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

Elliott Ash Michael Poyker

August 4, 2023

Abstract

Local 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. Fox News viewership also increases self-reported beliefs about the importance of drug crime as a social problem.

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): UT Austin, mikhail.poyker@austin.utexas.edu. We are grateful to Sergio Galletta and Greg Martin for help with understanding the datasets. We thank Christian Dippel, Anna Harvey, Mike Mueller-Smith, Daniel Nagin, Arnaud Philippe, Evan Rose, Yuan Tian, Ali Yurukoglu, conference participants at CELS, TWEC, and seminar participants at CMU Heinz, ETH Zürich, and Loyola Marymount University 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 (Arceneaux et al., 2016; Clinton and Enamorado, 2014). 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., Epstein et al., 2013; Posner, 2008). But recent empirical work has documented that judges do respond to non-legal influences, political and otherwise (Abrams et al., 2019, 2022; Ash and MacLeod, 2020; Ash et al., 2020; Berdejó and Chen, 2017; Berdejó and Yuchtman, 2013; Chen et al., 2016; Cohen and Yang, 2019; Dippel and Poyker, 2021). In addition, there is evidence suggesting that the judiciary has become more conservative over time (e.g., Bonica and Sen, 2021). 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 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 (Dippel and Poyker, 2023), paired with data on cable news viewership at the county level. The measure of conservative news exposure is Fox News viewership (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), channel-number variation can be used as an instrument for TV viewership across channels. Viewers tend to watch networks that have a lower position, and that position is exogenously assigned due to arbitrary historical factors. 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. 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).[1] Hence, Fox News influences the high-stakes policy outcome of increasing the length of incarceration of convicted criminals.

An immediate mechanism question is whether the effect is really going through judges, rather than other actors in the legal system. First, 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 within-state across-county variation, absorbing any state-level factors (legislative or otherwise). Second, it could be that police officers, after watching Fox News, might arrest more people or focus on more serious crimes, or third, prosecutors might become more aggressive in charging decisions. The evidence works against both of these channels. We do not find effects of Fox News on factors that can be affected by police and prosecutors but not judges: e.g., the number or types of charges that defendants face. Fourth and 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, an additional mechanism question is whether it works by changing judge preferences directly, or instead through the electoral process. In U.S. state judiciaries, the judges in almost half the courts are directly elected by voters (e.g., Ash and MacLeod, 2020; Dippel and Poyker, 2021). Fox News might influence voters to become more conservative on crime, as it does on other issues (e.g., Schroeder and Stone, 2015). The tough-on-crime attitudes in the electorate might then put electoral pressure on judges to be harsher on crime, perhaps amplified by partisan prosecutors (Krumholz, 2019; Okafor, 2021). 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 harshness only for elected judges, and not for appointed judges. These results are consistent with voter attitudes providing a pivotal mechanism for our effects. In addition, these results hold with judge fixed effects using changes in Fox exposure over time, suggesting that the effect is driven in part by retention pressures and not just due to selection of harsher judges. We then separately investigate whether Fox also affects selection into harsher judges, and find evidence in support of this channel. The selection effect is stronger in states with harsher electoral competition and partisan elections.

To better connect the dots between media messaging and sentencing, we apply text analysis to Fox News transcripts and measure mentions of crime issues on the network's

[1] We find no effect on the extensive margin; i.e., probability of being sent to prison. The results hold in Dippel and Poyker's (2023) newly collected sentencing dataset, which contains the universe of sentencing decisions for ten states with information on probation, more detailed case controls, and judge identifiers.

programs. We show that the effect of Fox News on sentencing is strongest in months where the network paid more attention to crime issues in its broadcasts. In a placebo analysis using transcripts from CNN or MSNBC, we find no effect on sentencing. Thus, it is not just Fox News driving sentencing harshness; it is Fox's crime-related messaging that has an effect.

We can further validate that the effect goes through shifts in voter attitudes toward drugs and crime, which in turn, affects judges facing re-election. Using Gallup survey data, we look at the effect of Fox News on respondents' opinions on the perceived importance of crime issues. We find that in places that watch more Fox News due to the channel position, survey respondents report that drug crime is a more important policy issue. Thus we can draw a causal chain from Fox News exposure to voter attitudes to the behaviors of elected judges.

Consistent with the survey findings that Fox viewership causes voters to worry about illegal drugs, we find in heterogeneity analysis that the effect on sentencing is largest for drug-related crimes. Further, we find that the effect is larger for black defendants than for White or Hispanic defendants. Following up with additional text analysis of Fox transcripts, we show that the Fox News effect is stronger when there are more mentions of illegal drugs, or when the word "Black" is mentioned together with crime or drug words. This supporting evidence suggests that some of the observed effects of Fox News comes through racialized crime messaging.

These results will be of interest to scholars in empirical political economy, and in particular for those who study how the mass media influence the justice system. One of the closest papers is Lim et al. (2015), who find that higher newspaper coverage increases sentence lengths imposed by elected judges. A second related paper by Philippe and Ouss (2018), using French data, find that defendants whose cases received more attention in the media receive longer sentences. We show complementary evidence for the ideological direction, rather than volume, of coverage making a difference. Partisan TV news influences elected judges, driven not by attention to specific cases but by an overall strengthening of voter concerns about crime and drugs.[2]

Further, our results add to the literature on racial disparities in the criminal justice system. We complement the cross-sectional evidence for a partisan gap in racial sentencing disparities (Cohen and Yang, 2019), as our estimates have a causal interpretation for

[2]There are three other related papers that are worth mentioning. Using data from Italy, Mastrorocco and Minale (2018) show that the introduction of digital TV, which covers crime less than local TV, reduced concerns about crime among the population. In Switzerland, Couttenier et al. (2023) show that increased news coverage of immigrant crimes increased voting in a "minaret ban" referendum. In the U.S., Mastrorocco and Ornaghi (2020) show that local media acquisitions by a national conglomerate reduce coverage of local crime, which reduces policing performance in those localities.

shifts in ideology. Thus we contribute to understanding of the causes of disproportionate incarceration rates among African-Americans.

From a policy standpoint, the findings are relevant to recent debates on how judges should be selected, retained, and compensated (Ash and MacLeod, 2020; Cameron et al., 2019; Epstein et al., 2013; Iaryczower et al., 2013; Lim, 2013; Mehmood, 2022). We provide causal evidence that elected judges are more responsive to shifts in the preferences of the electorate than are appointed judges (Kessler and Piehl, 1998). The findings have implications for democratic institutions and judicial independence more broadly.[3] Finally, we add to recent debates on polarization and media regulation (Allcott and Gentzkow, 2017; Boxell et al., 2017).

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. First, Figure 3 from Martin and Yurukoglu (2017), shows that Fox News tends to use politicized phrases associated with Republican politicians. Second, Figure 1 shows two additional pieces of evidence on this point. In Panel A, we see that for the years 2005–2008, Fox speakers mention crime more often than speakers on CNN and MSNBC.[4] In Panel B, we show in our data that places with higher Fox News viewership tend to impose longer criminal sentences.[5] 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.

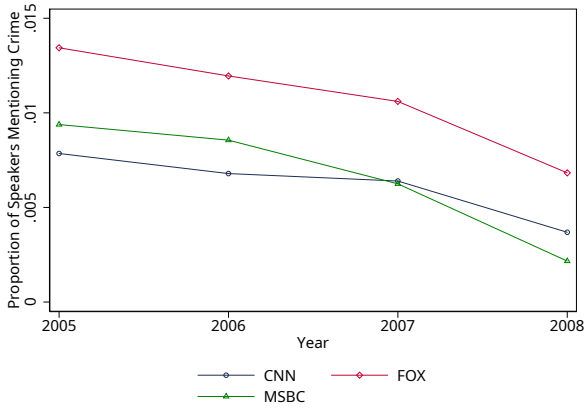
[3]In particular, we add theoretically relevant evidence on elected versus appointed judges, related to the literature on democratic accountability (e.g., Maskin and Tirole, 2004). We show evidence for a mechanism of changing voter preferences about crime, which influences judges through elections.

[4]These 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.

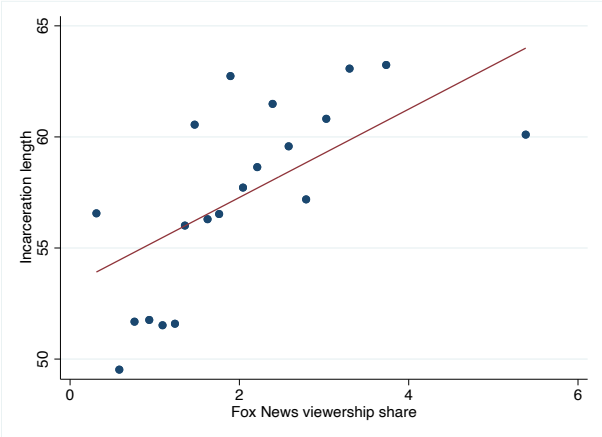
[5]We don't take a position on whether Fox News policy advocacy is "biased" away from some optimum; we are only speaking relative to the CNN and MSNBC reference point. In addition, we don't take a position on the motivations underlying this advocacy; it could be due to political motivations, due to trying to get more viewers, or for other reasons.

Figure 1: Fox News is Conservative; Viewership is Correlated with Sentencing Length

Panel A



Panel B



Notes: Illustrations for Fox News conservatism. Panel A is the number of references to crime per sentence spoken in cable news transcripts, obtained from Lexis. Panel B is a binscatter for the OLS correlation between incarceration length and Fox News viewership. Appendix Figures B.4 contains results for CNN and MSNBC.

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 neighbouring 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 similar 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. Word clouds illustrating the most crime-associated words for each of the three networks are reported in Appendix Figure B.1. Strikingly, one can see that at Fox (Panel A), 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, with CNN focusing on organized crime ("mobsters," "underworld") and MSNBC focusing on sensationalist murders.[6]

3 Data

3.1 Sentencing Data

We use two sources of data on sentencing, the NCRP dataset and the Dippel-Poyker dataset. The datasets are complementary for our analysis.[7] We discuss them each in turn.

NCRP Data. The first dataset on sentencing comes from the National Corrections Reporting Program (ICPSR 36373, hereafter NCRP). This is a standard dataset for the literature and it contains data on state prison admissions in the United States from 1991 to 2014. The data do not include relatively short jail (rather than prison) sentences (Neal and Rick, 2016). Federal (rather than state) prison sentences are not included.

The main advantage of NCRP is its breadth. It spans almost all U.S. states during the time period of our study (2005–2014). Thus, we can obtain a large sample size for the regressions. A disadvantage of NCRP is that it relies on voluntary submissions of information from each state (U.S. Bureau of Justice Statistics, 2021). That voluntary process may introduce bias, if for example partisan news influences the degree to which courts voluntarily submit information.

[6]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. E.g., Murder of Imette St. Guillen, in Fahim (2009).

[7]Here we do not use U.S. Sentencing Commission (USSC) data — another "off the shelf" dataset on criminal sentencing — because it covers only federal courts. We are measuring the effect of local media consumption on local sentencing decisions, so the local state courts are a better setting for our research question. Furthermore, we do not use Uniform Crime Reporting Program Data because it includes no data on an individual crime's severity and the defendant's prior criminal records.

Our main outcome variable for NCRP is the length of sentence imposed. We also include 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 fixed effects for the 180 offenses with the longest sentence length. We also have variables on criminal history (recidivism), education, military background, and demographic characteristics, including age at conviction, gender, and race (Asian, Black, Hispanic, Native American, White, and other).

Dippel-Poyker Data. The second dataset, from Dippel and Poyker (2023), is what we will refer to as the Dippel-Poyker dataset. This data was collected directly from state sentencing commissions. Information on how the data was collected is provided in Appendix A.

The main downside of the Dippel-Poyker dataset, relative to NCRP, is that it only available for ten states. These are Alabama, Colorado, Georgia, Kentucky, Minnesota, North Carolina, Pennsylvania, Tennessee, Virginia, and Washington. Thus, in regressions using the Dippel-Poyker dataset, we will have a smaller sample and less statistical power.

The Dippel-Poyker data has two key upsides. First, it is more comprehensive. Unlike NCRP, which is voluntary, the Dippel-Poyker data contain the universe of cases for these states. Appendix Figure A.1 illustrates how much the datasets diverge for the 10 respective states. For example, Virginia did not submit any data to NCRP. In other states, the completeness varies substantially over time and ranges from 4.5% in Georgia in 2010 to 72% in Minnesota in 2011.

Second, the case data contain more information. Besides the case-level information on sentencing decisions and offense type, we have to additional useful pieces of information. First, we have information on the defendants that were found not guilty or did not go to prison, including those that went on probation. Those defendants are dropped from NCRP. Second, we have information on the judges who decide the sentence.

The years covered by the Dippel-Poyker data vary from state to state, ranging from 1980 to 2017. In the regressions, we use the years for which we have overlap with Nielsen viewership data. Appendix Table A.1 provides details on data availability over time.[8]

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 re-

[8]We have as little as 4 years in Colorado and Georgia and as much as 11 years in Kentucky. In total, it is 78 years of data for 10 states.

quire many subjective decisions. Therefore, in our regressions we construct offense class fixed effects separately for each state. The data also contain information on recidivism and basic demographic characteristics, including age at conviction, gender, and race.

An advantage of having data on judges is that we can match them to information on judicial elections. To construct electoral cycles we use judges' biographies from ballotpedia.org linked to the sentencing dataset.[9]

Complementarity of the Datasets. Both the NCRP data and Dippel-Poyker data are crucial and play complementary roles in the analysis. NCRP gives the most geographical and time variation, providing the largest sample for the main analysis of how sentencing responds to Fox News exposure. In addition, it provides variation when comparing counties with appointed vs. elected judges. In particular, we use NCRP in our analysis in Kansas, where there is within-state variation in how judges are selected.

Meanwhile, the Dippel-Poyker data allows us to check whether our main results are biased by the selection of cases. This bias could come from voluntary submission of NCRP data, and the censoring of cases where no sentence was imposed. We also have better controls, including judge identifiers. Third, we can analyze the extensive margin of incarceration — the decision between getting a sentence or not. Fourth, we can use the judge data to better pinpoint the influence of judges as the mechanism.

3.2 Media Data

The data on channel positions and viewership 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,[10] we aggregate our treatment (viewership for Fox, CNN, and MSNBC) and instrument (channel positions) to the county level.[11]

The viewership data are for all shows on the networks, so they include a collection

[9] See details of the construction of electoral cycles and linking in Dippel and Poyker (2021).

[10] Their 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.

[11] We use weighted averages of channel position and viewership by zipcode, where the weighting is by zipcode population, to aggregate up to the county level. All results hold if we don't weight channel position by zipcode population (Table 1 Column VI).

of "news" shows (which claim to report straight news) and "pundit" shows (which have an acknowledged political viewpoint).[12] For the text illustrations (shown previously), we downloaded full-text transcripts of prime time shows on Fox News, CNN, and MSNBC from LexisNexis.

3.3 Other Data

We use individual-level survey data from the Gallup Poll Social Series for 2010–2016 to construct respondent's attitude toward importance of crime. We use population controls from American Community Survey, crime rates from Uniform Crime Reporting dataset, and 1996 Republican vote share from Martin and Yurukoglu (2017). For the Google Trends results, the data are downloaded from the Google Trends web site (trends.google.com) 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 differences in channel ordering across localities. Viewers tend to spend most of their TV time watching non-news content on popular channels (e.g., NBC, ESPN, Discovery), which tend to be low in the channel order for historical reasons (especially the broadcast networks NBC, CBS, ABC). When viewers surf away from these popular channels, for example during a commercial break, they are more likely to land on proximate channels with a low position. Hence, ordinal position proxies for the likelihood of incidental exposure for someone watching non-news content.

Meanwhile, Fox News is located at widely different positions, sometimes as low as 10 and sometimes as high as 100. This positioning of Fox News has a big influence on how many people watch it across different places.[13] The channel position was set at arbitrary numbers in different localities based on local conditions — in particular, what channels were being removed from those local cable systems in the year that Fox was added. Because of the decentralized market and local variation in existing channels, the position varied widely and exogenously (Martin and Yurukoglu, 2017).[14]

[12] More precisely, the 2005–2008 viewership data is for all shows, while the 2010–2017 viewership data are for prime time shows. This is what Nielsen made available to us through the data purchase.

[13] Note that the compliers for the channel position instrument are the individuals with low news interest and in particular weak preferences on cable news. Conservative viewers with a strong preference for Fox will watch it even if it is at a high channel position, while left-wing viewers will search for MSNBC even if Fox is on channel 1.

[14] It could be that the channel position is more likely to be endogenous in the later years of the sample (e.g.,

We adapt the zipcode-level analysis of Martin and Yurukoglu (2017) to the county level, as counties are the lowest-level geographical unit for the court-level sentencing data.[15] We specify an observation as a sentencing decision i that took place in county c of state s at year t. The first-stage estimating equation is:

$$T_{ct} = \alpha_{st} + \gamma Z_{ct} + \mathbf{X}_{i(c)t}\beta + \eta_{i(c)t}, \tag{1}$$

where viewership T_{ct} is the share of television-watching time spent on Fox News for county c at time t,[16] α_{st} includes state-year (interacted) fixed effects,[17] Z_{ct} is the channel number for Fox News, and $\eta_{i(c)t}$ is an error term. $\mathbf{X}_{i(c)t}$ includes other covariates describing demographics and cable system characteristics. From Martin and Yurukoglu (2017) we expect a negative and significant estimate for γ.

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:

$$Y_{i(c)t} = \alpha_{st} + \rho\widehat{T}_{ct} + \mathbf{X}_{i(c)t}\beta + \epsilon_{i(c)t}, \tag{2}$$

where $Y_{i(c)t}$ is log of sentencing length of decision i in court/county c during year t, $\epsilon_{i(c)t}$ is the error term, and other terms are the same as in Eq. (1). As treatment is at the county level, we cluster standard errors by county.[18]

We require instrument relevance. Appendix Figure B.2 Panel A shows graphical evidence of the first-stage effect. There is a clear downward trend, with higher channel numbers having lower viewership. In the regression tables below, we report the F-statistic of

after 2010). This is why we provide as a robustness check a cross-sectional specification for the instrument, where we use the 2005–2008 positions for all years.

[15] As our geographic cell is a court/county, we can employ the 2SLS identification strategy from Martin and Yurukoglu (2017), but we cannot follow differences-in-differences approach based on gradual introduction of Fox News by zipcodes as done in DellaVigna and Kaplan (2007). All counties in our sample have at least one zipcode with Fox News and zipcode level treatment can't be unambiguously mapped to our outcome of interest — court's sentencing decision. In a situation when a county has zipcodes without Fox News, county-level aggregation means that they are assigned a (weighted) average of Fox viewership and channel position from the zipcodes where Fox News is present. As a result, judges there are counted as "treated" while they are not "treated" in reality. Such measurement error will attenuate our coefficient, making it work against us finding a positive effect of Fox News exposure on sentencing length.

[16] We also show robustness to using the amount of time on Fox News, rather than the share. See Appendix Tables B.3 and B.4.

[17] This means identification comes from within-year within-state between-county variation in Fox News consumption and Fox News channel position. In particular, legislated trends (at the state level) will be absorbed by α_{st}.

[18] Our results hold if we cluster by state, DMA, or by location and year; see Table B.8. Note that this is different from Martin and Yurukoglu (2017), who cluster by cable system. We cannot use system clustering because system varies at the zipcode level, whereas our data are at the county level.

excluded instruments for each regression, and they are consistently greater than 10.[19]

Martin and Yurukoglu (2017) provide a lengthy discussion and set of checks for exogeneity of the instrument. They show that Fox News channel position does not correlate with the 1996 Republican presidential vote share, nor does it correlate with a set of demographic variables that are predictive of vote share. We replicate this balance test in our county-level dataset in Appendix Table B.1, along with several additional checks. The instrument is not related to past Republican vote shares, nor is it related to a set of covariates that are predictive of Republican vote shares. In our data, moreover, channel position is not related to past crime rates, past sentencing harshness, or public opinion about crimes and drugs. In the specification checks, we also demonstrate robustness of our coefficient estimates to inclusion of these factors as controls.

Under exogeneity, two-stage least squares procures consistent estimates for ρ if the instrument satisfies an exclusion restriction. That is, channel position affects sentencing only through its effect on Fox viewership. We feel this assumption is reasonable in our context. It could be that Fox News affects sentencing through a number of mechanisms — which we explore in detail below — but all of those are downstream from local viewership.

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

This subsection reports the main results for the effect of Fox News on sentencing harshness. First, we report the baseline results using the NCRP dataset. Second, we show consistent results using the Dippel-Poyker dataset.

Results with NCRP Data. The main results are presented in Table 1, starting with OLS estimates in Column I. As we already saw in Figure 1 Panel C, Fox News viewership is positively associated with the sentence length. Column II reports corresponding 2SLS estimates. The first stage is strong ($F = 33.2$), and the coefficient is positive and statistically

[19]Monotonicity requires that lower Fox News Channel would not decrease Fox News viewership, which 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 (see Appendix Figure B.5).

significant.[20] When adjusting for case characteristics (Column III), the 2SLS result is robust and becomes somewhat more precisely estimated.[21]

Notably, the coefficient on Fox for 2SLS is somewhat larger than for OLS, as in previous work using the channel position instrument (Martin and Yurukoglu, 2017). This is perhaps surprising, as one would expect that more conservative counties with harsher judges might also tend to have a pre-existing preference for Fox News. The most likely explanation, noted in the previous work, is measurement error. Nielsen viewership statistics are measured with substantial error due to changes in the composition of the Nielsen panel over time, particularly at low levels of geographic aggregation. Hence, 2SLS might procure a larger coefficient by reducing attenuation bias due to measurement error. It could also be that 2SLS identifies a local average effect among viewers who are less politically committed and therefore more sensitive to channel position.

The subsequent columns of Table 1 provide some key checks on the design. To check whether Fox viewership and channel position are confounded with the other cable news networks (CNN and MSNBC), Column IV uses an alternative specification for viewership which is normalized relative to the other networks,[22] while Column V includes CNN and MSNBC viewership as (non-excluded) controls. In each of these alternative specifications, the 2SLS estimates are positive, significant, and comparable in magnitude to Column III.[23] Next, Column VI shows robustness when the county-level averages across zipcodes for viewership and channel position are not weighted by zipcode population. Finally, in Column VII we show robustness to using only cross-sectional variation across counties in the instrument (the average channel position from 2005–2008).[24]

[20] Reduced form results are shown in Column I of Appendix Table B.2. Appendix Figure B.2 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. Residualized binscatter plot for the baseline specification in Column III is shown in Figure B.3.

[21] A specification with aggregated data (by court c and year t) would mix different offense types across jurisdictions. We found that aggregated specifications yield similar results to those in Columns I–II; i.e., without case controls.

[22] It is constructed as Fox viewership, minus the average of CNN and MSNBC viewership.

[23] Appendix Figure B.4 reports graphical evidence for effects of CNN and MSNBC on sentencing. We do not see the same effects for these networks.

[24] This check is important because cable systems might be able to endogenously adjust the Fox News channel position in response to local characteristics in later years. Hence, our results are not driven by changes in channel position over time. Additionally, it helps to address the concern about strategic behavior of television broadcast systems interested in acquiring TV stations strategically based on the Fox News channel position. For example, Martin and McCrain (2019) show that TV stations under Sinclair become more conservative and they reduce local news in favour of national focus. Using cross-sectional channel positions address this issue. Another reason we know that Sinclair is not driving the results is that Sinclair started to become more conservative after 2017, while our NCRP data ends in 2014 and in the Dippel-Poyker data only Tennessee has observations in 2017.

Table 1: Fox News and Sentencing Decisions: NCRP Data

	Dependent variable: Log sentencing length in months						
	I	II	III	IV	V	VI	VII
	OLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
Fox viewership	0.016***	0.054*	0.075***		0.075***	0.080***	0.066**
	(0.0044)	(0.029)	(0.028)		(0.029)	(0.024)	(0.026)
Fox - (CNN+MSNBC)/2				0.078***			
				(0.030)			
CNN viewership					-0.004		
					(0.009)		
MSNBC viewership					-0.009		
					(0.012)		
State-year FEs	✓	✓	✓	✓	✓	✓	✓
Demographic controls	✓	✓	✓	✓	✓	✓	✓
Case controls			✓	✓	✓	✓	✓
Unweighted						✓	
Cross-sectional instrument							✓
Partial R-squared		0.029	0.029	0.026	0.028	0.041	0.050
Anderson-Rubin p-value		0.0837	0.0134	0.0155	0.0169	0.0043	0.0455
F-stat. of excl. inst.		33.2	33.1	36.4	34.7	55.9	21.7
Observations	4,974,207	4,974,207	4,974,207	4,974,207	4,974,207	4,974,207	4,974,207

Notes: Regression results using NCRP data. The dependent variable is the log of the sentencing length. All columns include state-year FEs and demographic controls: age, age squared, race dummies (Black, Hispanic, Asian, Native American, and other), dummy for military background, recidivism, and education dummies. Case controls include a set of offense code dummies. In Column IV, viewership is computed relative to the average for CNN and MSNBC. In Column VI, viewership and channel position are unweighted averages across zipcodes (not weighted by zipcode population). Column VII uses only cross-sectional variation in the channel position (the average Fox News channel position from 2005–2008, applied across all years). Standard errors clustered by county are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

These estimates are economically significant. According to the estimate from Column II, increasing Fox News viewership by 1 standard deviation (1.3 percent of total TV-watching time) would increase average sentencing length by approximately 7 percentage points (about 3 months).[25] As a baseline, one could consider the evidence from Cohen and Yang (2019) that in federal courts, Republican-appointed judges give about 2-months-longer sentences on average than Democrat-appointed judges.

Results with Dippel-Poyker Data. We further demonstrate robustness of the effect of Fox News on sentencing harshness by providing a complementary analysis in the Dippel-Poyker dataset, containing the universe of cases in 10 U.S. states. These results are reported in Table 2. First, Column I replicates the baseline NCRP-data specification, where we only include sentencing decisions with a non-zero outcome (i.e., no probation). The

[25]The equivalent effect for OLS (Column I) is smaller: 37 days.

2SLS estimate is significant and comparable in magnitude to its counterpart in Table 1 Column III. This consistency is somewhat reassuring, as it suggests that the NCRP's limitations in terms of reporting are not significantly biasing the results when using that dataset.[26]

Table 2: Fox News and Sentencing Decisions: Dippel-Poyker Data

	Dependent variable:					
	Log sentencing length in months				D(Incarceration)	
Sample	No probations and acquittals		All			
	I	II	III	IV	V	VI
Fox viewership	0.058*	0.070**	0.061*	0.068*	-0.013	0.001
	(0.030)	(0.034)	(0.031)	(0.036)	(0.013)	(0.016)
Judge FE & tenure		✓		✓		✓
State-year FEs	✓	✓	✓	✓	✓	✓
Demographic controls	✓	✓	✓	✓	✓	✓
Case controls	✓	✓	✓	✓	✓	✓
Partial R-squared	0.067	0.030	0.094	0.046	0.094	0.046
Anderson-Rubin p-value	0.032	0.012	0.074	0.061	0.074	0.061
F-stat. of excl. inst.	63.4	33.0	76.9	41.1	76.9	41.1
Observations	2,007,519	2,007,519	2,521,509	2,521,509	2,521,509	2,521,509

Notes: Regression results using Dippel-Poyker data. Includes data from Alabama, Arkansas, Colorado, Georgia, Kentucky, Minnesota, North Carolina, Pennsylvania, Tennessee, Virginia, and Washington. The dependent variable in Columns I–II is the log of the sentencing length. The dependent variable in Columns III–IV is the inverse hyperbolic sine of the sentencing length. The dependent variable in Column V is a dummy equal to one if the defendant is sentenced to a term in prison and zero otherwise (i.e., probation or acquittal). All columns use the baseline specification from Column III of Table 1. Judge's tenure is a judge-specific linear trend in years. Offense codes, crime severity, and recidivism variables are included as fixed effects which may vary across states. Columns I–II include only observations with non-zero sentencing length, so they are comparable with NCRP regressions from Table 1. Columns III–VI include defendants that got zero sentence. Standard errors clustered by county are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

An advantage of the Dippel-Poyker dataset is that we have judge identifiers. We explore further the within-judge effects of Fox News exposure by adding judge fixed effects and identifying judge responses to changes in instrumented Fox viewership over time. The main result holds when adding judge fixed effects and judge's tenure on the court (Column II).[27] As a preview to our discussion of mechanism, the within-judge result is important because it suggests that conservative media has an incentive effect on incumbent

[26] We note that including judge tenure (years on court) as a control does not change estimates (to the fourth decimal place). Table B.4 shows robustness of our results to using number of minutes of Fox consumption as the endogenous variable.

[27] Most judges only appear in a single county or judicial district (lowest criminal court territorial division, constituting one or several counties). Among our ten states, only judges in North Carolina rotate between counties within a judicial district. Results are robust to dropping North Carolina from the sample. Table B.6 replicates specifications with judge fixed effects from Table 2 without North Carolina. The significance of the resulting coefficients and of the first stage decreases as we now not using variation from judges changes their counties during their tenure. Nevertheless, our main results for intensive margin effect hold.

judges, rather than working through the selection of different types of judges. We separately test, whether Fox News consumption also affects selection into election of harsher judges (and thus affecting the sentencing via judge fixed effects) in Section 6.3 below.

For Columns III–VI of Table 2, we add back to the sample the defendants that did not receive a jail sentence (now included with $Y_{i(c)t} = 0$).[28] Columns III–IV show that our previous results (that only included the intensive margin of sentencing length), are similar to a broader specification that has both extensive and intensive margin.

Finally, Columns V–VI show no evidence for an effect of Fox News on the probability of incarceration. Thus the effect exists on the intensive margin but not on the extensive margin. This difference is intuitive, as judges have less discretion at the extensive margin. Generally judges do not send people to prison when they should not. Instead, judges have more discretion over the length of the sentence, so that is where we see an effect.

5.2 Robustness and Sensitivity Checks

Now we describe an additional sequence of robustness and sensitivity checks.

Inclusion of additional controls. To address possible endogeneity of the instrument, Appendix Table B.7 reports a number of additional regressions controlling for variables that may correlate with channel position and conservatism on crime.[29] Column II shows robustness to adjusting for pre-existing political conservatism, as measured by 1996 Republican vote share. Results also hold when we control for county's population (Column III), the share of county's zipcodes' population that had access to Fox News (Columns IV), the county's lagged average sentence length (Column V), and the county's lagged crime rate (Column VI). The results also hold when we include all of these controls together in Column VII.

Alternative specification for viewership. In the previous, we use as the endogenous regressor the share of TV time spent on Fox News Channel. We prefer the share, rather

[28] Here, because we have zero-values (i.e., probations and acquittals) for sentence lengths, so rather than logs we use the inverse hyperbolic sine transformation, $\log(y_i + (y_i^2 + 1)^{1/2})$, 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 et al., 1988; Card and DellaVigna, 2020). For brevity, we often refer to the inverse hyperbolic sine of sentence length as log sentence length. Table B.5 shows that our results for Columns III–IV of Table 2 are also robust to usage of $\log(y_i + 1)$.

[29] We prefer not to add these variables in the baseline specification as these variables may be affected by Fox News and we don't want to introduce joint outcomes as controls. Hence, we also separately show absence of county-level correlation between Fox channel position and socio-economic, political, and crime variables in Table B.1.

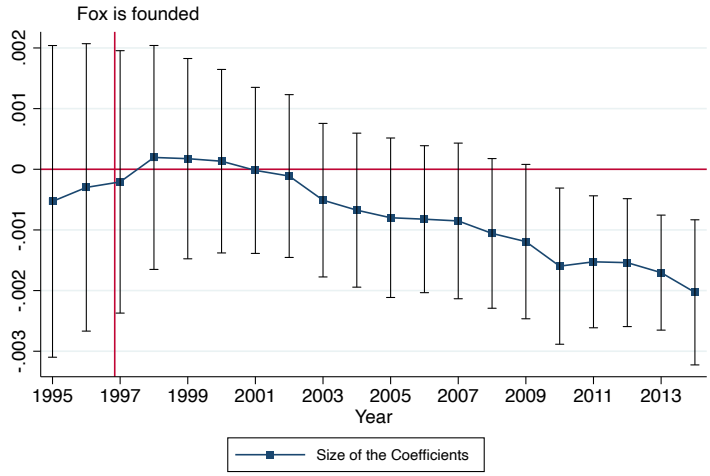
than time in minutes, as it adjusts for variation across counties in total TV watching time. Yet our results also hold if instead we use the number of minutes of Fox News watching as the endogenous dependent variable. The results corresponding to Table 1 are reported in Appendix Table B.3, while the results corresponding to Table 2 are reported in Appendix Table B.4.

Effect dynamics and placebo checks. We provide an additional placebo analysis by looking at the county-level effect of post-treatment Fox exposure on pre-treatment sentencing decisions. We estimate the reduced form year-by-year using the cross-sectional channel position (averaged from 2005–2008), starting even before the introduction of Fox News. Figure 2 plots the resulting coefficients. Reassuringly, there is a zero effect of Fox channel position on sentencing before Fox was introduced. The effect then gradually becomes more negative over time — that is, meaning a lower channel position (with higher viewership) starts to increase sentencing harshness. The effect is marginally significant (p-value = 0.12) by 2008, and significant at the 5% level by 2010. Overall, the graph suggests a cumulative effect of Fox exposure on sentencing harshness that increases over time. This analysis serves to check that Fox causes harsher sentencing, rather than places with already harsh sentencing tending to get a lower Fox channel position.

In the Appendix, we provide additional supporting regression results in this vein. When limiting lagging the outcome or focusing on early years, there are no effects of Fox on sentencing.[30] As a set of additional placebo exercises, Appendix Table B.2 shows 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 effect on sentencing decisions (Columns III–VI).

[30] Appendix Table B.10 estimates our main specification while lagging the sentencing decisions $Y_{i(c)t}$ to $t-10$ and $t-15$, always evaluated relative to a state-specific year fixed effect. Neither the reduced form nor the second stage is significant. As a similar placebo test, Appendix Table B.2 Column II shows that the Fox channel position does not have an effect on sentencing length when limiting the sample to the earliest years (2004 and earlier). Note that, here we can only show the reduced form because we don't have local viewership data before 2005. This null estimate for the early years is consistent with the argument from Martin and Yurukoglu (2017) that Fox News was less conservative in that earlier period. Meanwhile, the analogous reduced-form estimate for the later years is negative and significant, as expected (Column I).

Figure 2: Reduced-Form Effect of Fox News on Sentencing by Year



Notes: Reduced-form regressions for the effect on sentencing length of cross-sectional variation in Fox News channel position, interacted with year dummies. This figure can be viewed as a repeated cross section providing yearly coefficients. Error spikes give 90% confidence intervals. Regression uses NCRP data, with the same set of controls and fixed effects as the baseline specification from Table 1 Column III. Vertical line indicates establishment of Fox News in October 1996.

Permutation test. Next we provide an additional permutation test using the reduced form, to assess significance of the effect. We permute the Fox News channel positions with replacement, thereby comparing our true reduced-form estimates to the distribution of estimates obtained from regressing sentence length on fake Fox channel position 500 times. Appendix Figure B.6 shows the distribution of these estimates. The true coefficient has by far the largest magnitude.

5.3 Heterogeneous Effects by Defendant Characteristics

This section considers key dimensions for heterogeneous effects according to defendant characteristics. Here, we check whether media slant is disproportionately affecting minorities, women, or certain non-violent types of crime. Previous work has found heterogeneous responses by defendant race to pressures on the criminal-justice system, for example with electoral incentives (Kubik and Moran, 2003).

Our empirical approach is the same as before, where we adapt the 2SLS specification with an interaction term. The second stage is

$$Y_{i(c)t} = \alpha_{s(c)t} + \rho_1 T_{ct} + \rho_2 T_{ct} \times \mu_{i(c)} + \mu_{i(c)} + \mathbf{X}_{i(c)t}\beta + \epsilon_{i(c)t}, \tag{3}$$

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, ρ_1, plus the interaction effect with the defendant's characteristic that might be targeted by the media, ρ_2.

There are two endogenous variables. The first stage consists of

$$T_{ct} = \alpha_{s(c)t} + \gamma_1 Z_{ct} + \gamma_2 Z_{ct} \times \mu_{i(c)} + \mu_{i(c)} + \mathbf{X}_{i(c)t}\beta + \eta^1_{i(c)t} \tag{4}$$

$$T_{ct} \times \mu_{i(c)} = \alpha_{s(c)t} + \gamma_1 Z_{ct} + \gamma_2 Z_{ct} \times \mu_{i(c)} + \mu_{i(c)} + \mathbf{X}_{i(c)t}\beta + \eta^2_{i(c)t}, \tag{5}$$

where the terms are as above. The two excluded instruments are the channel position Z_{ct}, plus the channel position interacted with the defendant characteristic, $Z_{ct} \times \mu_{i(c)}$. For this heterogeneity analysis and the other regressions below that use an interaction term, we obtain equivalent results if instead of 2SLS we estimate the reduced form with the respective covariate interacted with the instrument.

Table 3: Fox News and Sentencing Decisions: Heterogeneity by Defendant Characteristics

	Dependent variable: Log sentencing length in months				
	I	II	III	IV	V
Characteristic	Black	Hispanic	Female	Drug-related crimes	DUI crimes
Fox viewership	0.006	0.078***	0.067**	0.014	0.075***
	(0.027)	(0.030)	(0.028)	(0.031)	(0.028)
Fox viewership Characteristic	0.168**	-0.040	0.055	0.218***	-0.001
	(0.073)	(0.070)	(0.036)	(0.074)	(0.076)
Characteristic	-0.328**	0.154	-0.265***	-	-
	(0.143)	(0.145)	(0.075)	-	-
Partial R-squared	0.013	0.031	0.023	0.017	0.029
F-stat. of excl. inst.	17.2	34.2	20.6	19.3	17.4
Observations	4,974,207	4,974,207	4,974,207	4,974,207	4,974,207

Notes: Regression results using NCRP data. Dependent variable is the log of the sentencing length. All columns use the baseline specification from Table 1 Column III. Characteristic coefficients in Columns IV and V are absorbed by crime type fixed effects. We report partial R^2 only for the non-interacted endogenous variable. Standard errors clustered by county are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

The heterogeneity estimates are reported in Table 3. First, Column I looks at heterogeneity for black defendants. The coefficient for ρ_1 becomes insignificant, while ρ_2 is large, suggesting that the effect of Fox News on sentencing is highly concentrated among black defendants.[31] We find no evidence for a disproportionate media effect toward Hispanics

[31]Notably, we find a negative coefficient on the black-defendant indicator, 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).

or toward female defendants (Columns II–III).

Motivated by previous work on how minorities are often disproportionately prosecuted for non-violent crimes, the second heterogeneity analysis is by crime type. Specifically, 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 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 and prosecutors. Subsection 6.2 provides evidence for a judicial elections channel. Subsection 6.3 studies whether in addition to the changes in judges' harshness within their term, there is also a selection into election harsher judges. Subsection 6.4 uses text analysis to pinpoint the effect of crime-related language in Fox News transcripts on sentencing. Subsection 6.5 provides validation of voter attitudes as an intermediate factor.

6.1 Mechanisms: Judges, Police, Prosecutors, or Crime Rates

Here we provide evidence that our results are not driven by the decisions of police, prosecutors, or potential offenders. To this end, Table 4 tests whether Fox News affected other sentencing-related outcomes. First, in Column I, we check whether Fox News consumption affects the number of offenses in each case, which is strongly influenced by police or prosecutors, but not typically influenced by the judge.[32] We find no effect, providing evidence against the non-judicial channels (policing and prosecutors). Second, in Column II, we find no effect of Fox News on the predicted sentence length based on charges, meaning that prosecutors do not appear to be bringing more serious charges due to media exposure.[33]

Next, we check another criminal-justice channel — the actions of potential criminal offenders. Fox News could affect criminality through changes in local welfare policies, for example, as documented in Ash and Galletta (2023). More specifically, if Fox News affected voter preferences and local welfare policies, that may affect the opportunity costs

[32] We do not observe the number of charges for Alabama, Virginia, and Washington.

[33] This result also holds if we include cases with zero-length sentence.

of crime. We find no effect on crime rates (Column III), however, nor (as just mentioned) on the severity of crimes as proxied by charges (Column II). Overall, there is no evidence of potential criminal choices being an important mechanism.[34]

Table 4: Fox News and Additional Criminal Justice Outcomes

	I	II	III
	Dependent variable:		
	# of charges	Predicted sentence length	Crime rates
Fox viewership	-0.930	-0.002	0.0001
	(0.959)	(0.004)	(0.0001)
Partial R-squared	0.106	0.030	0.007
F-stat. of excl. inst.	119.6	33.0	99.5
Observations	1,493,575	2,007,519	14,635

Notes: Observation in Columns I–II is a sentencing decision. In Column III, observation is a county-year. In Columns I–II we use data from Dippel and Poyker (2023) and for Column III we use number of crimes per 100,000 population from the Uniform Crime Reporting data (www.ucrdatatool.gov). Columns I–II use the baseline specification from Column III of Table 1. Column III only includes state-year fixed effects. Standard errors clustered by county are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

6.2 Heterogeneous Effects by Judicial Elections

To the extent that Fox News influences judges, that influence 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).[35] 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 5. 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

[34]Note, further, that these results also make it unlikely that the effects go through changes in victim reporting. Changes in victim reporting would likely be observed in changes in crime rates, and changes in the composition of charges.

[35]The list of states with appointed and elected judges can be found in Table 1 of Lim et al. (2015).

sign and insignificant. Column III reports the coefficient for the subsample of states with elected judges. The coefficient of interest is significant and only slightly larger than the baseline coefficient.[36] To further confirm a statistical difference between appointed-judge and elected-judge states, Column IV uses the full sample and estimates the effect with interaction terms for these state groups. The difference between coefficients is statistically significant.[37] As a placebo test for these results, we find no effect of Fox channel position on lagged outcomes in the subsamples of states with appointed or elected judges (Appendix Table B.10, Columns IX–XII).

Table 5: Fox News and Sentencing Decisions: Elected vs. Appointed Judges

	Dependent variable: Log sentencing length in months								
	I	II	III	IV	V	VI	VII	VIII	IX
Sample	All states				Kansas				
States/counties	All	Appointed	Elected	Both	All		Appointed	Elected	Both
	2SLS	2SLS	2SLS	2SLS	OLS	2SLS	OLS	OLS	OLS
Fox viewership	0.075***	-0.056	0.077***		0.007	0.248***	-0.014	0.022**	
	(0.028)	(0.076)	(0.028)		(0.008)	(0.086)	(0.014)	(0.009)	
Fox viewership x elected judges				0.086***					0.026**
				(0.028)					(0.012)
Fox viewership x appointed judges				-0.124					-0.017
				(0.087)					(0.014)
p-value of Δ coef.				[0.030]**			[0.029]**		[0.022]**
Partial R-squared	0.029	0.010	0.031	0.01 & 0.03		0.013			
F-stat. of excl. inst.	31.9	8.1	31.9	5 & 18		8.5			
Inst. F-stat. p-value	0.000	0.004	0.000	0.01 & 0.00		0.004			
Observations	4,974,207	1,025,716	3,948,491	4,974,207	9,115	9,115	4,369	4,746	9,115

Notes: Regression results using NCRP data. The dependent variable is the log of the sentencing length. All columns use the baseline specification from Column III of Table 1. Standard errors clustered by county are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

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).[38] 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).[39]

[36] Appendix Table B.9 clarifies that results using the Dippel-Poyker data are also driven by states with elected judges. In this dataset, we limit to the 9 out of 10 states (dropping Virginia) with elected judges (initial and/or re-election) and show similar results from Table 2.

[37] To test the significance of this interaction, we use Fox viewership interacted with a dummy for appointed judge county and a dummy for elected judge county as two endogenous variables, instrumented by Fox channel position also interacted with these two dummies. We then test for equality of coefficients.

[38] The list of counties with elected/appointed judges is available at Ballotpedia.org.

[39] 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).

For comparison, we show the OLS and 2SLS effects of Fox News for all of Kansas in Table 5 Columns V and VI, respectively. OLS is insignificant and 2SLS is positive and significant, but the first stage is weak with F-statistic of 8.5. Next, we split the state in samples of counties where judges are appointed (Column VII) and counties where judges are elected (Column VIII). While we do not have enough power in the first stage to report 2SLS estimates, we find positive and significant OLS effects for the subsample of counties with elected judges, but no effect for counties where they are appointed. Finally, in Column IX we show the coefficients are statistically different from each other (p-value = 0.022).

Overall, Fox seems to influence the sentencing of elected judges but not appointed judges. These results provide support for the effects being driven by an election channel, where voters become more tough on crime and thence influence judges through retention pressures. Moreover, the absence of an effect in states with appointed judges provides additional evidence against policing and prosecutor channels, which presumably would still be active in appointed-judge states.[40]

In the appendix, we report a supplementary analysis interacting the Fox effect with the judicial election cycle. We find that both Fox and proximity to election increase sentencing harshness, but there is no interaction effect between the two (Appendix Table B.11). The effects of media and the election cycle appear to be additive rather than multiplicative.

6.3 Selection Effects on Judges

In Section 5 we showed sentencing effects driven by changes in judges harshness during their term. Our most conservative specification with judge fixed effects in Table 2 allowed us to absorb any effects of *selection* of harsher judges and shows that incumbent judges become harsher in response to changes in Fox channel position. By absorbing the effect of selection, we are unable to separately identify it in that specification. Yet Fox News may also affect sentencing length via *selection* of judges. We test this hypothesis here.

To estimate the effect of Fox News on selection of judges, we use the Dippel-Poyker sample with judge IDs and regress individual sentence length on the set of demographic and crime controls and judge fixed effects.[41] We save the judge fixed effects and aggregate average severity of judges in county c year t. If a new judge come with higher harshness

[40]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.

[41]In particular we control for defendant's demographics, prior criminal history, set of state-specific crime severity fixed effects, and state-year fixed effects.

than the average (replacing a judge with lower harshness), then counties average judge harshness increases and we can detect *selection* into harsher judges.

We estimate the following second-stage specification:

$$\text{Judge harshness}_{ct} = \alpha_{s(c)t} + \theta_1 \hat{T}_{ct} + \theta_2 \hat{T}_{ct} \times \text{Election characteristic}_{s(c)} + \epsilon_{ct}, \tag{6}$$

where Judge harshness$_{ct}$ is average judge harshness in county c nested in state s in year t that we normalize to have zero mean and standard deviation of one. $\hat{T}_{ct}$ is Fox News consumption instrumented with the Fox channel position. The coefficient θ_2 estimates the additional Fox News consumption effect as a function of the competitiveness of judicial elections (Election characteristic$_{s(c)}$), measured at the state level. $\alpha_{s(c)t}$ are state-year fixed effects. We cluster standard errors at the county level.

While many states have elections, each state has its own cultural norms within the judicial profession regarding the electoral competition, with some states' incumbent judges never being challenged and some having tough competition. Thus we may expect some heterogeneity. We consider four determinants of competitiveness. First, variation in electoral rules (i.e., partisan, non-partisan elections) likely determines the competitiveness of elections. Second, variation in the length of electoral cycles might also determine election competitiveness because longer cycles generate a stronger incumbency advantage. Third, donor activity (from Bonica, 2016) will be a key determinant of a race's competitiveness. Fourth, differences in the baseline (state-wide) probability of having a challenger will determine the electoral pressures faced by incumbents (from Dippel and Poyker, 2021).

Column I of Table 6 recaps the estimate without the interaction with electoral competitiveness. According to the estimate, increasing Fox News viewership by 1 standard deviation (1.75 percent of total TV-watching time in this sample of counties) would increase average sentencing length by approximately 9 percent of a standard deviation.[42] We interact Fox News consumption (and its instrument in the first stage) with the state-level number of donors per judge-race in Column II.[43] The coefficient for Fox become insignificant; however, the estimate for the interaction is positive and significant.[44] This result suggest that the selection effect is larger in states with tougher electoral competition among circuit court judges. Column III shows results for the interaction with the average

[42]These results are not driven by correlations with judges' tenure or changes in judge turnover. In Table B.12, we adjust for tenure by controlling for starting year of the court's judges, and the results hold. Table B.13 shows that Fox News consumption does not affect judge turnover, in the sense of increasing/decreasing the share of new judges.

[43]Electoral contribution data for Colorado and Tennessee is not available.

[44]The average number of donors per race is 26. At this average, the effect of Fox News is positive: $-0.112 + 26 \times 0.007 = 0.07$.

share of contested races; while the sign of the coefficient is correct, the interaction is not statistically insignificant.[45] In Column IV we find no evidence of longer cycles mediating the effect of Fox News on selection of judges, in line with our evidence that electoral cycles do not exacerbate the effect of Fox on sentencing (Table B.11). Finally, in Column V we see that Fox affects selection judges more in states with partisan elections.

Table 6: Locations with Higher Consumption of Fox Also Elect Harsher Judges

	Dependent variable: Average judges' harshness				
	I	II	III	IV	V
Fox viewership	0.050**	-0.112	0.009	0.043	-0.010
	(0.022)	(0.090)	(0.030)	(0.043)	(0.023)
Fox viewership					
x # donors per judge-race		0.007*			
		(0.004)			
x Prob. electoral challenge			0.002		
			(0.002)		
x Length of electoral cycle				0.001	
				(0.007)	
x Partisan election					0.494*
					(0.271)
Anderson-Rubin p-value	0.021	0.009	0.069	0.051	0.013
F-stat. of excl. inst.	164	74.0	110.0	92.0	5.2
Observations	6,793	5,701	6,793	6,793	6,793

Notes: Observation in this table is county-year. The dependent variable is an average harshness of judges with mean zero and standard deviation of one. Individual harshness of judges is computed as fixed effects from the regression of judge fixed effects, case and demographic controls, and state-year fixed effects on sentencing using Dippel-Poyker sample of states with judge IDs. All columns include state-year fixed effects. Standard errors clustered by county are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

To further explore the selection effect, we focus on the year that a judge is first elected. Channel position does not change very often, so instead we take an approach closer to DellaVigna and Kaplan (2007) and construct a variable Fox at selection$_{j(c)}$, equal to the population-weighted share of zip-codes in county c where Fox News is available to cable subscribers, in the year when judge j was selected for service. We then compare judges to their peers on the same court who were selected when Fox was more or less available. As there are several judges per county and they are elected at different times. we can add county or even county-year fixed effects. Specifically, we estimate

$$Y_{i(c,j)t} = \alpha_{st} + \rho \text{Fox at selection}_{j(c)} + \mathbf{X}_{i(c)t}\beta + \mu_c + \mu_{ct} + \epsilon_{i(c,j)t}, \tag{7}$$

[45]The average average share of contested races is 26%. At this average, the effect of Fox News is also positive: $0.009 + 28 \times 0.002 = 0.065$.

where besides the availability treatment, we have state-year, county, and county-year FE's, as well as other controls as described previously.

Appendix Table B.14 reports the results on this additional selection analysis. We find that judges starting in years with higher Fox News availability also are harsher in terms of sentence length (Column I). In Column II, we add county-year fixed effects, which measures differences between judges working on the same court at the same time, but selected in years with different Fox availability. The coefficient is positive and larger in magnitude, yet not quite statistically significant (p-value=0.14).

6.4 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 explore whether Fox News affects judges through political messaging on crime, we applied a basic text analysis approach to the news show transcripts to measure the volume of crime-related language over time.[46] 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 (Phencyclidine)*, 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*).

This analysis uses the Dippel-Poyker dataset. NCRP only contains the date of prison admission, while Dippel-Poyker has the date of the judge's sentencing decision. Thus, we can count the number of crime mentions during the month before the sentencing decision was announced. The NCRP data would have significant measurement error because the month of prison admission is often later than the month of the sentencing decision.

We employ our most specification with judge fixed effects from Column IV of Table 2, to capture within-judge responses to messaging over time. As done with the heterogeneity analysis, we use an interaction term. The second stage is:

$$Y_{i(c)t} = \alpha_{s(c)t} + \rho_1 T_{ct} + \rho_2 T_{ct} \times L_t + \rho_3 L_t + \mu_{i(c)} + \mathbf{X}_{i(c)t}\beta + \epsilon_{i(c)t}, \tag{8}$$

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. The two excluded instruments are the channel position Z_{ct} and the channel

[46] Other work using a dictionary approach in economics includes Baker et al. (2016) and Enke (2020).

Table 7: Fox News and Sentencing Decisions: Impact of Crime and Drug Messaging

	Dependent variable: Log sentencing length in months							
	I	II	III	IV	V	VI	VII	VIII
Fox viewership	0.040	0.041	0.054	0.053	0.040	0.041	0.054	0.056
	(0.032)	(0.033)	(0.034)	(0.034)	(0.032)	(0.033)	(0.034)	(0.034)
Fox viewership x								
Log # crimes mentions	0.012***				0.012***			
at month-year t	(0.004)				(0.004)			
Log # drugs mentions		0.013***				0.013***		
at month-year t		(0.004)				(0.004)		
Log # crimes & black mentions			0.027***				0.028***	
at month-year t			(0.010)				(0.010)	
Log # drugs & black mentions				0.049*				0.041*
at month-year t				(0.028)				(0.024)
Month FEs					✓	✓	✓	✓
Log # X mentions	✓	✓	✓	✓	✓	✓	✓	✓
F-stat. of excl. inst.	42	43	36	36	42	43	36	37
Observations	2,521,068	2,521,068	2,521,068	2,521,068	2,521,068	2,521,068	2,521,068	2,521,068

Notes: The dependent variable is the log of the sentencing length. All columns use the specification from Column IV of Table 2 but add an interaction of Fox viewership with the log number of text pattern mentions, where the counted patterns are crime words, drug words, crime words co-occurring with black-race words, and drug-words co-occurring with black-race words (see text for details). We additionally control for the text mentions on their own (without the Fox interaction), as well as the log total number of sentences in transcripts. Columns V–VIII also include calendar month fixed effects. Standard errors clustered by county are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

position interacted with the Fox News language measure, $Z_{ct} \times L_t$. We include as a control the log of the total number of sentences said in Fox News transcripts in that month.[47]

Table 7 presents the 2SLS results for Equation (8). 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'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 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. All results are virtually unchanged.[48] As a placebo, in Tables B.17 and B.18 we replicate Table 7

[47]Results also hold if we control for the interactions of the log number of sentences with the state fixed effects (see Appendix Table B.15). Given that we are looking at changes in messaging over time, including judge fixed effects and looking at within-judge changes is the preferred specification. The results without judge fixed effects are of a positive sign but not statistically significant.

[48]In Appendix Table B.16 we report heterogeneous effects of Fox News and different drug types. We find

but use transcripts from CNN and MSNBC rather than Fox and Table B.19 shows that our results also hold when we separately control for the language of CNN and MSNBC. We find that interactions with their transcripts have no effect on sentencing, suggesting that the effect is driven by Fox's agenda rather than general media trends. Finally, we perform a permutation placebo on these estimates to check that the effect is specific to the month of the judge decision. For each of the four text measures, we replace the true value of the mention count with the value from a randomly selected month. We do this 100 times and plot the associated coefficients in Appendix Figure B.8. The true reduced form coefficients are at the edge of the distribution of estimates, suggesting that the effect is specific to the month when the judge makes the sentencing decision.

6.5 Effects on Public Opinion About Drug-Related Crimes

So far, we have demonstrated that the Fox News effect on sentencing goes through locations where judges are elected. Second, the effect is explained by Fox News explicitly talking about (drug-related) crimes. To corroborate that the effect goes through judges catering to changing voter opinions, we now show that Fox affects opinions about the importance of drug-related crime.

We use individual-level data from Gallup to construct two variables. The first is a dummy equal to one if a respondent thinks that "crime is the most important problem" and zero otherwise. The second dummy is a dummy equal to one if a respondents thinks that "drugs are the most important problem" and zero otherwise. The data have county identifiers and a number of demographic covariates used as controls.

We estimate the baseline specification (equation 1) on the sample of Gallup respondents in Table 8. Columns I–III report results for the importance of crime. We find no evidence that Fox News viewership affects people's attitudes toward the problem of crime generally. However, Columns IV–VI show that Fox exposure significantly increases the share of respondents who think that drugs are the most important problem. A one standard deviation in Fox News viewership increases this probability by 0.02-standard deviations, or 0.17 of the mean.

This result provides evidence for another link in the causal chain between Fox News viewership and judge sentencing harshness. Fox messaging on drug crime causes viewers to be more concerned about drug crime. Those attitudes work through the electoral process to pressure incumbent judges to be harsher on drug-related offenses.

that the interaction of most of the drugs with Fox News exposure appear to be positive and significant, with the strongest effect for cocaine, methamphetamine, and heroin.

Table 8: Fox News and Reported Attitudes toward Crime and Drugs

	Dependent variable:					
	D(Crime is the most important problem)			D(Drugs are the most important problem)		
	I	II	III	IV	V	VI
	OLS	OLS	2SLS	OLS	OLS	2SLS
Fox viewership	-0.001		-0.0018	0.001**		0.00113*
	(0.001)		(0.0013)	(0.000)		(0.0006)
Fox channel position		0.00003			-0.00002*	
		(0.00002)			(0.00001)	
F-stat. of excl. inst.			75.9			75.9
Observations	85,594	85,594	85,594	85,594	85,594	85,594

Notes: The dependent variable in Columns I–III is a dummy equal to one if the respondent thinks that crime is the most important problem and zero otherwise. Its mean is 0.015 and standard deviation is 0.12. The dependent variable in Columns IV–VI is a dummy equal to one if the respondent thinks that drugs is the most important problem and zero otherwise. Its mean is 0.01 and standard deviation is 0.08. Columns III and IV use Fox News channel position as the instrument for the Fox News viewership as specified in baseline Eq. (1). All columns include state-year FEs. The following variables are used as controls: gender, age, age squared, marital status, and race dummies. Standard errors clustered by county are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

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 pressures (Berdejó and Yuchtman, 2013; Dippel and Poyker, 2021), we have established that politicized information (and not just appointment institutions) matter for judge decision-making. Further, more so than the other work, we show that partisan media has concrete "downstream effects" on people's lives — that is, more time spent behind bars.

Yet partisan media exposure, by itself, is not sufficient to change judge sentencing decisions. The effect of Fox is observed among elected judges, but not appointed judges, consistent with election pressures as a pivotal mechanism. Rather than shifting the policy preferences of judges directly, Fox News affects judge decisions by shifting voter attitudes, which then influence judges through the electoral process. Hence, we have some evidence that is relevant to the broader policy issues of judicial independence and how that interacts

with democratic accountability. Elected judges may react more to extra-judicial pressures, such as public opinion, including on high-stakes decisions such as incarceration. This evidence will be useful to policymakers considering whether to reduce electoral pressures in state courts (Ash and MacLeod, 2020; Lim, 2013).

In the heterogeneity analysis, 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 Appendix Figure B.1 appears to sway judges on the ground to increase disparities. Thus, we establish a racial bias in the effect of conservative discourse on criminal justice, which is linked to drug crimes. This result adds to the large literature on racial discrimination in the U.S. criminal justice system, and specifically in the context of the war on drugs (e.g., Banks, 2003). 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 most persuasive to voters and judges.

As a step in this direction, we started to examine how Google search activity is related to Fox News viewership. In Appendix Figure B.9, 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). While there is no effect for crime, there is a negative effect for the slur (p-value $= 0.088$). That is, places with lower Fox channel position (and therefore higher Fox viewership) have more racism expressed in Google searches.

Future work could build on this evidence to explore further how media influences inter-group attitudes. Using more localized and fine-grained data on racial attitudes over time would be a promising first step. Beyond that, it could be fruitful to produce richer linguistic representations of the narratives used on partisan media outlets. Those narratives might influence both the decisions and the writings of judges.

References

Abrams, D., R. Galbiati, E. Henry, and A. Philippe (2019). Electoral sentencing cycles. Technical report.

Abrams, D., R. Galbiati, E. Henry, and A. Philippe (2022). When in Rome... on local norms and sentencing decisions. *Journal of the European Economic Association 20*(2), 700–738.

Allcott, H. and M. Gentzkow (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives 31*(2), 211–36.

Arceneaux, K., M. Johnson, R. Lindstädt, and R. J. Vander Wielen (2016). The influence of news media on political elites: Investigating strategic responsiveness in congress. *American Journal of Political Science 60*(1), 5–29.

Ash, E., D. L. Chen, and S. Naidu (2020). Ideas have consequences: The impact of economics on American justice. Technical report.

Ash, E. and S. Galletta (2023). How cable news reshaped local government. *American Economic Journal: Applied Economics*.

Ash, E. and W. B. MacLeod (2020). Reducing partisanship in judicial elections can improve judge quality: Evidence from us state appellate courts. *Center for Law & Economics Working Paper Series 5*.

Baker, S. R., N. Bloom, and S. J. Davis (2016). Measuring economic policy uncertainty. *Quarterly Journal of Economics 131*(4), 1593–1636.

Banks, R. R. (2003). Beyond profiling: Race, policing, and the drug war. *Stanford Law Review*, 571–603.

Berdejó, C. and D. L. Chen (2017). Electoral cycles among us courts of appeals judges. *The Journal of Law and Economics 60*(3), 479–496.

Berdejó, C. and N. Yuchtman (2013). Crime, punishment, and politics: an analysis of political cycles in criminal sentencing. *Review of Economics and Statistics 95*(3), 741–756.

Bonica, A. (2016). Database on ideology, money in politics, and elections (dime). public version 2.0 [computer file].

Bonica, A. and M. Sen (2021). Estimating judicial ideology. *Journal of Economic Perspectives 35*(1), 97–118.

Boxell, L., M. Gentzkow, and J. M. Shapiro (2017). Greater internet use is not associated with faster growth in political polarization among us demographic groups. *Proceedings of the National Academy of Sciences*, 201706588.

Burbidge, J. B., L. Magee, and A. L. Robb (1988). Alternative transformations to handle extreme values of the dependent variable. *Journal of the American Statistical Association 83*(401), 123–127.

Cameron, C. M., J. P. Kastellec, and L. A. Mattioli (2019). Presidential selection of supreme court nominees: The characteristics approach. *Quarterly Journal of Political Science 14*(4), 439–474.

Card, D. and S. DellaVigna (2020). What do editors maximize? evidence from four economics journals. *Review of Economics and Statistics 102*(1), 195–217.

Chen, D. L., T. J. Moskowitz, and K. Shue (2016). Decision making under the gambler's fallacy: Evidence from asylum judges, loan officers, and baseball umpires. *The Quarterly Journal of Economics 131*(3), 1181–1242.

Clinton, J. D. and T. Enamorado (2014). The national news media's effect on Congress: How Fox News affected elites in Congress. *The Journal of Politics 76*(4), 928–943.

Cohen, A. and C. S. Yang (2019). Judicial politics and sentencing decisions. *American Economic Journal: Economic Policy 11*(1), 160–91.

Couttenier, M., S. Hatte, M. Thoenig, and S. Vlachos (2023). The logic of fear-populism and media coverage of immigrant crimes. *Review of Economics and Statistics*.

DellaVigna, S. and E. Kaplan (2007). The Fox News effect: Media bias and voting. *The Quarterly Journal of Economics 122*(3), 1187–1234.

Dippel, C. and M. Poyker (2021). Rules versus norms: How formal and informal institutions shape judicial sentencing cycles. *Journal of Comparative Economics 49*(3), 645–659.

Dippel, C. and M. Poyker (2023). Do private prisons affect criminal sentencing? *Journal of Law and Economics*.

Enke, B. (2020). Moral values and voting. *Journal of Political Economy 128*(10), 3679–3729.

Epstein, L., W. M. Landes, and R. A. Posner (2013). *The behavior of federal judges: a theoretical and empirical study of rational choice*. Harvard University Press.

Fahim, K. (2009). In bouncer's murder trial, victim's friend recalls their last night. *New York Times, url: www.nytimes.com/2009/05/12/nyregion/12littlejohn.html*.

Gordon, S. and G. Huber (2007). The effect of electoral competitiveness on incumbent behavior. *Quarterly Journal of Political Science 2*(2), 107–138.

Iaryczower, M., G. Lewis, and M. Shum (2013). To elect or to appoint? bias, information, and responsiveness of bureaucrats and politicians. *Journal of Public Economics 97*, 230–244.

Kessler, D. P. and A. M. Piehl (1998). The role of discretion in the criminal justice system. *Journal of Law, Economics, and Organization 14*(2), 256–256.

Krumholz, S. (2019). The effect of district attorneys on local criminal justice outcomes. *Available at SSRN 3243162*.

Kubik, J. D. and J. R. Moran (2003). Lethal elections: Gubernatorial politics and the timing of executions. *The Journal of Law and Economics 46*(1), 1–25.

Lim, C. S. (2013). Preferences and incentives of appointed and elected public officials: Evidence from state trial court judges. *The American Economic Review 103*(4), 1360–1397.

Lim, C. S., J. M. Snyder Jr, and D. Strömberg (2015). The judge, the politician, and the press: newspaper coverage and criminal sentencing across electoral systems. *American Economic Journal: Applied Economics 7*(4), 103–135.

Martin, G. J. and J. McCrain (2019). Local news and national politics. *American Political Science Review 113*(2), 372–384.

Martin, G. J. and A. Yurukoglu (2017). Bias in cable news: Persuasion and polarization. *American Economic Review 107*(9), 2565–99.

Maskin, E. and J. Tirole (2004). The politician and the judge: Accountability in government. *American Economic Review 94*(4), 1034–1054.

Mastrorocco, N. and L. Minale (2018). News media and crime perceptions: Evidence from a natural experiment. *Journal of Public Economics 165*, 230–255.

Mastrorocco, N. and A. Ornaghi (2020). *Who watches the watchmen? Local news and police behavior in the United States*. University of Warwick, Department of Economics.

Mehmood, S. (2022). The impact of presidential appointment of judges: Montesquieu or the federalists? *American Economic Journal: Applied Economics*.

Mikolov, T., I. Sutskever, K. Chen, G. S. Corrado, and J. Dean (2013). Distributed representations of words and phrases and their compositionality. *Advances in neural information processing systems*, 3111–3119.

Neal, D. and A. Rick (2016). The prison boom and sentencing policy. *The Journal of Legal Studies 45*(1), 1–41.

Okafor, C. O. (2021). Prosecutor politics: The impact of election cycles on criminal sentencing in the era of rising incarceration. *arXiv preprint arXiv:2110.09169*.

Park, K. H. (2017). The impact of judicial elections in the sentencing of black crime. *Journal of Human Resources 52*(4), 998–1031.

Philippe, A. and A. Ouss (2018). "no hatred or malice, fear or affection": Media and sentencing. *Journal of Political Economy 126*(5), 2134–2178.

Posner, R. (2008). *How Judges Think*. Harvard University Press.

Schroeder, E. and D. F. Stone (2015). Fox News and political knowledge. *Journal of Public Economics 126*, 52–63.

Stephens-Davidowitz, S. (2014). The cost of racial animus on a black candidate: Evidence using google search data. *Journal of Public Economics 118*, 26–40.

U.S. Bureau of Justice Statistics (2021). National corrections reporting program, 1991-2017. retrievable from https://doi.org/10.3886/ICPSR37951.v1.

Online Appendix

to

"Conservative News Media and Criminal Justice:

Evidence from Exposure to Fox News Channel"

A Data Appendix

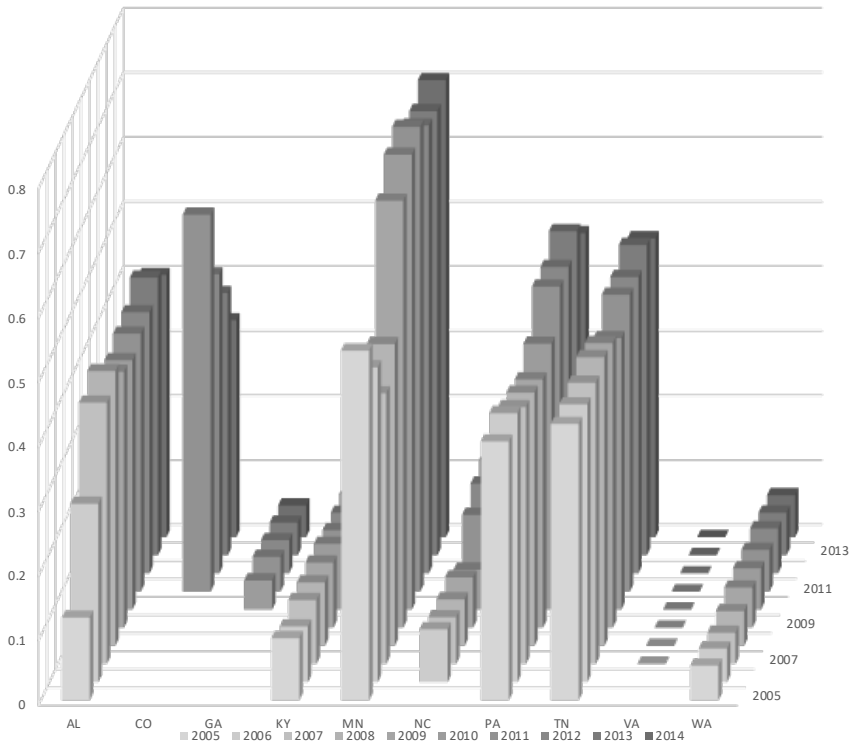
A.1 Sentencing Data

Table A.1: Range of Years with the Available Data on Sentencing in 10 States

#	State	Years y-start	y-end	#years	#years overlap with Nielsen data
1	AL	2002	2016	14	9
2	CO	2010	2016	6	4
3	GA	2010	2016	6	4
4	KY	2005	2018	13	11
5	MN	1991	2014	23	7
6	NC	2006	2016	10	8
7	PA	1991	2016	25	9
8	TN	1980	2017	37	10
9	VA	2006	2016	10	8
10	WA	2004	2015	11	8

Notes: Data for Colorado in 2010 is only available for the second half of the year.

Figure A.1: Comparison of NCRP/NJRP Data and with Full Data, 2005–2014



Notes: This figure shows the share of observations appearing in the NCRP data relative to the universe of the sentencing decision in that state-year. NCRP data ends in 2014. We show the data from 2005 just for the sake of readability of the graph, in order not to overwhelm the figure.

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 (2021) and Dippel and Poyker (2023). Below we provide the description of the data and how it was obtained. Table A.1 and Figure A.1 provide statistics comparing the datasets.

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.[49] The following reports show 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.
   - 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.

[49] We also obtained sentencing data from Arkansas, Maryland, Mississippi, Nevada, Oregon, and Texas, but these states' data does not include judge identifiers.

- 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

- Initial contact though www.ncdoj.gov/Top-Issues/Public-Integrity/Open-Government/Understanding-Public-Records.aspx.
- 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

- In Pennsylvania, sentencing data can be requested from the Sentencing Commission at http://pcs.la.psu.edu/data/request-and-obtain-data-reports-and-data-sentencing/data-sets.
- 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.

A.2 Judicial Biography Data

All data about judge elections is 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.

Ten of the states in this paper include judge names or identifiers in the sentencing data: Alabama, Colorado, Georgia, Kentucky, Minnesota, North Carolina, Pennsylvania, Tennessee, Virginia, 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.

B Additional Results

Figure B.1: Crime Discourse in U.S. Cable News Channels

Panel A. Most Similar Words to "Crime": Fox News

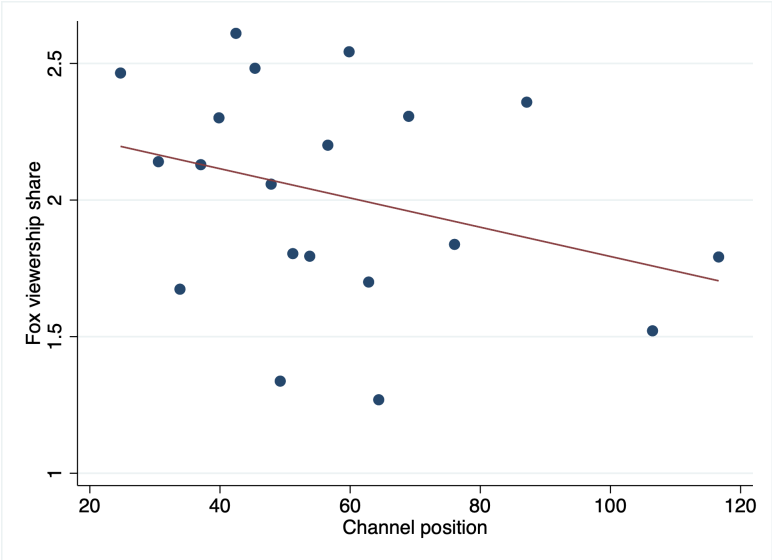
Panel B. Most Similar Words to "Crime": CNN

Panel C. Most Similar Words to "Crime": MSNBC

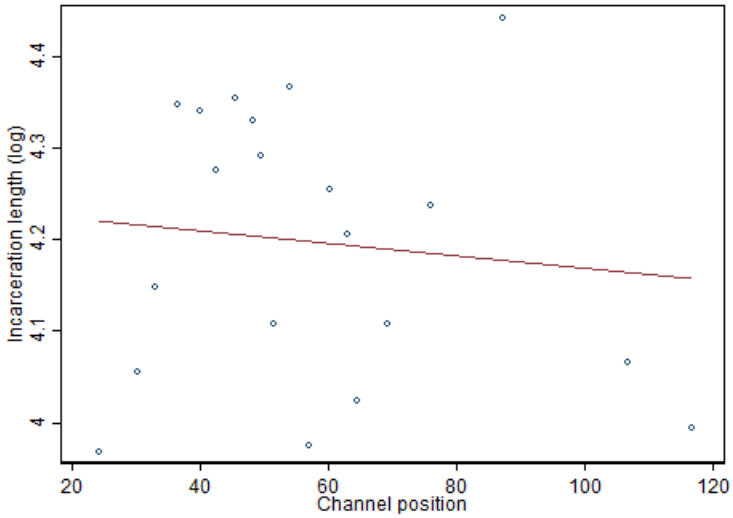
Notes: Most closely related terms to "crime," in Fox, CNN, and MSNBC, respectively. Similarities computed from word2vec models trained separately on the transcript corpora for each network. Larger words mean the word has higher similarity for the indicated network and lower similarity for the other two networks.

Figure B.2: First Stage and Reduced Form: Graphical Results

Panel A. First Stage

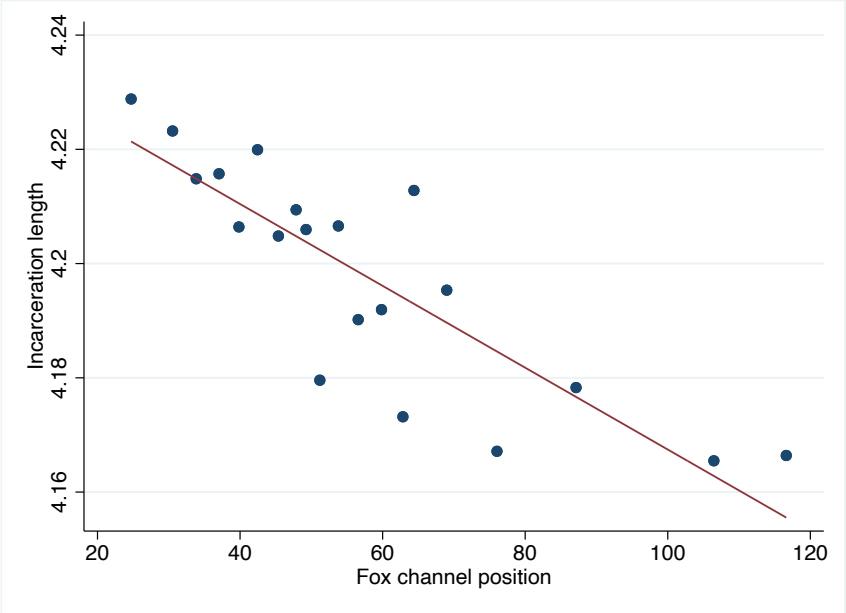


Panel B. Reduced Form



Notes: Binscatter diagrams for the first stage (Panel A) and reduced form (Panel B) of the baseline specification in Table 1 without any controls and fixed effects.

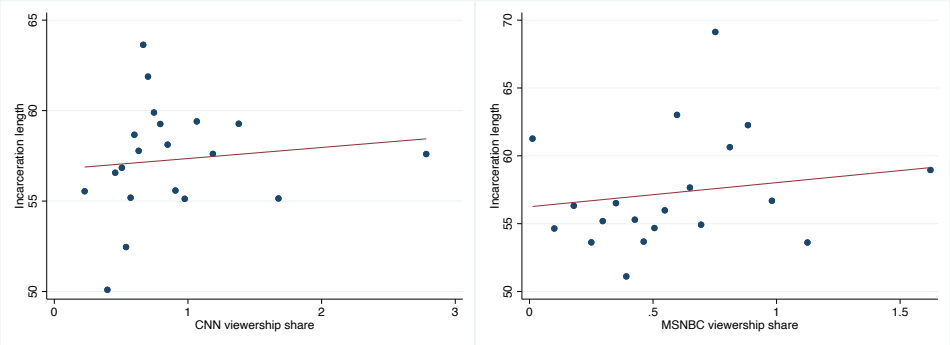
Figure B.3: Residual Binscatter Plot for Baseline Specification in Column III of Table 1



Notes: Residual Binscatter Plot for Baseline Specification in Column III of Table 1.

Figure B.4: OLS relation: CNN/MSNBC Viewership and Incarceration Length

Panel A (CNN) Panel B (MSNBC)



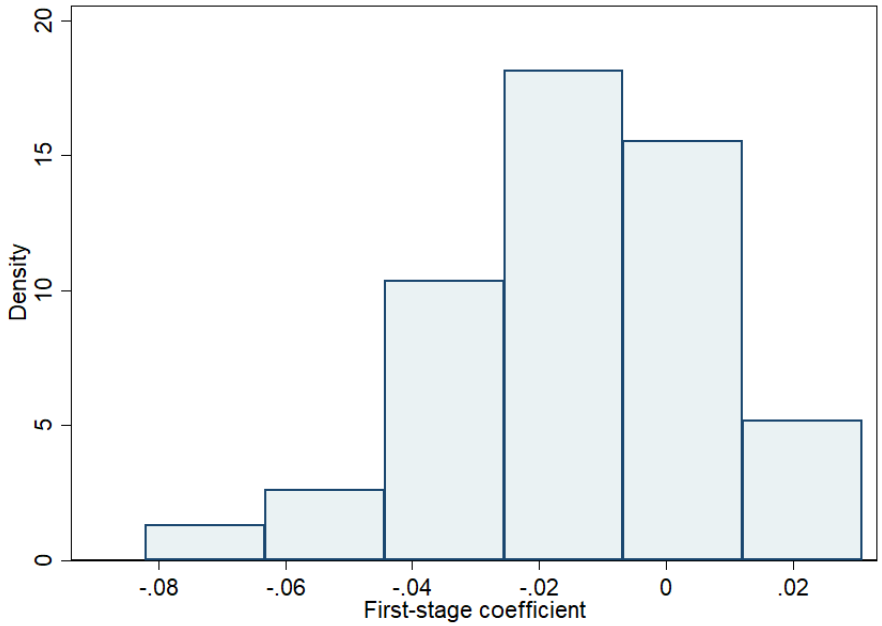
Notes: This Figure shows that higher CNN and MSNBC viewership are associated with higher sentencing cross-sectionally, although the relationship is weaker than with Fox News.

Table B.1: Balance Tests

	I	II	III
	Coefficient	S.E.	P-value
Socio-Demographic Controls (2010):			
Population	0.244	(0.146)	[0.101]
Poverty	-0.013	(0.009)	[0.189]
Urban/rural	0.030	(0.018)	[0.114]
Share nonwhite	0.155	(0.385)	[0.690]
Dom. migration	0.356	(0.320)	[0.271]
No high school	0.046	(0.139)	[0.739]
Median income	0.011	(0.007)	[0.101]
Republican Vote Shares:			
Presidential election Republican vote share, 1996	-0.108	(0.192)	[0.576]
Pres. el. Rep. vote share (predicted from demog), 1996	0.014	(0.110)	[0.903]
Crime & Sentencing Controls:			
Crime rates per 100k pop., 1991-2004	-0.0002	(0.0002)	[0.276]
Crime rates per 100k pop., 1996	-0.00002	(0.0001)	[0.889]
Avg. sentencing length, 1991-2004	0.0010	(0.002)	[0.653]
Avg. sentencing length, 1996	-0.0012	(0.005)	[0.804]
Perceptions about Crime and Drugs:			
D(Crime is the most important problem), 2000	-0.0003	(0.001)	[0.605]
D(Drugs are the most important problem), 2000	0.0001	(0.001)	[0.880]

Notes: Each observation is a county. Column I contains coefficient of the bivariate regression of Fox News channel position in 2005 on various outcomes. All regressions include state fixed effects. Column II reports standard errors clustered on county level. Column III reports p-values. *** p<0.01, ** p<0.05, * p<0.1

Figure B.5: First-Stage Coefficients (by State)



Notes: This figure plots first-stage coefficients estimated separately for each state.

Table B.2: Channel Positions and Sentencing: Reduced Form and Placebo Tests

	Dependent variable: Log sentencing length in months					
	I	II	III	IV	V	VI
Channel name	FOX News		Golf	Playboy	Trinity BN	A&E (crime ch.)
Sample	Baseline, [2005;2014]	t<2005				
Channel position	-0.0009**	-0.0006	0.0001	-0.00002	-0.0002	-0.0004
	(0.0003)	(0.001)	(0.0002)	(0.0001)	(0.0001)	(0.001)
R-squared	0.393	0.461	0.460	0.493	0.461	0.461
Observations	4,974,207	3,003,437	2,946,300	1,809,236	3,002,924	3,003,437

Notes: Regression results using NCRP data. The dependent variable is the log of the sentencing length. All columns use baseline specification from Column III of Table 1. The number of observations varies in Columns III–VI because some counties do not have these channels in some years. Standard errors clustered by county are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table B.3: Robustness for Table 1: Minutes of Watching Fox News as Endogenous Variable

	Dependent variable: Log sentencing length in months						
	I	II	III	IV	V	VI	VII
	OLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
Fox viewership (minutes)	0.036**	0.112*	0.155***		0.163***	0.150***	0.196**
	(0.015)	(0.059)	(0.056)		(0.063)	(0.046)	(0.079)
Fox - (CNN+MSNBC)/2				0.166***			
				(0.062)			
CNN viewership					-0.056		
					(0.045)		
MSNBC viewership					-0.108*		
					(0.059)		
State-year FEs	✓	✓	✓	✓	✓	✓	✓
Demographic controls	✓	✓	✓	✓	✓	✓	✓
Case controls			✓	✓	✓	✓	✓
Unweighted						✓	
Cross-sectional instrument							✓
Partial R-squared		0.063	0.062	0.061	0.058	0.113	0.052
Anderson-Rubin p-value		0.0837	0.0134	0.0155	0.018	0.0043	0.0455
F-stat. of excl. inst.		27.919	27.819	35.158	36.802	66.225	20.683
Observations	4,974,207	4,974,207	4,974,207	4,974,207	4,974,207	4,974,207	4,974,207

Notes: This table replicates Table 1 but uses ratings (proportional to number of minutes of Fox News watched) instead of share of viewership. Standard errors clustered by county are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table B.4: Robustness for Table 2: Minutes of Watching Fox News as Endogenous Variable

	Dependent variable:					
	Log sentencing length in months				D(Incarceration)	
Sample	No probations and acquittals		All			
	I	II	III	IV	V	VI
Fox viewership (minutes)	0.099*	0.100**	-0.018	-0.030	0.100*	0.100*
	(0.051)	(0.045)	(0.020)	(0.027)	(0.051)	(0.053)
Judge FE & tenure		✓		✓		✓
State-year FEs	✓	✓	✓	✓	✓	✓
Demographic controls	✓	✓	✓	✓	✓	✓
Case controls	✓	✓	✓	✓	✓	✓
Partial R-squared	0.170	0.124	0.234	0.171	0.234	0.171
Anderson-Rubin p-value	0.032	0.012	0.047	0.061	0.047	0.061
F-stat. of excl. inst.	68.508	56.503	96.989	71.309	96.989	71.309
Observations	2,007,519	2,007,519	2,521,509	2,521,509	2,521,509	2,521,509

Notes: This table replicates Table 2 but uses ratings (proportional to number of minutes of Fox News watched) instead of share of viewership. Standard errors clustered by county are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table B.5: Robustness for Columns III and IV of Table 2: Logs Instead of Inverse Hyperbolic Sine

	Dependent variable: Log	
Sample	All	
	I	II
Fox viewership	0.054*	0.060*
	(0.028)	(0.032)
Judge FE & tenure		✓
State-year FEs	✓	✓
Demographic controls	✓	✓
Case controls	✓	✓
Partial R-squared	0.094	0.046
Anderson-Rubin Wald test p-value	0.074	0.061
F-stat. of excl. inst.	76.9	41.1
Observations	2,521,509	2,521,509

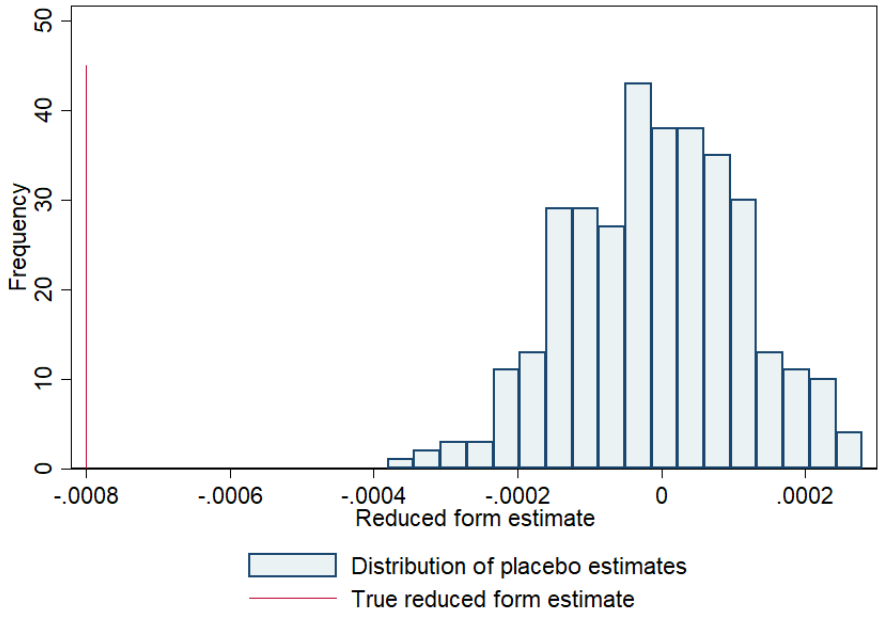
Notes: This table replicates Columns III and IV of Table 2 but uses $\log(x+1)$ instead of inverse hyperbolic sine transformation. Standard errors clustered by county are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table B.6: Robustness for Table 2: Results for Specifications with Judge Fixed Effects without North Carolina

	Dependent variable:		
	Log sentencing length in months		D(Incarceration)
Sample	No probations and acquittals	All	
	I	II	III
Fox viewership	0.078**	-0.025	0.074*
	(0.037)	(0.021)	(0.038)
Judge FE & tenure	✓	✓	✓
State-year FEs	✓	✓	✓
Demographic controls	✓	✓	✓
Case controls	✓	✓	✓
Partial R-squared	0.052	0.034	0.034
Anderson-Rubin p-value	0.057	0.009	0.009
F-stat. of excl. inst.	34.9	26.4	26.4
Observations	1,785,333	2,280,380	2,280,380

Notes: This table replicates Columns II, IV, and VI of Table 2 but omits North Carolina (the state, where judges are rotating between counties withing judicial district). Standard errors clustered by county are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Figure B.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 1.

Table B.7: Fox and Sentencing: Robustness to Inclusion of Additional Controls

	Dependent variable: Log sentencing length in months						
	I	II	III	IV	V	VI	VII
Fox News viewership	0.075***	0.074**	0.081**	0.063**	0.071***	0.081**	0.070**
	(0.028)	(0.029)	(0.033)	(0.032)	(0.023)	(0.033)	(0.0333)
1996 Republican vote share		0.166					0.0166
		(0.147)					(0.122)
Log population			-0.036				-0.0375
			(0.065)				(0.0686)
Share population w Fox News				0.103*			0.0712
				(0.056)			(0.0512)
t-1 avg. sentencing length					0.011***		0.010***
					(0.001)		(0.000356)
t-1 crime rates						-0.027	-0.011
						(0.049)	(0.0503)
Partial R-squared	0.029	0.028	0.017	0.021	0.022	0.017	0.021
F-stat. of excl. inst.	33.1	32.6	25.2	25.9	27.1	25.2	16.9
Observations	4,974,207	4,974,207	4,974,207	4,974,207	4,974,207	4,974,207	4,974,207

Notes: Regression results using NCRP data. The dependent variable is the log of the sentencing length. All columns use baseline specification from Column III of Table 1 but add additional controls. Republican vote share in 1996, population in 2010, and share of population with Fox News in 2005 only have cross-sectional variation. Share of population with Fox News is computed as total population of zipcodes that have Fox News divided by the total population of all zipcodes in that county. Lagged sentencing and crime rates vary on county-year level. Standard errors clustered by county are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table B.8: Fox and Sentencing: Alternative Clustering

	Dependent variable: Log sentencing length in months					
	I	II	III	IV	V	VI
Clustering	Baseline/county	State	County-year	State-year	DMA	DMA-year
Fox News viewership	0.075***	0.075**	0.075**	0.075**	0.075**	0.075**
	(0.028)	(0.033)	(0.027)	(0.032)	(0.024)	(0.023)
F-stat. of excl. inst.	33.1	22.0	27.4	21.0	17.7	16.6
Observations	4,974,207	4,974,207	4,974,207	4,974,207	4,974,207	4,974,207

Notes: Regression results using NCRP data. The dependent variable is the log of the sentencing length. All columns use baseline specification from Column III of Table 1. Each column has exactly the same specification but use different clustering for computing standard errors. *** p<0.01, ** p<0.05, * p<0.1

Table B.9: Robustness for Table 2: Results are Driven by the States with Elected Judges (i.e., w/o Virginia)

	Dependent variable:		
	Log sentencing length in months		D(Incarceration)
Sample	No probations and acquittals	All	
	I	II	III
Fox viewership	0.097**	0.060	-0.038
	(0.049)	(0.045)	(0.026)
Judge FE & tenure	✓	✓	✓
State-year FEs	✓	✓	✓
Demographic controls	✓	✓	✓
Case controls	✓	✓	✓
Partial R-squared	0.03	0.05	0.05
Anderson-Rubin p-value	0.054	0.010	0.010
F-stat. of excl. inst.	18.0	25.8	25.8
Observations	1,845,979	2,294,908	2,294,908

Notes: This table replicates Columns II, IV, and VI of Table 2 but omits Virginia (the state, where judges are appointed or re-appointed). Standard errors clustered by county are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

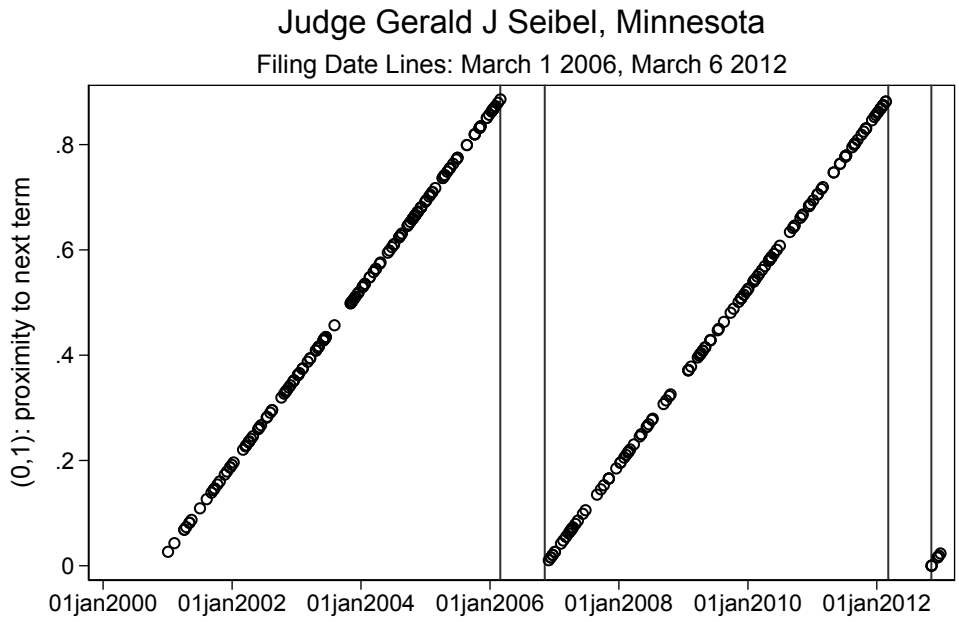
Table B.10: Fox and Sentencing: Placebo with Lagged Sentencing

	Dependent variable: Log sentencing length in months (lagged)											
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
Lag Y_t	t-15				t-10				t-10			
Sample	All				All				Appointed		Elected	
Years	1990-2005		1990-2000		1995-2005		1995-2000		1995-2005		1995-2005	
	RF	2SLS	RF	2SLS	RF	2SLS	RF	2SLS	RF	2SLS	RF	2SLS
Fox channel position	-0.0001		-0.0001		-0.0001		0.0014		0.0008		-0.0002	
	(0.0006)		(0.0006)		(0.0004)		(0.0009)		(0.0008)		(0.0004)	
Fox viewership		0.0050		0.0117		0.0101		-0.2947		-0.0678		0.0143
		(0.0447)		(0.0662)		(0.0357)		(0.5590)		(0.0670)		(0.0364)
Partial R-squared		0.092		0.037		0.048		0.001		0.053		0.047
F-stat. of excl. inst.		13.8		5.4		17.1		0.4		14.7		15.8
Observations	3,039,291	3,039,291	1,906,782	1,906,782	3,209,113	3,209,113	1,014,006	1,014,006	442,747	442,747	3,004,141	3,004,141

Notes: Regression results using NCRP data. All columns use baseline specification from Column III of Table 1 but shifts the time t of the case by 15 and 10 years for columns I–IV and V-VIII, respectively. Columns IX–XII do the same for the 10 year shift but split the sample by locations with appointed and elected judges. The dependent variable is the log of the sentencing length. Standard errors clustered by county are in parentheses. *** $p<0.01$, ** $p<0.05$,* $p<0.1$

Electoral Cycles Analysis

Figure B.7: Illustration of Election Cycle Data



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, 2021) 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.

To dig further into the election mechanism, we looked 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 Dippel-Poyker's 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 τ_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 $T_{NC} = 8 \times 365 + 2$.[50] Thus τ_j is scaled as $\tau_j \in [0; 1]$. Appendix Figure B.7 provides a visualization of what an electoral cycle looks like in the data.

[50] 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 (2021).

Table B.11: Fox Viewership and Electoral Cycles

	Dependent variable: Log sentencing length in months		
	I	II	III
Fox viewership	0.097**	0.100**	0.073**
	(0.049)	(0.048)	(0.036)
Proximity-to-election	0.009*	0.131	
	(0.005)	(0.124)	
Proximity-to-election		-0.050	
x Fox viewership		(0.050)	
Proximity-to-election in NC, PA, and WA			0.238**
			(0.121)
Proximity-to-election in NC, PA, and WA			-0.097*
x Fox viewership			(0.053)
Partial R-squared	0.03	0.03 & 0.04	0.04 & 0.05
F-stat. of excl. inst.	18	16 & 3	21 & 4
Observations	1,381,186	1,381,186	1,381,186

Notes: The dependent variable is the inverse hyperbolic sine of sentencing length. All columns use the specification from Column IV of Table 2. Column I adds proximity-to-election as a control. Columns II and III additionally add the interaction with Fox viewership. In Column III proximity-to-election is set to zero for all states for the exception of three states with strong electoral cycles: North Carolina, Pennsylvania, and Washington. This interaction is instrumented by the interaction of proximity-to-election with Fox channel position. The number of observations is smaller than Table 2 Column II because not all judges can be correctly merged to their biography at the end points of the time interval for which sentencing data is available. Standard errors clustered by county are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

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:

$$T_{ct} = \alpha_{st} + \gamma_1 Z_{ct} + \gamma_2 Z_{ct} \cdot \tau_j + \beta\tau_j + X_{i(c)t}\beta + \mu_j + \eta_{1i(c)t}, \tag{9}$$

$$T_{ct} \cdot \tau_j = \alpha_{st} + \gamma_1 Z_{ct} + \gamma_2 Z_{ct} \cdot \tau_j + \beta\tau_j + X_{i(c)t}\beta + \mu_j + \eta_{2i(c)t} \tag{10}$$

where τ_j defines proximity to elections. The second stage is

$$Y_{i(c)t} = \alpha_{st} + \rho_1 \widehat{T}_{ct} + \rho_2 \widehat{T_{ct} \cdot \tau_j} + \beta\tau_j + X_{i(c)t}\beta + \mu_j + \epsilon_{i(c)t}, \tag{11}$$

which now includes judge fixed effects (μ_j).

The electoral cycles results are reported in Appendix Table B.11. We first replicate the most conservative specification from Column IV of Table 2, but now add the proximity-to-election as a separate control.[51] The coefficient for Fox News viewership is significant,

[51]Note, that the number of observation decreases, because not all judges can be uniquely matched to their biographies at the end points of data. See Section 2 in Dippel and Poyker (2021) for complete discussion of construction of electoral cycles.

and the proximity to election is also positive and marginally significant. To get at the interaction between Fox News and the judicial election cycle, we produce estimates for the 2SLS System (Equations 9, 10, and 11) in Column II. 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 and elections is marginally smaller as elections become more immediate; however, it is insignificant on any conventional level.[52] Finally, in Column III we only try to find heterogeneous effects on the sample of three states with the strongest electoral cycles; Dippel and Poyker (2021) find that only electoral cycles in North Carolina, Pennsylvania, and Washington exhibit robust significance in our sample of ten states. While the coefficient for electoral cycle is indeed becoming larger in magnitude, the interaction remains negative.

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.

[52]We can't estimate the specification in Column II 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.

Table B.12: Robustness for Table 6: Adding Judge's Starting Year Fixed Effects

	Dependent variable: Average judges' harshness				
	I	II	III	IV	V
Fox viewership	0.052**	-0.098	0.031	0.037	-0.004
	(0.022)	(0.085)	(0.028)	(0.043)	(0.022)
Fox viewership					
x # donors per judge-race		0.007*			
		(0.004)			
x Prob. electoral challenge			0.001		
			(0.002)		
x Length of electoral cycle				0.002	
				(0.007)	
x Partisan election					0.465*
					(0.268)
Judge starting-year FEs	✓	✓	✓	✓	✓
Anderson-Rubin p-value	0.016	0.014	0.038	0.045	0.018
F-stat. of excl. inst.	164	76.9	78.6	89.6	5.3
Observations	6,793	5,701	6,793	6,793	6,793

Notes: This table replicates Table 6 but add in all columns a matrix of dummies for the years when all judges in in county c year t appeared in the Dippel-Poyker dataset for the first time. Standard errors clustered by county are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B.13: Locations with Higher Fox Consumption do not Have Higher Judge Turnover

	Dependent variable: D(Judge changed)	
	I	II
Fox viewership	-0.016	0.028
	(0.017)	(0.057)
Judge starting-year FEs		✓
Anderson-Rubin p-value	0.02	0.02
F-stat. of excl. inst.	167	78.3
Observations	6,793	6,793

Notes: Observation in this table is county-year. The dependent variable is a dummy equal to one if a new judge appears in county c year t and zero otherwise. All columns include state-year fixed effects. Column II includes matrix of dummies for the years when all judges in in county c year t appeared in the Dippel-Poyker dataset for the first time. Standard errors clustered by county are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B.14: Selection Effect of Fox News Based on Availability at Judge Starting Year

	Dependent variable:					
	Log sentencing length in months				D(Incarceration)	
Sample	No probations and acquittals		All			
	I	II	III	IV	V	VI
Share of zip-codes with Fox News	0.150*	0.170	0.092	0.153	-0.006	-0.008
at the year judge is selected	(0.081)	(0.114)	(0.081)	(0.111)	(0.016)	(0.019)
County FEs	✓		✓		✓	
County-year FEs		✓		✓		✓
State-year FEs	✓		✓		✓	
Demographic controls	✓	✓	✓	✓	✓	✓
Case controls	✓	✓	✓	✓	✓	✓
R-squared	0.68	0.68	0.51	0.52	0.36	0.37
Observations	2,007,519	2,007,519	2,521,509	2,521,509	2,521,509	2,521,509

Notes: Regression results using Dippel-Poyker data. Includes data from Alabama, Arkansas, Colorado, Georgia, Kentucky, Minnesota, North Carolina, Pennsylvania, Tennessee, Virginia, and Washington. The dependent variable in Columns I and II is the log of the sentencing length. The dependent variable in Columns III and IV is the inverse hyperbolic sine of the sentencing length. The dependent variable in Columns V and VI is a dummy equal to one if the defendant is sentenced to a term in prison and zero otherwise (i.e., probation or acquittal). All columns use the baseline specification from Column III of Table 1. Offense codes, crime severity, and recidivism variables are included as fixed effects which may vary across states. Treatment variable is county's population-weighted share of zip codes with Fox News at the year when the judge is selected. In cases when judge is selected for the first time before Fox News is established, we set the share to be zero. Appointed judges (in Virginia) are assigned with the Fox News channel position at the year of their appointment. Standard errors clustered by county are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table B.15: Robustness for Table 7: Adding State-Specific log Total # of Sentences

	Dependent variable: Log sentencing length in months			
	I	II	III	IV
Fox viewership	0.040	0.041	0.054	0.053
	(0.032)	(0.033)	(0.034)	(0.034)
Fox viewership x				
Log # crimes mentions	0.012***			
at month-year t	(0.0044)			
Log # drugs mentions		0.013***		
at month-year t		(0.0039)		
Log # crimes & black mentions			0.027***	
at month-year t			(0.0103)	
Log # drugs & black mentions				0.052*
at month-year t				(0.0283)
F-stat. of excl. inst.	76 & 96	80 & 94	60.3 & 31	67 & 7.4
Observations	2,521,068	2,521,068	2,521,068	2,521,068

Notes: This table contains robustness checks for Table 7. It replicates Columns I–IV of Table 7 but includes the total monthly number of sentences spoken on Fox transcripts, interacted with state fixed effects. Standard errors clustered by county are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B.16: Fox News and Mentions of Various Types of Drugs

	Dependent variable: Log sentencing length in months						
	I	II	III	IV	V	VI	VII
Fox viewership	Marijuana	Cocaine	Crack	Ecstasy	Meth	PCP	Heroin
	0.045	0.045	0.042	0.053	0.055	0.067*	0.044
	(0.033)	(0.032)	(0.033)	(0.076)	(0.035)	(0.036)	(0.033)
Fox viewership x Log # mentions at month-year t							
Marijuana	0.016***						
	(0.005)						
Cocaine		0.027**					
		(0.011)					
Crack			0.018***				
			(0.006)				
Ecstasy				0.530			
				(3.110)			
Meth					0.027***		
					(0.010)		
PCP						0.097	
						(0.092)	
Heroin							0.028**
							(0.011)
Log # X mentions	✓	✓	✓	✓	✓	✓	✓
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 IV of Table 7, but only counts number of mentions for a particular drug. The dependent variable is the log of the sentencing length. Standard errors clustered by county are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B.17: Text Analysis Placebo: Table 7 with CNN Transcripts

	Dependent variable: Log sentencing length in months							
	I	II	III	IV	V	VI	VII	VIII
Fox viewership	0.298	-0.045	0.121**	0.121**	0.302	-0.064	0.119**	0.129**
	(0.435)	(0.244)	(0.052)	(0.051)	(0.440)	(0.241)	(0.051)	(0.050)
Fox viewership x								
Log # crimes mentions in CNN	-0.022				-0.022			
at month-year t	(0.058)				(0.059)			
Log # drugs mentions in CNN		0.025				0.028		
at month-year t		(0.035)				(0.035)		
Log # crimes & black mentions in CNN			0.010				0.012	
at month-year t			(0.026)				(0.025)	
Log # drugs & black mentions in CNN				0.005				0.001
at month-year t				(0.016)				(0.014)
Month FEs					✓	✓	✓	✓
Log # X mentions	✓	✓	✓	✓	✓	✓	✓	✓
F-stat. of excl. inst.	34	43	35	35	34	43	35	35
Observations	2,521,068	2,521,068	2,521,068	2,521,068	2,521,068	2,521,068	2,521,068	2,521,068

Notes: This table replicates Table 7 but uses CNN's transcripts. Standard errors clustered by county are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B.18: Text Analysis Placebo: Table 7 with MSNBC Transcripts

	Dependent variable: Log sentencing length in months							
	I	II	III	IV	V	VI	VII	VIII
Fox viewership	-0.180	0.058	0.123**	0.128**	-0.162	0.054	0.123**	0.128**
	(0.187)	(0.190)	(0.054)	(0.057)	(0.182)	(0.192)	(0.054)	(0.056)
Fox viewership x								
Log # crimes mentions in MSNBC	0.058				0.055			
at month-year t	(0.037)				(0.037)			
Log # drugs mentions in MSNBC		0.012				0.013		
at month-year t		(0.033)				(0.033)		
Log # crimes & black mentions in MSNBC			0.010				0.009	
at month-year t			(0.014)				(0.014)	
Log # drugs & black mentions in MSNBC				0.002				0.002
at month-year t				(0.015)				(0.014)
Month FEs					✓	✓	✓	✓
Log # X mentions	✓	✓	✓	✓	✓	✓	✓	✓
F-stat. of excl. inst.	34	43	35	35	34	43	35	35
Observations	2,521,068	2,521,068	2,521,068	2,521,068	2,521,068	2,521,068	2,521,068	2,521,068

Notes: This table replicates Table 7 but uses MSNBC's transcripts. Standard errors clustered by county are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

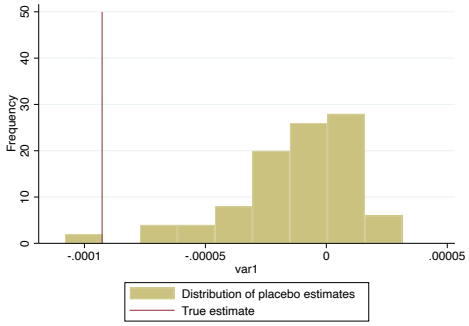
Table B.19: Robustness for Table 7: Additional Controls for Transcripts of CNN and MSNBC

	Dependent variable: Log sentencing length in months							
	I	II	III	IV	V	VI	VII	VIII
Fox viewership	0.051	0.052	0.054	0.054	0.051	0.052	0.053	0.057
	(0.038)	(0.039)	(0.035)	(0.035)	(0.038)	(0.039)	(0.035)	(0.035)
Fox viewership x								
Log # crimes mentions	0.010**				0.010**			
at month-year t	(0.005)				(0.005)			
Log # drugs mentions		0.010**				0.010**		
at month-year t		(0.005)				(0.005)		
Log # crimes & black mentions			0.028***				0.029***	
at month-year t			(0.011)				(0.011)	
Log # drugs & black mentions				0.048*				0.040*
at month-year t				(0.028)				(0.024)
CNN + mentions x CNN + mentions	✓	✓	✓	✓	✓	✓	✓	✓
MSNBC + mentions x MSNBC + mentions	✓	✓	✓	✓	✓	✓	✓	✓
Month FEs					✓	✓	✓	✓
Log # X mentions	✓	✓	✓	✓	✓	✓	✓	✓
F-stat. of excl. inst.	27	31	35	37	25	26	36	37
Observations	2,521,068	2,521,068	2,521,068	2,521,068	2,521,068	2,521,068	2,521,068	2,521,068

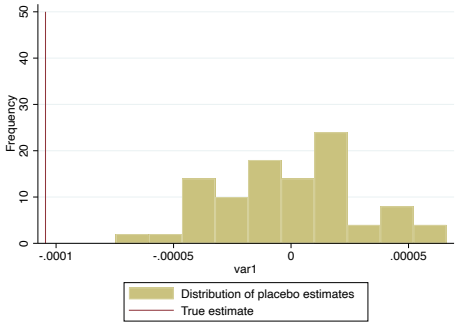
Notes: TThis table replicates Table 7 but additionally controls for the CNN and MSNBC viewership, log number of X mentions (crime, drugs, crime & black, or drugs & black), and their interactions. Standard errors clustered by county are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Figure B.8: Permutation Test: Reduced Form Effect of the Interaction Term from Columns I–IV of Table 7

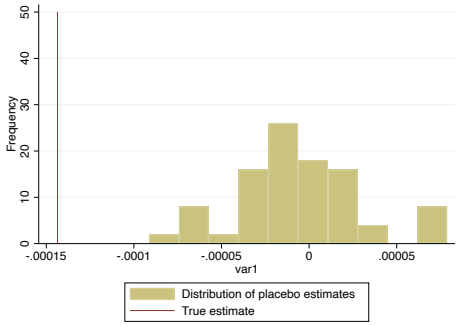
Panel A (crime × Fox channel position)



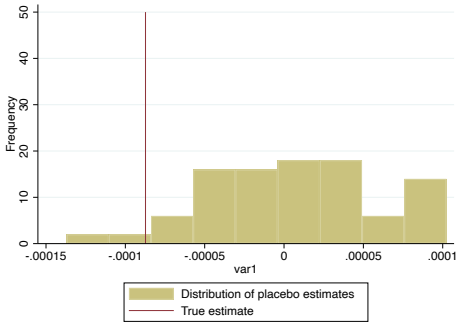
Panel B (drugs × Fox channel position)



Panel C (crime & Black × Fox channel position)



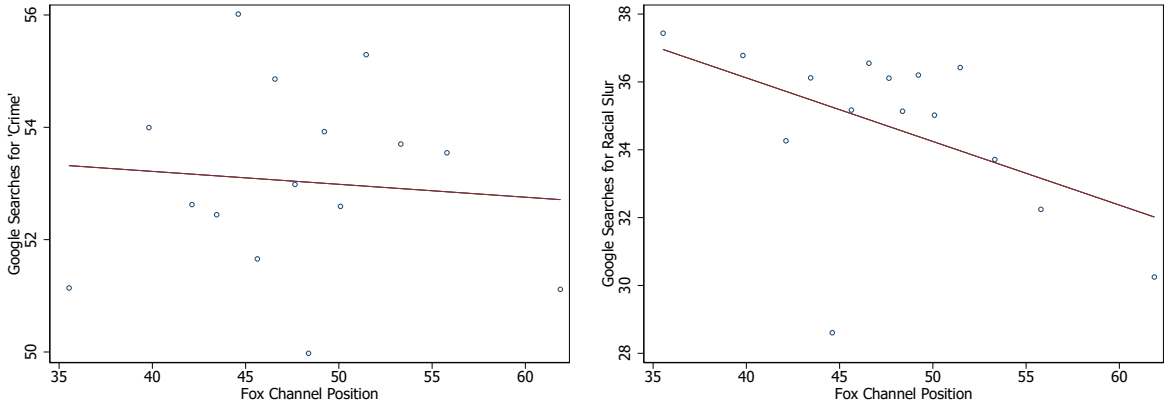
Panel D (drugs & Black × Fox channel position)



Notes: Based on 100 placebo Fox News channel positions interacted with log number of mentions of crime, drugs, crime and Black, or drugs and Black. All regressions use baseline specifications from Columns I–IV of Table 7. The true coefficient is shown with the vertical red line.

Figure B.9: Effects of Fox News Channel on Google Search

Panel A (Searches for "crime") Panel B (Searches for racial slur)



Notes: Binscatter diagrams relating Fox channel position (at DMA level) to Google searches on "crime" (Panel A) and Google searches on a racial slur (Panel B), 2005–2008. Includes state fixed effects and controls for position of CNN and MSNBC.