneity assumption, in Robustness Section 5.2 we include these covariates as controls.

Under exogeneity, two-stage least squares procures consistent estimates for $\rho$ if the instrument satisfies an exclusion restriction. That is, the channel position affects sentencing decisions only through its effect on Fox News viewership. We feel this is a reasonable assumption in our context. In particular, we include state-year fixed effects in all specifications to control for the changes in state-level legislation: it allows us to rule out an effect of conservative media through changes in legislation.

A final assumption is monotonicity. That is, a lower Fox News Channel would not decrease Fox News viewership. Again, we feel this is a reasonable assumption in our institutional context. Still, we have performed a series of checks to see that the first stage is satisfied in subsets of the data.\footnote{We report a histogram of the first-stage coefficients in Online Appendix Figure 5.}

5 Results

This section reports the main results of the regression analysis. Subsection 5.1 reports the main results. Subsection 5.2 contains robustness and sensitivity checks. Subsection 5.3 reports key heterogeneity analysis.

5.1 Main Results

The main results are presented in Table 1. We report OLS estimates for the specification with only demographic controls and state-year fixed effects in Column I. As we already saw in the binscatter from the bottom panel of Figure 1, the share of Fox News viewers is positively associated with the sentence length without any controls. We report corresponding 2SLS estimates in Column II. The first stage is strong ($F = 33.2$), and the coefficient is positive and statistically significant.\footnote{Online Appendix Figure 1 contains first stage and reduced form binscatter plots. Reduced form results are shown in Column I of 3.} In Column III we add case controls, which make the coefficient about 38% larger.\footnote{A specification with aggregated data (by court $c$ and year $t$) would mix different offense types across jurisdictions. However, as a sanity check, aggregated specifications yield similar results to those in Columns I–II; i.e., without case controls.} Defendant characteristics and case characteristics are likely correlated with Fox News viewer share, so they soak up the residual variation and strengthen estimates.

The coefficient for 2SLS is somewhat larger than for OLS. Martin and Yurukoglu (2017) observe a similar pattern in their results. They offer two explanations: that the IV reduces the attenuation bias due to measurement error in the ratings data, and that it estimates a local average affect among complier zip codes, i.e., those whose Fox News viewership is most sensitive to channel position. Complier zip codes are likely to be those with the weakest attachments to either party, and hence
Table 1: Fox News and Sentencing Decisions

<table>
<thead>
<tr>
<th>Nielsen share</th>
<th>OLS</th>
<th>2SLS</th>
<th>2SLS</th>
<th>2SLS</th>
<th>2SLS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fox news</td>
<td>0.016***</td>
<td>0.054*</td>
<td>0.075***</td>
<td>0.075***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0044)</td>
<td>(0.029)</td>
<td>(0.028)</td>
<td>(0.029)</td>
<td></td>
</tr>
<tr>
<td>Fox - (CNN+MSNBC)/2</td>
<td>0.078***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.030)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CNN</td>
<td>-0.004</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSNBC</td>
<td>-0.009</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.012)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demographic controls</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Case controls</td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Partial R-squared</td>
<td>0.029</td>
<td>0.029</td>
<td>0.026</td>
<td>0.026</td>
<td></td>
</tr>
<tr>
<td>F-stat. of excl. inst.</td>
<td>33.2</td>
<td>33.1</td>
<td>36.4</td>
<td>34.7</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>4,974,207</td>
<td>4,974,207</td>
<td>4,974,207</td>
<td>4,974,207</td>
<td>4,974,207</td>
</tr>
</tbody>
</table>

Notes: All columns include state-year FEs. The dependent variable is the inverse hyperbolic sine of the sentencing length. The following variables are used as controls: age, age squared, and race dummies (Black, Hispanic, Asian, Native American, and other), dummy for military background, recidivists and set of crime severity and education dummies. Standard errors clustered by county are in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1

partisan persuasion effects are plausibly larger in those places. In our setting, the former bias may be less of a concern because when aggregating to the county level there is less idiosyncratic measurement error in the Nielsen ratings (compared to zip codes, where Nielsen can have low-single-digit numbers of surveyed households). The latter explanation still holds, as judges may be more sensitive to the voters’ opinion in the swing counties.

The estimates could be confounded by viewership for other cable news networks, CNN and MSNBC. To zero in on Fox News, Column IV provides results with an alternative specification for viewership which is normalized relative to CNN and MSNBC.\(^{16}\) Finally, Column V includes CNN and MSNBC viewer share as (non-excluded) controls. In each of these alternative specifications, the 2SLS estimates are comparable to Column III: positive and statistically significant.\(^{17}\)

Are these estimates economically significant? According to the estimate from Column V, increasing Fox News share by 1 percent (one standard deviation is 1.3 percent) would increase average sentencing length by 7 percentage points (about 4 months).\(^{18}\) As a baseline, consider the evidence from Cohen and Yang (2018) that in federal courts, Republican-appointed judges give about 2-months-longer sentences on average than Democrat-appointed judges.

---

\(^{16}\)It is constructed as \(\text{share}_{\text{FOX}} = \text{share}_{\text{FOX}} - \frac{1}{2}(\text{share}_{\text{CNN}} + \text{share}_{\text{MSNBC}})\).

\(^{17}\)The Online Appendix Figure 2 reports graphical evidence for effects of CNN and MSNBC on sentencing. We do not see the same effects for these networks.

\(^{18}\)The equivalent effect for OLS (Column I) is smaller: 28 days.
5.2 Robustness and Sensitivity Checks

To address possible exclusion restriction violations, we report a number of additional regressions in Table 2. In Columns II–V we add control variables that may correlate with channel position and conservativism on crime: past Republican vote share, share of rural population, past crime rates, and past sentencing. While some of these variables appear to be significant and have expected signs, our results hold. They also hold when we include all of these controls together in Column VI.

Table 2: Fox and Sentencing: Robustness to Inclusion of Additional Controls

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>VI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nielsen share (Fox)</td>
<td>0.075***</td>
<td>0.074**</td>
<td>0.081**</td>
<td>0.071***</td>
<td>0.081**</td>
<td>0.079***</td>
</tr>
<tr>
<td></td>
<td>(0.028)</td>
<td>(0.029)</td>
<td>(0.033)</td>
<td>(0.023)</td>
<td>(0.033)</td>
<td>(0.028)</td>
</tr>
<tr>
<td>1996 Republican vote share</td>
<td>0.166</td>
<td>-0.008</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.147)</td>
<td>(0.110)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log population</td>
<td>-0.036</td>
<td>-0.037</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.065)</td>
<td>(0.069)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-1 avg. sentencing length</td>
<td>0.011***</td>
<td>0.010***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.0003)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-1 crime rates</td>
<td>-0.027</td>
<td>-0.002</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.049)</td>
<td>(0.050)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partial R-squared</td>
<td>0.029</td>
<td>0.028</td>
<td>0.017</td>
<td>0.022</td>
<td>0.017</td>
<td>0.016</td>
</tr>
<tr>
<td>F-stat. of excl. inst.</td>
<td>33.1</td>
<td>32.6</td>
<td>25.2</td>
<td>27.1</td>
<td>25.2</td>
<td>23.2</td>
</tr>
<tr>
<td>Observations</td>
<td>4,974,207</td>
<td>4,889,187</td>
<td>3,384,386</td>
<td>4,454,828</td>
<td>3,384,386</td>
<td>3,314,318</td>
</tr>
</tbody>
</table>

Notes: All columns include state-year FEs. The dependent variable is the inverse hyperbolic sine of the sentencing length. All Columns use baseline specification from Column III of Table 5.1. Standard errors clustered by county are in parentheses. ** p<0.01, * p<0.05, * p<0.1

Next, we provide additional support for a significant effect through a permutation test. To this end, we permute the Fox News channel positions with replacement 500 times. Online Appendix Figure 6 compares our true reduced form estimate to the distribution of estimates obtained from regressing sentence length on fake Fox channel position 500 times; the true coefficient has by far the largest magnitude.

To further probe the exclusion restriction, we use the fact that Fox News became most conservative only starting in 2005 (Martin and Yurukoglu, 2017). Before 2005, Fox News should not contribute much to increases in sentencing. In Column II of Table 3 we show that Fox News channel position had no reduced form effect on sentencing length before 2005.19. As a set of placebo exercises, we also show that the positions for Golf channel, Playboy, Trinity Broadcast Network (Christian TV), and A&E (specializes in broadcasting true crime shows) have no reduced-form

19Column I provides the baseline reduced-form estimate for comparison. We use the reduced form, rather than 2SLS, for this check because we don’t have local viewership data before 2005.
Table 3: Channel Positions and Sentencing: Reduced Form and Placebo

<table>
<thead>
<tr>
<th>Channel name</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>VI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample</td>
<td>FOX News</td>
<td>Golf</td>
<td>Playboy</td>
<td>Trinity BN</td>
<td>A&amp;E (crime ch.)</td>
<td></td>
</tr>
<tr>
<td>Channel position</td>
<td>-0.0009**</td>
<td>-0.0006</td>
<td>0.0001</td>
<td>-0.00002</td>
<td>-0.0002</td>
<td>-0.0004</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.393</td>
<td>0.461</td>
<td>0.460</td>
<td>0.493</td>
<td>0.461</td>
<td>0.461</td>
</tr>
<tr>
<td>Observations</td>
<td>4,974,207</td>
<td>3,003,437</td>
<td>2,946,300</td>
<td>1,809,236</td>
<td>3,002,924</td>
<td>3,003,437</td>
</tr>
</tbody>
</table>

Notes: All columns include state-year FEs. In Column II we use FOX News channel position before 2005 as a placebo because it was not so exceptionally conservative before 2005 (Martin and Yurukoglu, 2017). The dependent variable is the inverse hyperbolic sine of the sentencing length. All columns use baseline specification from Column III of Table 5.1. Standard errors clustered by county are in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1

5.3 Heterogeneous Effects by Defendant Characteristics

We are also interested in estimating heterogeneous effects. In particular, we check whether media slant is disproportionately affecting minorities, women, or certain non-violent types of crime. To do so we adapt the 2SLS specification as follows. The second stage is:

\[ Y_{i(c)t} = \alpha_{s(c)t} + \rho_1 T_{ct} + \rho_2 T_{ct} \times \mu_{i(c)} + \mu_{i(c)} + X_{i(c)t} \beta + \epsilon_{i(c)t}, \]  

(3)

where \( \mu_{i(c)} \) is a characteristic for defendant \( i \) (e.g., race category or crime category). The coefficients of interest are the baseline effect of media consumption, \( \rho_1 \), plus the interaction effect with the defendant’s characteristic that might be targeted by the media, \( \rho_2 \).

There are two endogenous variables. The first stage consists of

\[ T_{ct} = \alpha_{s(c)t} + \gamma_1 Z_{ct} + \gamma_2 Z_{ct} \times \mu_{i(c)} + \mu_{i(c)} + X_{i(c)t} \beta + \eta_{i(c)t}^1 \quad \text{and} \]  

(4)

\[ T_{ct} \times \mu_{i(c)} = \alpha_{s(c)t} + \gamma_1 Z_{ct} + \gamma_2 Z_{ct} \times \mu_{i(c)} + \mu_{i(c)} + X_{i(c)t} \beta + \eta_{i(c)t}^2, \]  

(5)

where the two excluded instruments are the channel position \( Z_{ct} \), plus the channel position interacted with the defendant characteristic, \( Z_{ct} \times \mu_{i(c)} \). These first stages are used to predict the endogenous regressors.

Table 4 reports possible heterogeneous effects by providing estimates for Eq. 3. Column I shows that our effect goes almost entirely through black defendants. The coefficient for \( \rho_1 \) becomes insignificant, while \( \rho_2 \) is large, suggesting racial bias in the effect of Fox News on sentencing.
Table 4: Fox News and Sentencing Decisions: Heterogeneity by Defendant Characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>I (Black)</th>
<th>II (Hispanic)</th>
<th>III (Female)</th>
<th>IV (Drug-related crimes)</th>
<th>V (DUI crimes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nielsen share (Fox)</td>
<td>0.006</td>
<td>0.078***</td>
<td>0.067**</td>
<td>0.014</td>
<td>0.075***</td>
</tr>
<tr>
<td></td>
<td>(0.027)</td>
<td>(0.030)</td>
<td>(0.028)</td>
<td>(0.031)</td>
<td>(0.028)</td>
</tr>
<tr>
<td>Nielsen share x Characteristic</td>
<td>0.168**</td>
<td>-0.040</td>
<td>0.055</td>
<td>0.218***</td>
<td>-0.001</td>
</tr>
<tr>
<td></td>
<td>(0.073)</td>
<td>(0.070)</td>
<td>(0.036)</td>
<td>(0.074)</td>
<td>(0.076)</td>
</tr>
<tr>
<td>Characteristic</td>
<td>-0.328**</td>
<td>0.154</td>
<td>-0.265***</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.143)</td>
<td>(0.145)</td>
<td>(0.075)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Partial R-squared</td>
<td>0.013</td>
<td>0.031</td>
<td>0.023</td>
<td>0.017</td>
<td>0.029</td>
</tr>
<tr>
<td>F-stat. of excl. inst.</td>
<td>17.2 &amp; 15.2</td>
<td>34.2 &amp; 3.0</td>
<td>20.6 &amp; 20.4</td>
<td>19.3 &amp; 14.7</td>
<td>17.4 &amp; 19.4</td>
</tr>
<tr>
<td>Observations</td>
<td>4,974,207</td>
<td>4,974,207</td>
<td>4,974,207</td>
<td>4,974,207</td>
<td>4,974,207</td>
</tr>
</tbody>
</table>

Notes: All columns include state-year FEs. The dependent variable is the inverse hyperbolic sine of the sentencing length. All Columns use baseline specification from Column III of Table 5.1. Characteristic coefficients in Columns IV–V are absorbed by crime type fixed effects. We report F-statistics and partial $R^2$ for the first endogenous variable (Nielsen share, Fox). Standard errors clustered by county are in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1

We find no evidence for a disproportionate media effect toward Hispanics or toward female defendants (Columns II–III).

The empirical literature on U.S. criminal justice has emphasized that minorities are often disproportionately prosecuted for non-violent crimes (e.g., Fagan and Ash, 2017). And anecdotally, conservative discourse often concerns itself with the risks posed by illicit drugs and the associated informal economy. Motivated by these points, we estimate the relative effects of Fox News exposure for drug-related crimes (Column IV) and a non-illegal-drug placebo, driving under the influence of alcohol (DUI, Column V). We find that while there is no differential effect of Fox News on DUI crimes, there is a large interaction effect for drug crimes. The estimates in Column IV suggest that the effect operates almost entirely through drug-related crimes.

6 Analysis of Mechanisms

This section probes the mechanism of how conservative media affects sentencing decisions. Subsection 6.1 provides evidence against non-judicial channels, such as policing, prosecutors, and effects on crime rates. Subsection 6.2 provides evidence in favor of a judicial elections channel. Subsection 6.3 uses text analysis to pinpoint the effect of crime-related language in Fox News transcripts on sentencing.

20Interestingly, we don’t find a positive relationship between being black and sentence length, which is somewhat different than other work in this literature. It could be that blacks are more often arrested with less serious crimes (even within observed charge categories).
6.1 Is it Judges, Police, or Prosecutors?

Here we provide evidence that our results are not driven by the decisions of police, prosecutors, or potential offenders. To this end, Table 5 tests whether Fox News affected other sentencing-related outcomes. Column I shows no evidence for an effect of Fox News on probability of incarceration. Thus the effect exists only on the intensive margin but not on the extensive margin.

In Column II, we check whether Fox News consumption affects the number of offenses in each case.\textsuperscript{21} Because the initial number of charges does not depend on a judge, finding evidence that Fox News increases the number of offenses would mean that the effect is partially driven by the police officers and prosecutors who prepare the case for the court. However, we find no effect, providing additional evidence against the non-judicial channels (policing and prosecutors). Similarly, in Column III we find no evidence that Fox News affect the predicted sentence length based on charges, meaning that prosecutors do not appear to be bringing more serious charges due to media exposure.\textsuperscript{22}

\begin{table}[h]
\begin{center}
\caption{Fox News and Additional Outcomes}
\begin{tabular}{lcccc}
\hline
 & I & II & III & IV \\
\hline
\text{Dependent variable:} &  &  &  &  \\
\text{Nielsen share (Fox)} & -0.014 & 0.701 & -0.001 & 0.00001 \\
 & (0.013) & (0.821) & (0.003) & (0.00003) \\
\text{Partial R-squared} & 0.094 & 0.157 & 0.094 & 0.017 \\
\text{F-stat. of excl. inst.} & 76.9 & 226.6 & 77.3 & 25.2 \\
\text{Observations} & 2,521,497 & 1,856,053 & 2,521,505 & 3,384,386 \\
\hline
\end{tabular}
\end{center}
\end{table}

Notes: In Columns I–III we use data from Poyker and Dippel (2019) and for the Column IV we use uniform crime reporting data (www.ucrdatatool.gov). All columns include state-year FE's. The dependent variable is the inverse hyperbolic sine of the sentencing length. All Columns use baseline specification from Column III of Table 5.1. Standard errors clustered by county are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

A final potential channel is through potential criminal offenders. This could occur through changes in local welfare policies, for example, as documented in Galletta and Ash (2019). More specifically, if Fox News affected voter preferences and local welfare policies, that may affect the opportunity costs of crime.\textsuperscript{23} However, we find no effect on crime rates (Column IV), nor on the severity of crimes as proxied by charges (Column III). Therefore there is no evidence of potential criminal choices being an important mechanism.

\textsuperscript{21}We do not observe the number of charges for Alabama, Virginia, and Washington.

\textsuperscript{22}Our main NRCP dataset only includes information on cases that received a non-zero sentence. In ten states we have newly collected data that includes all cases. Our results hold in these data, where we include cases with zero-length sentence. Therefore the effects do not seem to be due to differential selection of cases into jail time.

\textsuperscript{23}Voters’ effect on state legislation is absorbed by state-year fixed effects.
6.2 Judicial Elections

To the extent that Fox News influences judges, that could occur through two mechanisms. First, judges might themselves be watching Fox News and it affects their behavior directly. Second, voters might become more conservative due to Fox News and then influence judges through the election process. To explore this idea, we test whether the effect is different in states where judges are elected versus states with appointed judges.

Specifically, we split the sample in two (with elected-judge and appointed-judge states). We then estimate the effect of Fox News consumption on each subsample of states separately. Besides the sample split, the estimation approach is the same as that from Section 4.

The main results for judicial elections are reported in Table 6. Column I includes the baseline specification with controls (Table 1, Column III) for comparison. Column II reports results for the sample of states with appointed judges: the coefficient is the opposite sign and insignificant. Column III reports the coefficient for the subsample of states with elected judges. The coefficient of interest is significant and is not statistically different from the baseline coefficient. This result is consistent with the effects being driven by the election channel. Moreover, the absence of an effect in states with appointed judges provides additional evidence against policing and prosecutor channels. If police or prosecutors were important, we would expect appointment states to have positive (possibly smaller in magnitude) effects.

### Table 6: Fox News and Sentencing Decisions: Elected vs. Appointed Judges

<table>
<thead>
<tr>
<th>Sample</th>
<th>Dependent variable: Log sentencing length in months</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I</td>
</tr>
<tr>
<td></td>
<td>All</td>
</tr>
<tr>
<td></td>
<td>2SLS</td>
</tr>
<tr>
<td>Nielsen share (Fox)</td>
<td>0.075***</td>
</tr>
<tr>
<td></td>
<td>(0.028)</td>
</tr>
<tr>
<td>Partial R-squared</td>
<td>0.029</td>
</tr>
<tr>
<td>F-stat. of excl. inst.</td>
<td>31.9</td>
</tr>
<tr>
<td>Inst. F-stat. p-value</td>
<td>0.000</td>
</tr>
<tr>
<td>Observations</td>
<td>4,974,207</td>
</tr>
</tbody>
</table>

Notes: All columns include state-year FEs. The dependent variable is the inverse hyperbolic sine of the sentencing length. All columns use the baseline specification from Column III of Table 5.1. Standard errors clustered by county are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

24 The list of states with both appointed and elected judges can be found in Table 1 of Lim, Snyder and Strömberg (2015). We include four states (Arizona, Indiana, Kansas, and Missouri) with within-state variation in judge selection in both samples; however, results hold if we drop them or split by counties’ selection method.

25 Here, we do not discard the legislation channel, as it is conditioned by the state-year fixed effects. While we can’t rule it out, it does not drive our results.

26 The following states have appointed prosecutors: Alaska, Connecticut, New Jersey, and Washington D.C. The fact that we don’t find a significant coefficient might mean that media attention is mostly concentrated on judges rather than police or prosecutors.
Next we exploit the institutional context of Kansas, which has within-state variation in judicial selection. In roughly half of judicial districts, judges are elected in partisan elections, and in the other half judges are appointed (49 out of 105 counties). Identification using within-state variation in judge selection has been used in previous research, which has compared the counties in greater detail (Gordon and Huber, 2007; Lim, 2013; Park, 2017).

In Columns IV and V of Table 6, we respectively report OLS and 2SLS estimates of the baseline specification for Kansas. With only 9,104 observations from Kansas, OLS is insignificant. 2SLS is positive and significant, but the first stage is barely strong enough with F-statistic of 8.5. Finally, we split the state in samples of counties where judges are appointed (Column VI) and counties where judges are elected (Column VII). Similarly to the results in Columns II and III, we find positive and significant OLS effects for the subsample of counties with elected judges but no effect for counties where they are appointed. We do not report 2SLS results because we do not have enough power in the first stage.

To dig further into the election mechanism, we now look at dynamic effects of the electoral cycle. For this purpose we need judge identifiers, which are not available in the NCRP data. Therefore, we use newly collected data from ten states where those identifiers are included. Let $j$ be a judge. All judges are uniquely mapped to one county at any given time, and as a result case $i$ can be uniquely linked to judge $j$.

Following Berdejó and Yuchtman (2013) we construct a variable $\tau_j$ that measures proximity to election of judge $j$ at time $t$ as a linear running variable that is scaled from 0 to 1. It starts at 0 on the day after a general election, and equals 1 on the day of the next general election. It increases by $1/T_s$ each day, where $T_s$ is the length of state $s$’s electoral cycle, i.e., $T_{WA} = 4 \times 365 + 1$ and $T_{NC} = 8 \times 365 + 2$. Thus $\tau_j$ is scaled as $\tau_j \in [0; 1]$. Online Appendix Figure 7 provides a visualization of what an electoral cycle looks like in the data.

For this analysis we adapt the 2SLS specification by adding an interaction of Fox News viewership with a judge’s proximity to election as the second endogenous variable. In turn, we include an interaction of Fox channel position with electoral proximity as a second instrument. The first stages are as follows:

$$T_{ct} = \alpha_{st} + \gamma_1 Z_{ct} + \gamma_2 Z_{ct} \cdot \tau_j + \beta \tau_j + X_{i(c)t} \beta + \mu_j + \eta_{1i(c)t}, \quad (6)$$

$$T_{ct} \cdot \tau_j = \alpha_{st} + \gamma_1 Z_{ct} + \gamma_2 Z_{ct} \cdot \tau_j + \beta \tau_j + X_{i(c)t} \beta + \mu_j + \eta_{2i(c)t} \quad (7)$$

where $\tau_j$ defines proximity to elections. The second stage is

---

27 The list of counties with elected/appointed judges is available at [https://ballotpedia.org/Judicial_selection_in_Kansas](https://ballotpedia.org/Judicial_selection_in_Kansas).

28 Arizona, Indiana, and Missouri also have within state variation in selection of judges, but their counties are not comparable in terms of covariates (Park, 2017).

29 We trim the electoral cycles at the state-wide filing date, after which the electoral cycle effectively ends for the large majority of judges who have no challenger for their seat. More details on construction of the electoral cycles can be found in Dippel and Poyker (2019).
\[ Y_{i(c)t} = \alpha_{st} + \rho_1 T_{ct} + \rho_2 T_{c\cdot t} + \tau_j + X_{i(c)t} \beta + \mu_j + \epsilon_{i(c)t}, \]  

which now includes judge fixed effects \((\mu_j)\).

The electoral cycles results are reported in Table 7. First, Column I again includes the baseline effect with all the NCRP data. Column II reports the NCRP data results while subsampling only the ten states for which we have election cycles data. The effect is no longer significant, reflecting that the number of observations and number of clusters decreases by a factor of five. However, Column III shows the equivalent regression using self-collected data on the universe of sentences in these states, doubling the sample size. In this dataset there is a positive effect of Fox News, comparable in magnitude to Column I.

### Table 7: Fox News and Sentencing Decisions: Judge Fixed Effects and Electoral Cycles

<table>
<thead>
<tr>
<th>Sample</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nielsien share (Fox)</td>
<td>0.075***</td>
<td>0.037</td>
<td>0.061*</td>
<td>0.068*</td>
<td>0.077**</td>
</tr>
<tr>
<td></td>
<td>(0.028)</td>
<td>(0.057)</td>
<td>[0.0515]</td>
<td>(0.036)</td>
<td>(0.036)</td>
</tr>
<tr>
<td>Proximity-to-election</td>
<td>0.337***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.112)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proximity-to-election</td>
<td>-0.134***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>x Nielsien share</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.046)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Judge FE</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partial R-squared</td>
<td>0.029</td>
<td>0.033</td>
<td>0.094</td>
<td>0.047</td>
<td>0.05 &amp; 0.01</td>
</tr>
<tr>
<td>F-stat. of excl. inst.</td>
<td>33.1</td>
<td>27.8</td>
<td>76.9</td>
<td>41.1</td>
<td>24.8 &amp; 6.6</td>
</tr>
<tr>
<td>Observations</td>
<td>4,974,207</td>
<td>1,126,388</td>
<td>2,521,509</td>
<td>2,521,509</td>
<td>2,521,509</td>
</tr>
</tbody>
</table>

Notes: Columns I–II use baseline specification from Column III of Table 3.1. The following states are included in Columns II–V: Alabama, Arkansas, Colorado, Georgia, Kentucky, Minnesota, North Carolina, Pennsylvania, Tennessee, Virginia, and Washington. All columns include state-year FE. Because for the full sample of 10 states we have zero-values, in Columns III–V we use as the outcome inverse hyperbolic sin of sentence length \((\log(y_i + (y_i^2 + 1)^{1/2}))\), which is approximately equal to \(\log(2) + \log(y_i)\), and can be interpreted the same way as a standard logarithmic variable but without needing to fill in zero values (Burbidge, Magee and Robb 1988; Card and DellaVigna 2017). The following variables are used as controls in Columns III–V: age, age squared, and race dummies (Black, Hispanic, Asian, Native American, and other), dummy for recidivists and state-specific set of crime severity dummies. Standard errors clustered by county are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

The result holds even when adding judge fixed effects (Column IV), which measures within-judge, across-time variation. As judges generally reside within the same seat during the whole career, this specification is more restrictive than the specification with county fixed effects. Remarkably, even in this very restrictive and conservative specification our results are significant (p-value=0.057), and the magnitude of the estimate does not change much. Thinking in terms of mechanism, the within-judge result is important because it is consistent with conservative TV hav-

---

\(^{30}\)We did not use county fixed effects in Table 1 because there were few data points within state-year and within county over time in the NCRP data.
ing an incentive effect on incumbent judges, rather than working through the selection of different types of judges.

Finally, to get at the interaction between Fox News and the judicial election cycle, we produce estimates for the 2SLS System (6, 7, 8) in Table 7 Column V. The specification instruments for heterogeneity across judges in proximity to elections as an interaction term. The coefficient for Fox News remains positive and significant, while judges become harsher closer to reelection. However, we find a negative coefficient for the interaction: the effect of Fox News exposure on elections is marginally smaller as elections become more immediate.\(^{31}\)

An interpretation of this result is that electoral proximity and Fox exposure are substitutes (rather than complements) in their effects on sentencing. Both of these treatments work to politicize the judicial decision process, but their effects are additive rather than multiplicative. This notion could be useful for other partisan media researchers, who perhaps should consider the timing of associated elections.

### 6.3 Crime, Drugs, and Race in Fox News Language

If Fox News affects sentencing through the judicial election process, what content in the network is driving the effect on voters? To substantiate that Fox News affects judges through political messaging on crime, we applied text analysis to the news show transcripts to measure the volume of crime-related language over time. Specifically, we count the number of times the words crime, criminal, murder, or homicide are said on prime time shows during a month. To measure drug mentions, we count the number of times the words drug, drugs, marijuana, cocaine, crack, ecstasy, meth, PCP, or heroin were said. In addition, we computed separate counts for crime words and drug words including only sentences where there were any words referring to black race (black, african, or african-american).

For the regressions, we employ our most conservative specification with judge fixed effects from Column IV of Table 7. The second stage is:

\[
Y_{it} = \alpha_{it} + \rho_1 T_{ct} + \rho_2 T_{ct} \times L_t + L_t + \mu_i(c) + X_{it} \beta + \epsilon_{it},
\]

where \(L_t\) is the log of the count of words related to crime in Fox News transcripts (as just described). \(L_t\) is a nation-wide variable that varies by year-month (and therefore not absorbed by state-year fixed effects). The interaction \(T_{ct} \times L_t\) is the treatment effect of interest. Similar to the electoral-cycle specification, the two excluded instruments are the channel position \(Z_{ct}\) and the channel position interacted with the Fox News language measure, \(Z_{ct} \times L_t\). We assume that \(L_t\) can be considered exogenous to counties’ local crime rates. We include as a control the log of total number of sentences said in Fox News transcripts in that month.\(^{32}\)

\(^{31}\)We can’t estimate the specification in Column V separately for all 10 states because we would have 11 endogenous variables and 11 instruments and the first stages would not be strong enough. However, if we do a reduced form estimation, we do not find any heterogeneity in the interactions of electoral cycles and Fox News channel position by states. These results are available on request.

\(^{32}\)Results also hold if we control for the interactions of the log number of sentences with the state fixed effects (see
Table 8 presents the 2SLS results for Equation (9). In Column I we interact Fox News exposure with the log mentions of crime-related words during the month of the sentencing decision. Only the interaction is significant, and Fox News viewership loses significance, suggesting that the effect is concentrated in months with more crime-related language. In Column II, we reinforce our findings regarding the drug bias of the effect, by showing that Fox News channel’s interaction with the log mentions of drugs is significant. In Columns III and IV, we only use those words about crime or drugs that appear in the same sentence as black-race words, suggesting some racial bias of the Fox News’ crime messaging. Both appear to be significant. Finally, for robustness, in Columns V–VIII, we repeat the same specifications as in Columns I–IV with month fixed effects to control for within-year seasonality. However, all results remain virtually unchanged.  

7 Discussion and Concluding Remarks

This paper has shown evidence that conservative television media exposure has a causal effect on judge decision-making. When Fox News has higher viewership due to lower channel numbers, that makes judges harsher in their sentencing. This result adds to previous work showing that Fox News has an effect on voter attitudes (Martin and Yurukoglu, 2017); here we have established that it also has an influence on judges (in the high-stakes decision of how long to incarcerate a person).
Adding to work showing that judges respond to political incentives (Ash and MacLeod, 2017), we have established that politicized information (and not just appointment institutions) matter for judge decision-making.

In the heterogeneous effects section, we showed that the effect of Fox News is focused on black defendants and on drug-related crimes. The racialized discourse around crime indicated in Figure 2 appears to sway judges on the ground to increase disparities. This result adds to the large literature on racial discrimination in the U.S. criminal justice system (e.g., Fagan and Ash, 2017), and specifically in the context of the war on drugs (Banks, 2003). We establish a racial bias in the effect of conservative discourse on criminal justice decisions, and this is linked to drug crimes. As Blacks are disproportionately arrested for non-violent drug-related offenses, the effect could be driven by racial bias in media messaging. Alternatively, it could be that “tough-on-drugs” rather than “tough-on-crime” rhetoric matters in this setting. Future work could try to distinguish which types of rhetoric are more distinctive of Fox News.

As a step in this direction, we started to examine how Google search activity is related to Fox News viewership. In Figure 7, we visualize the reduced form at the DMA level for the effect of Fox News channel position on searches for “crime,” in Panel A, and on a racial slur, in Panel B (as done in Stephens-Davidowitz, 2014). These bincsatters are residualized on state fixed effects and controls for the channel positions of the other networks. While there is no effect for crime, there is a negative effect for the slur. That is, places with lower Fox channel position (and therefore higher Fox viewership) have more racism expressed in Google searches. The reduced form relationship is statistically significant with p-value $= 0.088$.

We see a Fox effect for elected, but not appointed, judges. This is consistent with an electoral mechanism, where Fox News affects judge decisions by shifting voter attitudes, rather than shifting the policy preferences of judges directly. Our findings suggest that judges are time-consistent in their desire to appeal to the voters and not trying to be tough-on-crime closer to the election.
date. These regressions work by subsampling states; however, so we cannot rule out selective differences in responsiveness across states.

Finally, the effect of Fox News on elected judges becomes weaker in the run-up to the election date. One interpretation of this result is that politicized information and politicized incentives are substitutes, rather than complements. As electoral pressures become stronger, media effects are reduced. Another possibility is that Fox News content becomes more election-focused, and less devoted to crime, in the run-up to elections. We will try to distinguish these explanations in future work.

More generally, follow-up work could use more sophisticated text-analysis methods to recover which ideas in Fox News are driving our effects. Using recent developments in high-dimensional instrumental-variables methods (e.g., Belloni et al., 2012; Ash, 2016), one could ask what features of cable news discourse have an impact on sentencing. The goal is a greater understanding of the political economy of media and criminal justice. The evidence produced by this research program will be useful to judges, policymakers, and the public.
References


Online Appendix

to

“Conservative News Media and Criminal Justice: Evidence from Exposure to Fox News Channel”
Online Appendix A  Data Appendix

Online Appendix A.1  Sentencing Data

We use two separate sources of data. The first one is National Corrections Reporting Program (ICPSR 36373). This dataset is restricted; however, one can apply for it with IRB and get access within a month. Other data comes from Dippel and Poyker (2019) and Poyker and Dippel (2019). Below we provide the description of the data and how it was obtained.

Sentencing data were collected separately from each state. 14 states were willing to share their data with us for free or at reasonable cost: Alabama, Arkansas, Georgia, Kentucky, Maryland, Minnesota, Mississippi, Nevada, North Carolina, Oregon, Pennsylvania, Tennessee, Texas, Virginia, and Washington.

We contacted each state with the following initial data request:

The data we are looking for has a court case (or 'sentencing event') as the unit of observation. In some states the data is organized by charge (with several charges making up the case or sentencing event) and that is equally fine. The key data that we need are:

- date, month and year of sentencing;
- type of crime;
- length of sentencing;
- type of sentencing (low-security, high security, etc);
- defendant's sex;
- defendant's race;
- court identifier;
- name of judge or judge identifier number;
- type of court that convicted (trial, appeal, etc);
- in what prison the person was sent;
- We do not seek any information that identifies defendants.

Sincerely, XXX

There were 10 states that (i) shared their sentencing data in digitized form and (ii) included the judge identifiers needed to estimate judge political cycles. The following reports for each state the office responsible for storing the data, as well as relevant contacts at the time we requested the data between late 2016 and late 2018. Some states had considerably longer processing times than others. These were typically due either to backlogs of data-technicians or to having to get our request vetted and signed off on by other individuals.

1. Alabama

- Initial contact with the Sentencing Commission at http://sentencingcommission.alacourt.gov/.
- After emailing sentencing.commission@alacourt.gov, Bennet Wright processed our request.

---

34We also obtained sentencing data from Arkansas, Maryland, Mississippi, Nevada, Oregon, and Texas, but these states’ data does not include judge identifiers.
• Time between data application and delivery: 16 months.

2. Colorado

• Initial contact with the Colorado Court Services Division, at https://www.courts.state.co.us/Administration/Division.
• Jessica Zender, the Court Programs Analyst at the Court Services Division processed our request.
• Time between data application and delivery: 1 month.

3. Georgia

• Initial contact with Department of Corrections at www.dcor.state.ga.us/Divisions/ExecutiveOperations/OPS/OpenRecords.
• After emailing open.records@gdc.ga.gov it was recommended we go through their ‘Media Inquiries’ under +1-478-992-5247, where Jamila Coleman coordinated our request with their data technicians.
• Time between data application and delivery: 3 months.

4. Kentucky

• We spoke on the phone to Cathy Schiflett at the Kentucky Courts Research and Statistics Department.
• She guided us to https://courts.ky.gov/Pages/default.aspx, where we had to select ‘Statistical Reports’ and then submit our data request.
• Daniel Sturtevant handled our request.
• Time between data application and delivery: 9 months.

5. Minnesota

• Initial contact with the Minnesota Sentencing Guidelines Commission at http://mn.gov/sentencing-guidelines/contact/contact-us.jsp.
• Email address: sentencing.guidelines@state.mn.us.
• Kathleen Madland was the Research Analyst who processed our request.
• Time between data application and delivery: 2 months.

6. North Carolina

• Then we were put in touch with the North Carolina Administrative Office of the Courts, where our data request was processed by the ‘Remote Public Access’ data technicians;
• Time between data application and delivery: 3 months.

7. Pennsylvania

Leigh Tinik processed our request.
Time between data application and delivery: 1 month.

8. Tennessee
- Initial contact with Tennessee’s Department of Corrections at www.tn.gov/correction/article/tdoc-prison-directory.
- Tanya Washington, the DOC’s Director of Decision Support (Research & Planning), processed our request.
- Time between data application and delivery: 6 months.

9. Virginia
- Initial contact was through a web form of the Virginia Criminal Sentencing Commission at www.vcsc.virginia.gov/.
- After being initially denied on the grounds that FOIA requests could only be processed for Virginia residents, we called +1-804-225-4398, and were eventually approved after speaking to the director Meredith Farrar-Owens.
- Time between data application and delivery: 3 months.

10. Washington
- Initial contact with the Department of Corrections at www.doc.wa.gov/aboutdoc/publicdisclosure.asp, where Duc Luu processed our request.
- We use essentially the same data as Berdejó and Yuchtman (2013).
- Time between data application and delivery: 2 weeks.

Online Appendix A.2 Judicial Biography Data

All data about judge electoral cycles was taken from ballotpedia.org. The site contains information about the judges of each circuit court (or equivalent) for each state. The individual page of each judge contains data for age and gender of a judge, the dates when she was appointed/elected, date of retirement (if already retired), name of governor by whom she was appointed (if appointed), and whom the judge replaced.

To collect the data research assistants started with the contemporary judges, collected their data, and proceeded with their predecessor judges. This procedure resulted in collecting information for approximately 80% of the judges mentioned in the sentencing data. For the states where the name of a judge was known we searched those judges individually on the sites of their courts and added them to the dataset.

Six of the states in this paper include judge names or identifiers in the sentencing data: Alabama, Georgia, Kentucky, North Carolina, Tennessee, and Washington. We coded up judge biographies, including when they are up for re-election. Where judges are identified by name, merging the judge biographies is straightforward. Where only judge identifiers are given, these identifiers still almost always include a variant of the judges’ initials. When they do not include initials, we match on entry and exit dates.
Online Appendix B  Additional Results

Online Appendix Figure 1: First Stage and Reduced Form: Graphical Results

Panel A. First Stage

Panel B. Reduced Form

Notes: Binscatter diagrams for the first stage (Panel A) and reduced form (Panel B) of the baseline specification in Table 1 without any controls.
Online Appendix Figure 2: OLS relation: CNN/MSNBC Viewership and Incarceration Length

Panel A (CNN)  
Panel B (MSNBC)

Notes: This Figure shows that higher CNN and MSNBC viewership are associated with higher sentencing crosssectionally, although the relationship is weaker than with Fox News.

Online Appendix Figure 3: First Stage: CNN/MSNBC Channel Positions and Viewership

Panel A (CNN)  
Panel B (MSNBC)

Notes: This Figure shows that we can get a first stage for CNN, but not for MSNBC.

Online Appendix Figure 4: Reduced Form: CNN/MSNBC Channel Positions and Incarceration Length

Panel A (CNN)  
Panel B (MSNBC)

Notes: This Figure shows that there are weak reduced form effects, which is negative for CNN (the same as Fox) but positive for MSNBC (the opposite of Fox).
Online Appendix Figure 5: First-Stage Coefficients (by State)
Online Appendix Figure 6: Permutation Test: Reduced Form Effect of Fox News Channel Position

Notes: Based on 500 placebo Fox News channel positions. All regressions use baseline specification from Column III of Table 5.1.
Online Appendix Table 1: Fox and Sentencing: Placebo with Lagged Sentencing

<table>
<thead>
<tr>
<th>Lag $Y_t$</th>
<th>Sample</th>
<th>Years</th>
<th>Channel position (Fox)</th>
<th>Nielsen share (Fox)</th>
<th>Partial R-squared</th>
<th>F-stat. of excl. inst.</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>I II III IV V VI VII VIII IX X XI XII</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I II III</td>
<td>All</td>
<td>1995-2005</td>
<td>RF 2SLS RF 2SLS RF 2SLS RF 2SLS RF 2SLS RF 2SLS RF 2SLS RF 2SLS RF 2SLS RF 2SLS</td>
<td>0.0050 (0.0006)</td>
<td>0.0117 (0.0006)</td>
<td>0.0101 (0.0004)</td>
<td>-0.2947 (0.0009)</td>
</tr>
<tr>
<td>t-10</td>
<td>All</td>
<td>1995-2000</td>
<td>RF 2SLS RF 2SLS RF 2SLS RF 2SLS RF 2SLS RF 2SLS RF 2SLS RF 2SLS RF 2SLS RF 2SLS</td>
<td>0.092 (0.0447)</td>
<td>0.037 (0.0662)</td>
<td>0.048 (0.0357)</td>
<td>0.001 (0.0590)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1995-2005</td>
<td>RF 2SLS RF 2SLS RF 2SLS RF 2SLS RF 2SLS RF 2SLS RF 2SLS RF 2SLS RF 2SLS RF 2SLS</td>
<td>13.8</td>
<td>5.4</td>
<td>17.1</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1995-2000</td>
<td>RF 2SLS RF 2SLS RF 2SLS RF 2SLS RF 2SLS RF 2SLS RF 2SLS RF 2SLS RF 2SLS RF 2SLS</td>
<td>3,039,291</td>
<td>3,039,291</td>
<td>1,906,782</td>
<td>1,906,782</td>
</tr>
</tbody>
</table>

Notes: All columns include state-year FEs. The dependent variable is the inverse hyperbolic sine of the sentencing length. All Columns use baseline specification from Column III of Table 5.1. Standard errors clustered by county are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

With state-year fixed effects $\alpha_{st}$ included, identification of our baseline estimates in Table 1 comes from within-state within-year variation. To check that this variation truly estimates the effect of changes in Fox News consumption, rather than within-year trends, in Columns I–VIII of Online Appendix Table 1, we shift sentencing decision ($Y_{i(c)t}$) to $t-10$ and $t-15$, always evaluated relative to a state-specific year fixed effect. Neither reduced form nor the second stage is significant throughout the table. We also find no effect of Fox channel position on lagged outcomes in the subsamples of states with appointed and elected judges (Columns IX–XII).
Online Appendix Table 2: Fox News and Sentencing Decisions: Subsample Analysis

<table>
<thead>
<tr>
<th>Sample</th>
<th>Dependent variable: Log sentencing length in months</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>OLS Nielsen share (Fox)</td>
<td>0.018*** (0.003)</td>
<td>0.010 (0.073)</td>
<td>0.033*** (0.010)</td>
</tr>
<tr>
<td>Partial R-squared</td>
<td>0.004</td>
<td>0.147</td>
<td></td>
</tr>
<tr>
<td>F-stat. of excl. instrument</td>
<td>3.5</td>
<td>38.8</td>
<td></td>
</tr>
<tr>
<td>F-stat. p-value</td>
<td>0.061</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>2,387,252</td>
<td>2,387,252</td>
<td>2,586,946</td>
</tr>
</tbody>
</table>

Notes: All columns include state-year FEs. The dependent variable is the inverse hyperbolic sine of the sentencing length. All Columns use baseline specification from Column III of Table 5.1. Standard errors clustered by county are in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1

Since our Fox News data come in two pieces (2005–2008 and 2010–2017), in Online Appendix Table 2 we also check whether our results hold on the subsamples. Both OLS and 2SLS hold for the 2010–2017 sample; however, the first stage on the subsample of 2005–2008 is weak (F-statistic equals 3.5) and the second-stage coefficient is insignificant (although OLS coefficient is positive and significant). Our results are not driven by a particular statistical artifact and hold if we omit any combination of 5 states.
Figure 7: Reduced Form Effect of Fox News and Placebo Fox News Channel Positions

Judge Gerald J Seibel, Minnesota
Filing Date Lines: March 1 2006, March 6 2012

Notes: This figure shows an example electoral cycles in our data. The example is from Washington, where judges are elected for four-year cycles. This data is from (Dippel and Poyker, 2019) and was originally collected from ballotpedia.org. In Minnesota, judges are elected for six-year cycles. Proximity on the vertical axis is defined on a 0,1 scale, where proximity equals 1 on the day of the general elections in early November. We trim the electoral cycles at the state-wide filing date, after which the electoral cycle effectively ends for the large majority of judges who have no challenger for their seat. The time between filing date and general election date is sandwiched between two vertical lines. The electoral cycle restarts with the general election date. An observation is a day in which a judge passed a sentence.
### Online Appendix Table 3: Robustness for Table 7

<table>
<thead>
<tr>
<th>Sample</th>
<th>Log sentencing length in months</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>w/o 0 sentences</td>
<td>Full w/o 0 sent.</td>
<td>τ&lt;sub&gt;j&lt;/sub&gt;=0 for VA</td>
<td>τ&lt;sub&gt;j&lt;/sub&gt;≠0 for VA</td>
<td></td>
</tr>
<tr>
<td>Virginia</td>
<td>Nielsen share (Fox)</td>
<td>0.058*</td>
<td>0.070**</td>
<td>0.080**</td>
<td>0.077**</td>
<td>0.081**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.030)</td>
<td>(0.034)</td>
<td>(0.035)</td>
<td>(0.036)</td>
<td>(0.035)</td>
</tr>
<tr>
<td></td>
<td>Proximity-to-election</td>
<td>0.245**</td>
<td></td>
<td></td>
<td>0.337***</td>
<td>0.266**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.115)</td>
<td></td>
<td></td>
<td>(0.112)</td>
<td>(0.121)</td>
</tr>
<tr>
<td></td>
<td>Proximity-to-election x Nielsen share</td>
<td>-0.099**</td>
<td></td>
<td></td>
<td>-0.134***</td>
<td>-0.107**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.048)</td>
<td></td>
<td></td>
<td>(0.046)</td>
<td>(0.051)</td>
</tr>
<tr>
<td>Judge FE</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>F-stat. of excl. inst.</td>
<td>63.4</td>
<td>33.0</td>
<td>21.3 &amp; 5.4</td>
<td>24.8 &amp; 6.6</td>
<td>21.5 &amp; 5.4</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>2,007,519</td>
<td>2,007,519</td>
<td>2,007,519</td>
<td>2,521,509</td>
<td>2,007,519</td>
<td></td>
</tr>
</tbody>
</table>

**Notes:** This Table contains robustness checks for Columns III–V of Table 7. Columns I–III estimate the same specification as in Columns III–V of Table 7 but with zero sentences being set as missing. Column IV estimates the same specification (with zeroes) as Column V of Table 7 but without zero sentences and with proximity to election τ<sub>j</sub> ≠ 0 for Virginia. Column V estimates the same specification as Column V of Table 7 but with proximity to election τ<sub>j</sub> ≠ 0 for Virginia. Standard errors clustered by county are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

In comparison to the NCRP sample that has only information of defendants that were sentenced to a prison term, in the universe of cases for 10 states we also observe cases with zero sentence length. These include defendants found not guilty as well as those who received a fine or probation. In Columns I to III of Online Appendix Table 3 we replicate Columns III–V of Table 7 on a subsample of cases with non-zero sentence length. Next, because judges in Virginia are appointed rather than elected like in the nine other states, in Column V of Table 7 we set τ<sub>j</sub> = 0 for all Virginia’s judges. In Columns IV and V of Online Appendix Table 3 we replicate the specification from Column V of Table 7 with τ<sub>j</sub> ≠ 0, such as τ<sub>j</sub> reflects proximity-to-reappointment. All the results hold.
Online Appendix Table 4: Robustness for Table 8: with State-Specific log of Total # of Sentences in Transcripts

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nielsen share (Fox)</td>
<td>0.040</td>
<td>0.041</td>
<td>0.054</td>
<td>0.053</td>
</tr>
<tr>
<td></td>
<td>(0.032)</td>
<td>(0.033)</td>
<td>(0.034)</td>
<td>(0.034)</td>
</tr>
<tr>
<td>Nielsen share x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log # crimes mentions</td>
<td>0.012***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>at month-year t</td>
<td>(0.0044)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log # drugs mentions</td>
<td>0.013***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>at month-year t</td>
<td>(0.0039)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log # crimes &amp; black mentions</td>
<td>0.027***</td>
<td></td>
<td></td>
<td>0.052*</td>
</tr>
<tr>
<td>at month-year t</td>
<td>(0.0103)</td>
<td></td>
<td></td>
<td>(0.0283)</td>
</tr>
<tr>
<td>Log # drugs &amp; black mentions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>at month-year t</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F-stat. of excl. inst.</td>
<td>76 &amp; 96</td>
<td>80 &amp; 94</td>
<td>60.3 &amp; 31</td>
<td>67 &amp; 7.4</td>
</tr>
<tr>
<td>Observations</td>
<td>2,521,068</td>
<td>2,521,068</td>
<td>2,521,068</td>
<td>2,521,068</td>
</tr>
</tbody>
</table>

Notes: This Table contains robustness checks for Table 8. All columns include state-year FEs. The dependent variable is the inverse hyperbolic sine of the sentencing length. All Columns use baseline specification from Column III of Table 5.1. All columns include state-year FEs and judge fixed effects. The following variables are used as controls: log number of X mentions, state-specific log of total number of sentences in transcripts, age, age squared, and race dummies (Black, Hispanic, Asian, Native American, and other), dummy for recidivists, and state-specific set of crime severity dummies. Standard errors clustered by county are in parentheses. *** \( p<0.01 \), ** \( p<0.05 \), * \( p<0.1 \).
Online Appendix Table 5: Fox News and Mentions of Various Types of Drugs

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>VI</th>
<th>VII</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nielsen share (Fox)</td>
<td>0.045</td>
<td>0.045</td>
<td>0.042</td>
<td>0.053</td>
<td>0.055</td>
<td>0.067*</td>
<td>0.044</td>
</tr>
<tr>
<td>(0.033)</td>
<td>(0.032)</td>
<td>(0.033)</td>
<td>(0.076)</td>
<td>(0.035)</td>
<td>(0.036)</td>
<td>(0.033)</td>
<td></td>
</tr>
<tr>
<td>Nielsen share x Log # mentions at month-year t</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marijuana</td>
<td>0.016***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.005)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cocaine</td>
<td></td>
<td>0.027**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.011)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crack</td>
<td></td>
<td></td>
<td>0.018***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.006)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ecstasy</td>
<td></td>
<td></td>
<td></td>
<td>0.530</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3.110)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Meth</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.027***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.010)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.097</td>
<td></td>
</tr>
<tr>
<td>(0.092)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heroin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.028**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.011)</td>
</tr>
</tbody>
</table>

Log # X mentions: X X X X X X X

F-stat. of excl. inst.: 29 & 38 25 & 35 30 & 34 26 & 1 38 & 22 8.4 & 31 36 & 21

Observations: 2,521,068 2,521,068 2,521,068 2,521,068 2,521,068 2,521,068 2,521,068

Notes: This Table contains heterogeneous effects of Fox News and different drug types. It uses specification from Column II of Table 8, but only counts number of mentions for a particular drug. All columns include state-year FE's and judge fixed effects. The following variables are used as controls: log number of X mentions, log of total number of sentences in transcripts, age, age squared, and race dummies (Black, Hispanic, Asian, Native American, and other), dummy for recidivists, and state-specific set of crime severity dummies. Standard errors clustered by county are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Online Appendix Table 6: Text Analysis Placebo: Lagged and Lead # of Mentions of Crimes in Fox News

<table>
<thead>
<tr>
<th>Lagged months from the sentence</th>
<th>Dependent variable: Log sentencing length in months</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Nielsen share (Fox)</td>
<td>0.040</td>
</tr>
<tr>
<td></td>
<td>(0.0325)</td>
</tr>
<tr>
<td>Nielsen share x</td>
<td>0.0123***</td>
</tr>
<tr>
<td>Log # crime mentions (lag)</td>
<td>(0.0044)</td>
</tr>
<tr>
<td>Log # X mentions</td>
<td>X</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Leads months from the sentence</th>
<th>Dependent variable: Log sentencing length in months</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Nielsen share (Fox)</td>
<td>0.040</td>
</tr>
<tr>
<td></td>
<td>(0.0325)</td>
</tr>
<tr>
<td>Nielsen share x</td>
<td>0.0123***</td>
</tr>
<tr>
<td>Log # crime mentions (lead)</td>
<td>(0.0044)</td>
</tr>
<tr>
<td>Log # X mentions</td>
<td>X</td>
</tr>
</tbody>
</table>

Notes: This Table contains specification from Column I of Table 8, but uses lags and leads of the log # of crime mentions. All columns include state-year FEs and judge fixed effects. The following variables are used as controls: log number of X mentions, log of total number of sentences in transcripts, age, age squared, and race dummies (Black, Hispanic, Asian, Native American, and other), dummy for recidivists, and state-specific set of crime severity dummies. Standard errors clustered by county are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.