Bursztyn, Leonardo; Rao, Aakaash; Roth, Christopher; Yanagizawa-Drott, David

Working Paper

Opinions as facts

ECONtribute Discussion Paper, No. 159

Provided in Cooperation with:
Reinhard Selten Institute (RSI), University of Bonn and University of Cologne

ECONtribute
MARKETS & PUBLIC POLICY

ECONtribute Discussion Paper No. 159

Opinions as Facts

Leonardo Bursztyn
Aakaash Rao
Christopher Roth
David Yanagizawa-Drott

April 2022
www.econtribute.de

UNIVERSITÄT BONN
Universität zu Köln

Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany´s Excellence Strategy – EXC 2126/1– 390838866 is gratefully acknowledged.

Opinions as Facts*

Leonardo Bursztyn[†]
Aakaash Rao[‡]
Christopher Roth[§]
David Yanagizawa-Drott[¶]

April 1, 2022

Abstract

The rise of opinion programs has transformed television news. Because they present anchors' subjective commentary and analysis, opinion programs often convey conflicting narratives about reality. We experimentally document that people across the ideological spectrum turn to opinion programs over "straight news," even when provided large incentives to learn objective facts. We then examine the consequences of diverging narratives between opinion programs in a high-stakes setting: the early stages of the COVID-19 pandemic in the US. We find stark differences in the adoption of preventative behaviors among viewers of the two most popular opinion programs, both on the same network, which adopted opposing narratives about the threat posed by the COVID-19 pandemic. We then show that areas with greater relative viewership of the program downplaying the threat experienced a greater number of COVID-19 cases and deaths. Our evidence suggests that opinion programs may distort important beliefs and behaviors.

JEL Codes: C90, D83, D91, Z13
Keywords: Opinion programs, Media, Narratives

*This draft supersedes a previous draft circulated under the title "Misinformation During a Pandemic." We thank the editor (Nicola Gennaioli) and three anonymous reviewers for excellent and constructive comments. We thank Alberto Alesina, Davide Cantoni, Bruno Caprettini, Ruben Durante, Eliana La Ferrara, Ed Glaeser, Nathan Nunn, Ricardo Perez-Truglia, Andrei Shleifer, David Yang, Noam Yuchtman, and numerous seminar participants for very helpful comments and suggestions. We thank Silvia Barbareschi, Aditi Chitkara, Jasurbek Berdikobilov, Hrishikesh Iyengar, Rebecca Wu, Alison Zhao, and especially Vanessa Sticher for outstanding research assistance. We are grateful to the Becker Friedman Institute for financial support. The experiment was pre-registered on the AEA RCT registry under ID AEARCTR-0006958. Roth: Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germanys Excellence Strategy EXC 2126/1-390838866.
[†]University of Chicago and NBER, bursztyn@uchicago.edu
[‡]Harvard University, arao@g.harvard.edu
[§]University of Cologne, Econtribute, CAGE Warwick, CESifo, CEPR, briq, roth@wiso.uni-koeln.de
[¶]University of Zurich and CEPR, david.yanagizawa-drott@econ.uzh.ch

1 Introduction

Over the past two decades, opinion programs have come to dominate cable television news. Unlike "straight news," which professes to impartially report "just the facts," opinion programs convey their anchor's perspective on the news of the day. They typically feature little original reporting; instead, they focus on story-telling, entertainment, and *subjective* commentary (Kavanagh et al., 2019), at the expense of objective factual reporting (Kavanagh and Rich, 2018). Consequently, different opinion programs often present distinct, and often conflicting, narratives about reality.

Cable networks themselves distinguish their "hard" or "straight" news reporting from their opinion content. For example, when defending a leading anchor from defamation claims, Fox News successfully argued that "the 'general tenor' of the show should then inform a viewer that the host is not 'stating actual facts' about the topics he discusses and is instead engaging in 'exaggeration' and 'non-literal commentary'."[1] MSNBC successfully adopted the same approach: "For her to exaggerate the facts... was consistent with her tone up to that point, and the court finds a reasonable viewer would not take the statement as factual given this context".[2] Emphasizing the difference between straight news and opinion, Fox News President Jay Wallace wrote:

> "We've always said that we have strong opinion and strong news. And, again, I think that's part of the success. You know what you're getting."[3]

Do viewers know what they're getting? If viewers interpret opinion programs appropriately, then such programs may make valuable contributions to political discourse: they are generally more engaging than straight news programs and they can distill complex issues into easy-to-understand narratives (Jacobs and Townsley, 2011). On the other hand, if viewers trust the literal statements made on opinion programs just as they would those made on straight news, failing to distinguish between opinion and fact and to appropriately discount hyperbole and speculation, then diverging narratives across programs can lead different segments of the population to hold dramatically different views of reality. Commenting on this phenomenon in 2010, veteran journalist Ted Koppel wrote:

> "Daniel Patrick Moynihan's oft-quoted observation that 'everyone is entitled to his own opinion, but not his own facts,' seems almost quaint in an environment that flaunts opinions as though they were facts."[4]

In this paper, we demonstrate that viewers turn to opinion programs for information about objective facts, and we explore the consequences of this trust for high-stakes outcomes.

[1] See "McDougal v. Fox News Network." *JUSTIA US Law*, 2020.

[2] See "Herring Networks, Inc. v. Maddow." *Casetext: Smarter Legal Research*, May 22, 2020.

[3] See "Fox News Exec Jay Wallace Gets Candid About Ratings, White House Access (Q&A)." *The Hollywood Reporter*, January 2, 2018.

[4] See "Ted Koppel: Olbermann, O'Reilly and the death of real news." *The Washington Post*, November 14, 2010.

We begin with a pre-registered motivating experiment conducted with a sample of regular viewers of the two most popular cable news networks, Fox News and MSNBC. We tell participants that they will provide their best guess about an objective statistic relating either to the spread of the COVID-19 pandemic or to one of four dimensions of the country's economic performance, all as of a randomly-selected recent date. In order to inform their guess, respondents can choose one of four TV clips, which were all excerpted from shows broadcast on the same week as the date pertinent to their guess. These four clips comprise the two most popular straight news programs and the two most popular opinion programs on their network. 75% of Fox News viewers choose an opinion program over a straight news program, as do 60% of MSNBC viewers. Varying the reward for a correct answer from $10 to $100 has a precisely estimated zero effect, suggesting that viewers trust opinion programs to reveal factual information even when making choices with relatively higher stakes.

Programs that report objective content can differ primarily in their choices of what to cover, perhaps shaping viewers' perceptions of what is important (Bordalo et al., 2021) but with less scope for shaping viewers' beliefs about *facts* (Bennett, 2016). Because they are unconstrained by objectivity, on the other hand, opinion programs can not only diverge more in their selection of topics to cover, but they can also present different narratives about the *same* set of underlying facts. This is particularly important given the dominant — and growing — role of opinion content in primetime cable news: different anchors, each drawing weekly audiences of several million, can present dramatically different narratives about reality.[5] Do these diverging narratives have consequences for real-world outcomes? Identifying the causal effect of these narratives on behavior is challenging for several reasons: most importantly, ruling out alternative explanations for behavioral differences among consumers of different opinion programs — such as different prior beliefs, different ideologies, or different preferences — generally requires a setting in which two opinion programs that are *ex ante* similar, both in their content and in the characteristics of their viewers, suddenly and sharply diverge in their coverage of a given topic, and moreover that this topic can be linked to naturally-occurring outcomes.

To overcome these empirical challenges, we examine the two most popular opinion programs in the United States: *Hannity* and *Tucker Carlson Tonight*. These shows are aired back-to-back on the same network (Fox News) and had similar content prior to January 2020; indeed, we show using natural language processing techniques that both in their selection of topics and the way they covered these topics, these shows were *more* similar than almost any other pair of primetime shows across Fox News and MSNBC. However, we document that the programs differed sharply along both margins in their reporting about COVID-19. While both narratives were consistent with the anchors' right-wing slant, they had very different implications for viewers' beliefs and behavior.

[5] See "Fox News Changes Up Daytime Lineup, Adds New Opinion Show at 7 p.m." *The Hollywood Reporter*, January 11, 2021.

Carlson emphasized the severity of the threat as early as January while placing blame on China for its lack of transparency with the international community, later hosting a Chinese virologist who alleged that COVID-19 is a bio-weapon created by the Chinese Communist Party.[6] In contrast, Hannity largely ignored or downplayed the threat posed by the virus through February and early March, blaming Democrats for using it as a political weapon to undermine the administration.[7] In the narratives they presented about the dangers of COVID-19, Carlson and Hannity were largely outliers (in opposite directions), not only on Fox News, but on broadcast and cable television as a whole — a striking divergence given the two programs' prior similarities. Focusing on these two opinion programs within the same network enables us to compare two *ex ante* similar viewer populations, allowing us to examine how exposure to diverging narratives broadcast on opinion programs drives beliefs, behavior, and downstream health outcomes.

To shed light on the timing of common behavioral adjustments at the early stages of the pandemic (such as washing hands more often, cancelling travel plans, and avoiding large events), we fielded a survey among 1,045 Fox News viewers aged 55 or older. Consistent with a persuasive effect of content on behavior, we find that viewership of *Hannity* is associated with changing behavior four days later than other Fox News viewers, while viewership of *Tucker Carlson Tonight* is associated with changing behavior three days earlier (controlling for demographics and viewership of other programs and networks). Given the critical importance of early preventive measures (Bootsma and Ferguson, 2007; Markel et al., 2007), these differences in the timing of adoption of cautious behavior may have significant consequences for health outcomes.[8]

Motivated by our survey evidence, we examine disease trajectories in the broader population using county-level data on COVID-19 cases and deaths. We first show that, controlling for a rich set of county-level demographics (including the local market share of Fox News), greater local viewership of *Hannity* relative to *Tucker Carlson Tonight* is associated with a greater number of COVID-19 cases starting in early March and a greater number of deaths resulting from COVID-19 starting in mid-March. We then employ an instrumental variable approach that shifts relative viewership of the two programs, yet is plausibly orthogonal to local preferences for the two programs and to any other county-level characteristics that might affect the virus' spread. In particular, we predict this difference in viewership using the product of (i) the fraction of TVs on during the start

[6]See "Tucker Carlson: Racist for Saying 'Chinese Coronavirus'? Now's Not the Time for the Dumbest Identity Politics." *Fox News*, March 12, 2020. "Tucker Carlson Blames Media for Coronavirus Spread: 'Wokeness Is A Cult. They'd Let You Die' Over Identity Politics." *Newsweek*, February 24, 2020. Yan et al. (2020).

[7]See "Hannity Claims He's 'Never Called the Virus a Hoax' 9 Days After Decrying Democrats 'New Hoax'." *Vox*, March 20, 2020.

[8]For example, Pei et al. (2020) estimate that approximately half of all COVID-19 deaths in the United States at the early stages of the pandemic could have been prevented had non-pharmaceutical interventions (NPIs) such as mandated social distancing and stay-at-home orders been implemented one week earlier. While the behavioral changes our survey respondents report are likely not as extreme, and our survey is representative only of Fox News viewers over the age of 55, this evidence nonetheless suggests that these differences in timing may have directly affected the spread of the pandemic.

time of *Hannity* (leaving out TVs watching *Hannity*) and (ii) the local market share of Fox News (leaving out *Hannity* and *Tucker Carlson Tonight*). The logic of our instrument is simple: if people like to turn on their TVs to watch *something* when *Hannity* happens to be on instead of *Tucker Carlson Tonight*, the likelihood that viewers are shifted to watch *Hannity* is disproportionately large in areas where Fox News is popular in general. We show that the interaction term is conditionally uncorrelated with any among a larger number of variables that might independently affect the local spread of COVID-19, and we show that it strongly predicts viewership in the hypothesized direction. Using this instrument, we confirm the OLS findings that greater exposure to *Hannity* relative to *Tucker Carlson Tonight* is associated with a greater number of COVID-19 cases and deaths. Consistent with the gradual convergence in scripts between the two shows beginning in late February, the effects on cases plateau and begin to decline in mid-March, while effects on deaths follow two weeks later.

Turning to the underlying mechanisms, we find that differential viewership affects stay-at-home behavior (as measured by cell phone GPS data from two different sources), although this is unlikely to be the primary mechanism driving our effects. The sequential timing of differences in coverage, followed by differences in behavioral change, followed by differences in COVID-19 outcomes is inconsistent with several alternative potential drivers of our estimated treatment effects, such as time-invariant unobservables correlated with our instrument and differential effects of exposure to the programs that are unrelated to their reporting about COVID-19. Instead, the timing strongly suggests a causal chain from content differences to behavioral differences to COVID-19 outcomes. Taken together, our results suggest that viewers indeed trust opinion programs as sources of *facts*, beyond these programs' entertainment value.[9] Indeed, our findings indicate that this trust shapes important beliefs and behaviors.

Our work contributes to a large literature on the economic and social effects of the media (DellaVigna and La Ferrara, 2016). This literature has examined a wide range of political, behavioral, and health outcomes (Durante and Zhuravskaya, 2018; Eisensee and Strömberg, 2007; La Ferrara, 2016; Bursztyn et al., 2019; Muller and Schwarz, 2018; Martinez-Bravo and Stegmann, 2021; Yanagizawa-Drott, 2014; Levy, 2021), including the effect of Fox News on voting behavior (DellaVigna and Kaplan, 2007; Martin and Yurukoglu, 2017). Insofar as opinion shows are more entertaining than straight news shows (Berry and Sobieraj, 2013), our paper also relates to work on the effects of entertainment media on social and political outcomes (La Ferrara et al., 2012; Durante et al., 2019), particularly work examining the effects of *specific* television shows (Kearney and Levine, 2015; Banerjee et al., 2019). Methodologically, our work relates to a literature analyzing media content (Gentzkow and Shapiro, 2010; Djourelova, 2020).[10]

[9] Consistent with this interpretation, as we discuss in Section 2, viewers in our experiment believe that opinion programs are more informative than straight news programs.

[10] Related to our study is work (Simonov et al., 2020; Ash et al., 2020; Ananyev et al., 2020) using the channel numbers instrument developed by Martin and Yurukoglu (2017) to establish a causal effect of exposure to Fox

We provide the first direct evidence on the importance of opinion shows in driving high-stakes behaviors. Our approach holds fixed important mechanisms that may operate through exposure to biased media, such as increased partisanship or lower trust in science, which allows us to identify the effect of contemporaneous exposure to diverging narratives on behavior. Our incentivized experiments demonstrate that people seek out opinion programs when given incentives to *get the facts right*.

The remainder of this paper proceeds as follows. In Section 2, we show that viewers across the ideological spectrum turn to opinion programs over straight news even in the presence of incentives to learn objective facts. In Section 3, we examine the role of diverging narratives on opinion programs in shaping beliefs and behavior during the early stages of the COVID-19 pandemic. Section 4 discusses implications and concludes.

2 Trust in Opinion Shows

In this section, we examine trust in opinion and "straight news" programs. Viewers might seek out opinion programs for several reasons: opinion programs tend to be more emotional and engaging than straight news (Kavanagh et al., 2019), and they can distill complex issues into easy-to-grasp summaries, expose viewers to partisan perspectives, and provide a frame through which to interpret the news of the day. Widespread distrust in "straight news" also plays a role in driving demand for opinion programs: only 36% of adults (11% of Republicans and 68% of Democrats) report a "great deal" or a "fair amount" of trust in the news to report fully, accurately, and fairly (Gallup, 2021).[11] Both a cause and consequence of opinion programs' popularity is that they dominate "prime time," the window between 8pm and 11pm when TV viewership as a whole is highest.

Whatever mechanisms drive the demand for opinion programs, however, we would expect viewers seeking *objective facts* about the world to exhibit greater relative demand for "straight news" than viewers seeking entertainment or analysis. By their very nature, opinion programs are centered around conveying anchors' commentary on and interpretation of the news of the day rather than "just the facts". Indeed, Kavanagh and Rich (2018) summarize the growing dominance of opinions over factual reporting during prime-time between 2000 and 2017 as follows:

> We found a starker contrast between broadcast news presentation and prime-time cable

News as a whole on health-related behaviors. Our work differs in its focus on a specific mechanism: the role of diverging narratives on opinion shows in driving differences in behavior and health outcomes, holding partisanship fixed. Ananyev et al. (2020) additionally finds that Fox News viewership has a statistically significant effect on death rates, while Ash et al. (2020) estimates a positive but statistically insignificant effect. These results are largely consistent with our evidence that media exposure affected health outcomes, but the *narrative* portrayed mattered, with different programs on Fox News presenting dramatically different narratives about COVID-19.

[11] Sociologist Sarah Sobieraj comments that "Some fans believe [opinion] shows are the only sources they can trust for information. That's a pretty impressive feat, really, to have convinced someone that an opinion program offers the truth, while conventional news is riddled with bias." See "Wrath of the talking heads: How the Outrage Industry affects politics", *PBS News Hour*, February 28, 2014.

> programming in the post-2000 period. Compared with news presentation on broadcast television, programming on cable outlets exhibited a dramatic and quantifiable shift toward subjective, abstract, directive, and argumentative language and content based more on the expression of opinion than on reporting of events.

Do consumers believe that prime-time opinion programs are less likely to accurately cover facts than straight news programs? We conduct a pre-registered motivating experiment examining via revealed preference (1) the extent to which viewers turn to opinion programs rather than straight news in order to learn objective facts, and (2) how this preference is affected by substantial incentives for accuracy.[12]

Sample and Design In December 2020, we targeted a sample of 1,000 US-based respondents — 500 Republican Fox News viewers and 500 Democrat MSNBC viewers — in cooperation with Luc.id, a survey provider widely used in social science research (Bursztyn et al., 2022).[13] We inform respondents that at the end of the survey, they will provide a guess about a historical statistic relating to a particular domain. The domain varies by treatment group: respondents are told that they will guess either about (i) a general fact relating to the US economy, (ii) the unemployment rate in the US, (iii) annualized GDP growth in the US, (iv) median weekly earnings in the US, or (v) the number of COVID-19 cases, all as of a specific, randomly-selected date from recent years. We further inform respondents that if their guess lies within 5 percent of the official value, they will win an Amazon gift card. Respondents are told that the date about which they are guessing will be revealed only a few seconds before they need to make their guess, ensuring that they do not expect they will be able to find the answer by web search. We cross-randomize the value of the gift card: half of the respondents are offered a $10 gift card and half a $100 gift card. Respondents are further told that in order to inform their choice, they can choose one of four TV clips, which were all excerpted from shows broadcast on the same week as the randomly-selected date pertinent to their guess.

Fox News viewers are offered segments from the two most popular straight news programs on the network — *The Story with Martha MacCallum* and *Special Report with Bret Baier* — and from the two most popular opinion programs on the network — *Hannity* and *Tucker Carlson Tonight*. MSNBC viewers are similarly offered segments from the two most popular straight news programs and the two most popular opinion programs on the network: *MSNBC Live* and *The Beat with Ari Melber*, and *The Rachel Maddow Show* and *The Last Word with Lawrence O'Donnell*, respectively. Our key outcome of interest is whether the viewer chooses an opinion show or a news show.

Two aspects of the design merit further discussion. First, we deliberately ask respondents to

[12] The preregistration is available on the AEA RCT registry under ID AEARCTR-0006958, available at `https://www.socialscienceregistry.org/trials/6958`.

[13] The survey instrument is reproduced in Appendix F.

make guesses about historical statistics rather than to make predictions about future statistics. Since opinion programs focus relatively more than straight news on prediction about the future and relatively less on reporting about the current state of the world (Jacobs and Townsley, 2011), this design choice pushes us towards identifying a lower bound on trust in opinion shows. Moreover, this design choice allows us to deliver gift cards to respondents immediately if they guess correctly, avoiding the possibility that respondents believe researchers will fail to deliver a gift card in the future. Second, we deliberately choose objective economic statistics that are often covered in the news media (or, in the case of the COVID-19 statistic, were extensively covered during the period of interest). In contrast, there are far fewer political, cultural, or social statistics that are frequently covered. Moreover, economic statistics about the past are a domain in which we would expect to see the *lowest* selection of opinion shows, given that these shows generally favor political or cultural issues over economic issues (Berry and Sobieraj, 2013).

Results Figure 1 presents the fraction of respondents in each treatment choosing an opinion show, separately for Fox News and MSNBC viewers. The levels are relatively similar across all domains and reveal a substantial preference for opinion programs: roughly 75% of Fox News viewers and 60% of MSNBC viewers choose one of the two opinion shows. For none of the five outcomes in either of the two populations does the $100 incentive significantly reduce the fraction choosing an opinion show. Indeed, for the COVID-19 condition — the condition most directly relevant to the empirical application of our paper — the higher incentive *increases* the fraction choosing an opinion show, though the effect is not statistically significant. Table 1 replicates this analysis in regression table form and confirms that controlling for a range of individual demographics, including age, a set of race indicators, a Hispanic indicator, a male indicator, a set of education indicators, a set of household income indicators, and a set of employment indicators does not significantly affect the estimates.[14]

Our results indicate that respondents across the political spectrum do not internalize the differences in informativeness between news and opinion shows when making their choice of program. Thus, whatever other factors may influence their choice (for example, preferences for popular programs, entertainment value, or habit formation), the vast majority of respondents end up trusting opinion programs over straight news programs even in the presence of substantial incentives to learn objective facts — despite the fact that both Fox News and MSNBC have argued in court that viewers should not interpret their opinion programs as factual.

One related interpretation is that respondents believe that neither straight news nor opinion programs are at all informative for their guess. Under this belief, respondents' decision of which

[14]Through manual coding of episode scripts during the week relevant to the experiment, we find that straight news programs are indeed substantially more likely to cover the statistics of interest than opinion programs. In turn, viewers who choose a straight news program also make more accurate guesses than viewers who choose an opinion program, though this may reflect selection into the shows.

news source to consume is (ex-ante) payoff-irrelevant, so respondents may choose opinion programs simply because they are more entertaining. Given the low levels of trust in conventional media sources documented above, this is perhaps true for some respondents. Yet we consider it unlikely that this drives our results given that the patterns are highly robust across the five domains, and thus respondents would have to believe that none of the programs convey useful information across any of the dimensions we study. As an additional benchmark, we directly elicit respondents' beliefs about the likelihood that each program contained the information necessary for the guess. We find that 70% of Fox News viewers and 57% of MSNBC viewers believed that an opinion program was weakly *more* informative than either of the straight news shows, confirming the hypothesis that both conservative and liberal viewers indeed see opinion programs as more informative for objective facts.

3 Opinion Programs and High-Stakes Behavior in the Field

What are the consequences of this trust in opinion programming? In this section, we examine how opinion programs shaped beliefs and behavior during the early stages of the COVID-19 pandemic in the United States. This setting is ideally suited to exploring the role of opinion programs for two reasons: first, because the stakes involved in acquiring accurate information were relatively high; and second, because there was substantial disagreement about the threat posed by COVID-19 across different opinion programs.

3.1 Diverging Narratives about COVID-19

We focus on media coverage of COVID-19 on Fox News during the early stages of the COVID-19 pandemic. Fox News is the most watched cable network in the United States, with an average of 3.4 million total primetime viewers in the first quarter of 2020, compared to 1.9 million for MSNBC and 1.4 million for CNN (the other two of the "Big Three" US cable news networks).[15] Moreover, the median age of primetime Fox News viewers is 68, substantially higher than that of CNN and MSNBC viewers.[16] Both due to its reach and the fact that more than half of its audience is over the age of 65 — a group that the CDC warns is at elevated risk from COVID-19 — Fox News may exert substantial influence on COVID-19 outcomes. This is particularly true given that the elderly both watch more TV in general than the average US citizen and because they disproportionately rely on television for news and information (Pew, 2019).

[15] "Fox News Channel Ratings for First Quarter of 2020 Are the Highest in Network History." *Fox News*, March 31, 2020.

[16] "Half of Fox News' Viewers Are 68 and Older." *The Atlantic*, January 27, 2014.

Narratives adopted by Carlson vs. Hannity Our paper focuses on the two most widely-viewed cable news programs in the United States, both of which are opinion programs: *Hannity* and *Tucker Carlson Tonight*. These shows had an average of 4.2 million and 4 million daily viewers, respectively, during the first quarter of 2020.[17] Before COVID-19 began to spread in the United States in January 2020, *Hannity* and *Tucker Carlson Tonight* were relatively similar in content and viewership: both covered the news from a conservative perspective and were broadly supportive of President Trump's policy agenda (see the end of Section 3.1 for evidence on pre-2020 coverage). Yet as we document using qualitative evidence, text-analysis methods, and human coding of the shows' scripts, the two shows adopted very different narratives about COVID-19.

News outlets and politicians across the ideological spectrum, and even experts such as National Institute of Allergy and Infectious Diseases director Anthony Fauci, suggested throughout much of February that COVID-19 would likely be safely contained.[18] While most programs on broadcast and especially cable networks occasionally discussed COVID-19 in January and early-to-mid February, the topic did not comprise a substantial fraction of coverage until late February.[19] Tucker Carlson thus stood out not only among his colleagues at Fox News, but more broadly among both broadcast and cable news anchors, for his repeated insistence as early as late January that COVID-19 posed a serious threat to the United States.[20] For example, on January 28 — more than a month before the first COVID-19-related death in the US — Tucker Carlson spent a large portion of his show discussing the subject, and continued to do so throughout February.

In contrast, Hannity covered COVID-19 and its consequences substantially less than Carlson and other Fox shows, particularly during February, when the virus was first beginning to spread in the United States. Even after he began discussing it more prominently in February, he downplayed the threat the virus posed and emphasized that Democrats were politicizing the virus. By mid-March, after President Trump declared a national emergency in response to COVID-19, Hannity's coverage had converged to that of Carlson and other Fox News shows, emphasizing the seriousness of the situation and broadcasting CDC guidelines.

Extensive margin of COVID-19 coverage To more systematically evaluate differences in the extensive margin of coverage between primetime Fox News shows, we turn to a simple word-counting procedure. For each of the seven shows on Fox News airing between 5pm and 11pm local time across the four major time zones, we download episode transcripts from LexisNexis. We count

[17] Authors' calculations based upon Nielsen data.

[18] See "What Went Wrong with the Media's Coronavirus Coverage?" *Vox*, April 13, 2020.

[19] Budak et al. (2021) estimate that 10-15% of airtime was devoted to coverage of the pandemic on cable networks in mid-February, compared to approximately 15-20% on broadcast networks. This fraction began to rise sharply in the last week of February and stabilized at approximately 70% across both broadcast and cable networks by March 16.

[20] See, for example, "His Colleagues at Fox News Called Coronavirus a 'Hoax' and 'Scam.' Why Tucker Carlson Saw It Differently." *The LA Times*, March 23.

the number of times any of a small list of coronavirus-related terms are mentioned on each day and plot the results in Panel A of Figure 2.[21] In particular, the y-axis of the panel displays the log of one plus the word count on each day.

Compared to the other primetime shows, both *Hannity* and *Tucker Carlson Tonight* stand out. Both anchors first discussed COVID-19 in late January when the first US case was reported, but Carlson continued to discuss the subject extensively throughout February whereas Hannity did not again mention it on his show until the end of the month. The other shows fell somewhere between these two extremes. By early March, the word counts of all shows had converged.[22]

However, this simple procedure does not entirely capture differences in how shows discussed COVID-19. The qualitative evidence above suggests that while Hannity discussed COVID-19 as frequently as Carlson during early March, he downplayed its seriousness and accused Democrats of using it as a partisan tool to undermine the administration. To capture these differences in the intensive margin of coverage, we turn to human coding of the scripts.

Human coding of scripts Between April 2 and April 6, we recruited workers on Amazon Mechanical Turk to assess how seriously each of the seven shows portrayed the threat of COVID-19 between early February and mid-March. For each episode that contained at least one coronavirus-related term, five MTurk workers read the entire episode script and answered "Yes" or "No" to the following question: "Did [the show] indicate that the virus is likely to infect many people in the US, causing many deaths or serious illnesses, or that many have already become infected and have died or become seriously ill?" We explicitly asked respondents to answer the question based only on the scripts, not their own views on the subject. We impute "No" for each script that does not mention any coronavirus-related terms, and we code "Yes" as 1 and "No" as 0.[23]

Panel B of Figure 2 displays one-week rolling means of this variable for Carlson, Hannity, and the other shows. Throughout almost the entire period, MTurk workers rate Carlson as portraying the threat of COVID-19 more seriously than the other shows, and in turn rate the other shows as portraying the threat more seriously than Hannity. In line with the qualitative evidence highlighted above, Hannity converges to Carlson in early to mid-March.

Pre-period similarity between *Hannity* and *Tucker Carlson Tonight* Finally, we establish that *Hannity* and *Tucker Carlson Tonight* featured similar content prior to their divergence on COVID-19 by analyzing all 2019 transcripts of these programs, as well as the primetime programs on Fox News and MSNBC used in the experiment described in Section 2. Because human-coding

[21]The words are "coronavirus", "virus," "covid," "influenza", and "flu."

[22]We also conduct a similar content analysis of all major primetime shows on CNN and MSNBC and find little variation across shows in terms of the coverage of COVID-19 (see Appendix Figure A2).

[23]We calculate Fleiss' Kappa of inter-rater agreement, a commonly used measure to assess the reliability of agreement among more than two sets of binary or non-ordinal ratings, as $\kappa = 0.629$ ($p < 0.001$), suggesting "substantial agreement" (Landis and Koch, 1977).

exercises are infeasible given the large number of transcripts and word-counting methods are infeasible given that we do not have priors over how (if at all) the two programs diverge in their coverage of other topics, we turn to natural language processing techniques to establish that *Hannity* was very close to *Tucker Carlson Tonight* both in the topics covered and the way they covered them, relative to the similarity between *Hannity* and the other Fox News and MSNBC programs. We provide a brief summary of our methodology and findings here, relegating a more detailed discussion to Appendix C.

To quantify divergence in the selection of topics, we employ Latent Dirichlet Allocation (Blei et al., 2003; Schwarz, 2018), an unsupervised topic model that finds underlying "topics" and then outputs vectors for each transcript characterizing the extent to which the transcript is composed of each topic. The resulting 10 topics are intuitively coherent: for example, one topic containing terms such as "Mueller," "prosecutor," "obstruction," and "judge" pertains to the Mueller investigation, while another, containing terms such as "storm," "hurricane," "water," and "Bahamas" captures hurricanes and other natural disasters.[24] For each week, we calculate the Euclidean (L2) distance between the topic vector for *Hannity* and that of each other program, plotting results in Panel A of Appendix Figure A1. The results show that *Hannity* is substantially more similar to *Tucker Carlson Tonight* than to any other program with the exception of *The Story*, which is equally similar. Reassuringly for our measure of distance, the figure also shows that programs on MSNBC are less similar to *Hannity* than programs on Fox News.[25]

To quantify divergence in the manner in which topics were covered, we use BERT (Devlin et al., 2019), a transformer that creates high-dimensional vector representations, or contextual embeddings, capturing documents' semantic meaning. After identifying the topic of each 512-token segment from the transcripts (the maximum length that BERT can process, corresponding to approximately 300 words) using the LDA procedure outlined above, we create a vector representation of the segment. For each topic, we calculate the L2 distance between all the contextual embeddings of segments on that topic and average to collapse to a program-by-topic-by-week measure of distance. Panel B of Appendix Figure A1 reports similarity between *Hannity* and the other programs (averaging across all 10 topics); *Tucker Carlson Tonight* is again the most similar program.[26]

Together, our evidence suggests that the two largest opinion shows in the United States adopted dramatically diverging narratives about the threat posed by COVID-19, despite presenting relatively similar content both before the COVID-19 pandemic and on non-COVID-19 topics during

[24] We provide a full list of topics in Appendix Table C1.

[25] Appendix Figure C1 reports program similarity for each topic individually. The topics where the two programs diverge the most are economic policy and Trump's impeachment; across most topics, differences are small. We also validate our measure by repeating this exercise for 2020, plotting results in Appendix Figure C3; as in our simpler word-counting exercise, we find large differences in COVID-19 coverage between *Tucker Carlson Tonight* and *Hannity* in January and February; these differences decrease to zero by early March. Differences in other topics, on the other hand, are much smaller.

[26] Appendix Figure C2 reports program similarity for each topic individually. Differences between *Tucker Carlson Tonight* and *Hannity* are again modest.

the pandemic. We next present survey evidence that these differences may have affected viewers' behavior during the period of initial spread of COVID-19 in the United States.

3.2 Timing of Behavioral Adjustment

Radical behavioral changes, such as stay-at-home behavior, did not become widespread until mid-to-late March, when the pandemic narrative gap between *Hannity* and *Tucker Carlson Tonight* had already closed.[27] To capture more subtle behavioral changes that may have occurred in February and March, and to shed light on which types of behavioral change were most common, we fielded a survey on April 3, 2020. Our survey targeted a representative sample of approximately 1,500 Republicans aged 55 or older both because this population is more likely to watch Fox News and because the elderly are at increased risk from COVID-19.[28] As we show in Appendix Table A1, our sample is broadly representative of Republicans aged 55 and older. All survey materials are available in Appendix F.

Survey design After eliciting demographics, we ask respondents which, if any, of the "Big Three" TV news stations (CNN, MSNBC, and Fox News) they watch at least once a week. 1,045 individuals reported that they watch any show on Fox News at least once a week; this is the sample we use in our analysis, given our focus on Fox News viewers. We ask respondents to indicate the frequency with which they watch the major prime-time shows on each network on a three-point scale ("never"; "occasionally"; "every day or most days").

We then ask our respondents about any changes in their behavior in response to COVID-19 outbreak. First, we ask whether they have changed any of their behaviors (e.g., cancelling travel plans, practicing social distancing, or washing hands more often) in response to COVID-19. For those respondents who answer that they have changed behavior, we elicit the date on which they did so. Finally, we ask an open-ended question asking respondents to describe which behaviors they changed.

Sample characteristics In Appendix Table A2, we plot demographic characteristics of exclusive *Tucker Carlson Tonight* and *Hannity* viewers. Although there are observable differences between the two groups of viewers, these differences appear to be modest.[29] In general, relatively few Fox

[27] See, e.g. "Social Distancing, but Mostly During the Workweek?" *Federal Reserve Bank of St. Louis*, May 26, 2020.

[28] The median age among Fox News viewers is 68. See, e.g. "Half of Fox News' Viewers Are 68 and Older." *The Atlantic*, January 27, 2014.

[29] As we discuss in Section 4, we conducted a larger-scale survey between December 2020 and January 2021, which included approximately 3,700 Fox viewers. Unlike our earlier survey, this survey was constructed to be representative of the US population as a whole, rather than only Republicans over the age of 55. As shown in Appendix Table A3, we again find that observable differences are modest. Of course, these results must be interpreted with caution, given that data were collected several months after the period we study.

viewers consume other sources of news, consistent with Pew survey data on viewership and on distrust of non-Fox media sources.[30]

Results To examine the correlation between viewership of different news shows and the timing of behavioral change, we estimate the following simple specification:

$$\text{TimingChange}_i = \alpha_0 + \beta S_i + \Pi X_i + \varepsilon_i,$$

where TimingChange_i is the number of days after February 1, 2020 on which the respondent reported having significantly changed any of their behaviors in response to COVID-19, S_i is a vector of indicators for whether the respondent occasionally or regularly watches each of the seven shows, and X_i is a vector of demographic controls.[31] The dependent variable for respondents who report that they have not changed any of their behaviors at the time of the survey is recoded to the date on which the survey was administered (April 3). We employ robust standard errors throughout our analysis.

Panel A of Figure 3 plots the smoothed density function of the reported date of behavioral change separately for viewers of Carlson, Hannity, and other Fox News shows. (The majority of viewers watch more than one show and thus appear in multiple panels.) We also display these results in regression table form in Table 2. Column 1 shows that viewers of *Hannity* changed their behavior four to five days later than viewers of other shows ($p < 0.001$), while viewers of *Tucker Carlson Tonight* changed their behavior three to four days earlier than viewers of other shows ($p < 0.01$); the difference in coefficients is also highly statistically significant ($p < 0.01$).[32] Column 2 reports a linear probability model in which the dependent variable is an indicator for whether the respondent reported changing behavior before March 1; Carlson viewers were 11.7 percentage points more likely and Hannity viewers 11.2 percentage points less likely to have changed their behavior before March 1 than viewers of other Fox shows.[33] We estimate identical linear probability models

[30] See "Five Facts about Fox News," *Pew Research*, April 8, 2020.

[31] Viewers who watch multiple shows will have multiple indicators set to one, while viewers that watch none of the five shows will have none of the indicators set to one.

[32] Ash et al. (2020) also find survey evidence that Republican *Hannity* viewers adopt social distancing measures significantly later than Republicans who do not watch *Hannity*, while Republican *Tucker Carlson Tonight* viewers adopt social distancing measures significantly earlier than Republicans who do not watch *Tucker Carlson Tonight.*

[33] To benchmark the plausibility of the estimated effects, we calculate the *persuasion rate* of viewership on the outcome of changing behavior by March 1, following the approach proposed by DellaVigna and Gentzkow (2010). The implied persuasion rate of *Hannity* viewership relative to *Tucker Carlson Tonight* viewership is 24.1 percent, well within the range of comparable estimates; for example, Martin and Yurukoglu (2017) find a Fox News persuasion rate on voting behavior of 58 percent in 2000, 27 percent in 2004, and 28 percent in 2008; Adena et al. (2015) find a persuasion rate of up to 36.8 percent; and Enikolopov et al. (2011) find persuasion rates rating from 7 to 66 percent. On one hand, we might expect a lower persuasion rate in our context because exposure is over a much shorter period; on the other hand, we might expect a higher persuasion rate (1) because the outcomes we study are arguably lower-stakes than the outcomes in other settings, (2) because viewers likely hold weak priors about the seriousness of the pandemic during the period under consideration, and (3) because regular viewers of a show likely place significant weight on the anchors' opinions.

for each day between February 1 and April 3 (the date on which we administered the survey) and report the coefficients on both *Hannity* viewership and *Tucker Carlson Tonight* viewership for each day in Panel B of Figure 3. By this measure, the difference between the two anchors peaks around March 1, then declines.

We also examine the timing of specific margins of behavioral adjustment by manually coding the open-ended responses to the question of which behaviors respondents changed. Appendix Figure A3 highlights that increased hand washing and physical distancing, including avoiding large events, are the most frequently mentioned behavioral changes, particularly in February, the period during which the differences in show content were largest. Cancelling travel plans and staying at home are also frequently mentioned, though primarily in mid and late March.[34]

Our survey suggests that show content may have affected individual behaviors relevant for the spread of COVID-19, shedding light on specific mechanisms that may explain the treatment effects on COVID-19 cases and deaths we document in Sections 3.3 and 3.4. However, the correlations might be driven by omitted variable bias or reverse causality: viewers who did not want to believe that the COVID-19 was a serious problem or viewers less inclined to changing their behavior may have selected into watching *Hannity*. Moreover, our outcome is self-reported, which may bias our estimates if respondents systematically misremember that they changed their behavior earlier or later than they actually did (and this tendency differs between *Hannity* and *Tucker Carlson Tonight* viewers). To address these issues, we turn to data on county-level COVID-19 cases and deaths, and later to an instrumental variable strategy shifting relative viewership of the two shows.

3.3 OLS Estimates on Health Outcomes

In this section, we discuss the empirical challenge in identifying causal effects. We then present OLS evidence on the effects of differential viewership of *Hannity* and *Tucker Carlson Tonight* on COVID-19 cases and deaths.

Data We employ several primary categories of data in our observational analysis, which we describe in detail in Appendix B. Our TV viewership data is provided by Nielsen at the Designated Market Area (DMA) level, of which there are 210 in the US. We focus on the continental United States, excluding the two DMAs in Alaska (Anchorage and Fairbanks) and the single DMA in Hawaii (Honolulu). Our dataset contains viewership data between 5pm and 11pm (local time) at

[34]The responses highlight the importance of distinguishing between two types of social distancing. Following the Federal Reserve, we distinguish *stay-at-home behavior* — remaining at home for all or a substantial part of the day — from *physical distancing* — continuing with day-to-day activities, but keeping a distance (e.g. of six feet) from others and avoiding large, potentially "superspreader" events such as sports games or concerts. While stay-at-home behavior becomes widespread only in mid-to-late March (see, e.g. Allcott et al. 2020b), our survey responses suggest that physical distancing and avoiding large events was widespread even in February among the population we survey.

the DMA-by-timeslot-by-day level (i.e. hourly ratings).[35] In addition to the fraction of TVs watching Fox News, we observe the fraction of TVs turned on during each timeslot. We supplement this dataset with 2018 data, previously acquired, on the local market share of each of the "Big Three" networks: CNN, MSNBC, and Fox News.[36] Our key outcome variables on county-level *confirmed* COVID-19 cases and deaths are drawn from Johns Hopkins University (Dong et al., 2020). Throughout our main analyses, we take the logarithm of one plus the cumulative number of cases and deaths, both to prevent outliers with a large number of cases from skewing the estimates and because the exponential nature by which a virus spreads makes the logarithm normalization natural. Finally, we compile a rich set of data on county level characteristics, including local vote shares, educational attainment, incomes, and the demographic age structure.

Data on COVID-19 cases are potentially subject to both classical and non-classical measurement error. For example, many COVID-19 cases are unreported (Lachmann, 2020; Stock et al., 2020), and if differential media coverage of the pandemic influences the rate of case detection, then our coefficient estimates will be biased. If viewers of *Hannity* are less concerned about the virus, and thus counties with greater viewership of *Hannity* have lower rates of case detection — this should bias our estimates *downward*. Classical measurement error will not bias our estimates, but will decrease their precision. Nonetheless, we urge caution in interpreting our estimated effects on cases given these potential data limitations. COVID-19 death counts are far less subject to either classical or non-classical measurement error.

OLS specification Our explanatory variable of interest is the DMA-level average difference between viewership of *Hannity* and viewership of *Tucker Carlson Tonight* across all days in January and February 2020 when both shows are aired. We scale this variable to take mean zero and standard deviation one for ease of interpretation. In our primary analysis, we estimate the following specification at the county level separately for each day between February 24 and April 15 (for cases) and between March 1 and April 15 (for deaths):

$$Y_{mct} = \alpha_t + \beta_t D_{mc} + \Pi_t X_{mc} + \varepsilon_{mct} \tag{1}$$

where Y_{mct} is an outcome (log one plus cases or log one plus deaths) in media market m, county c on day t, D_{mc} is the standardized difference between viewership shares of *Hannity* and *Tucker Carlson Tonight*, and X_{mc} is a vector of county-level controls.

[35] This is in contrast to previous work, e.g. Martin and Yurukoglu (2017), which uses data from 2005–2008 at the cable system-by-year level.

[36] Our primary analysis uses January and February viewership data; however, given the high degree of persistence in show viewership, our results are quantitatively extremely similar and qualitatively identical if we instead use only January data (to rule out concerns about reverse causality in our OLS estimates) or if we use data from January 1 through March 8 (the beginning of Daylight Savings Time, a natural stopping point given the structure of our identification strategy).

Identifying variation and potential confounders To see the potential threats to identifying causal effects, it is useful to understand where the variation in the main exposure variable, D_{mc}, comes from. By definition, it is the difference between the share of households that regularly watch *Hannity* ($v_{mc,H}$) and the share that regularly watch *Tucker Carlson Tonight* ($v_{mc,T}$). More broadly, for any show that airs at a certain hour-long time slot h in the evening, we can define the share of households that watch *any channel* on TV as $s_{mc,h}$ and, among those, the share at that moment that tunes in to Fox News as $f_{mc,h}$.

Thus, D_{mc} is driven by four factors:

$$D_{mc} = (s_{mc,H} \times f_{mc,H}) - (s_{mc,T} \times f_{mc,T})$$

This means that the OLS specification effectively exploits variation arising from differences in *timing preferences* and *channel preferences*:

$$Y_{mct} = \alpha_t + \beta_t(s_{mc,H} \times f_{mc,H} - s_{mc,T} \times f_{mc,T}) + \Pi_t X_{mc} + \varepsilon_{mct} \tag{2}$$

Since we are interested in examining the effects of differential exposure to two major shows on Fox News, Equation (2) makes it clear that if areas where Fox News is relatively popular experience more COVID-19 cases for any other (unobservable) reason — for example if populations in these areas live further away from high quality hospitals, tend to trust science less, or have certain life styles which make them more or less vulnerable to the virus — our estimate will be biased. To deal with this issue, we always control for the average evening TV market share of Fox News: $\overline{f}_{mc,h}$, where h denotes 8pm to 11pm Eastern Time. Moreover, since there may be selection into competing cable news networks specifically, rather than TV watching *per se*, we analogously always control for the "Big Three" cable TV market shares of Fox News and MSNBC (with CNN omitted since it is collinear with the other two). The inclusion of these controls hold fixed many potential confounders related to *channel preferences.*

Equation (2) also makes clear that if localities which have a tendency to watch evening TV *per se* around the time of *Hannity*, rather than *Tucker Carlson Tonight*, consist of populations which differ in their vulnerability to the virus, the OLS estimate could easily be biased. (Again, *ex ante* it is unclear which way the bias would go, given that we are comparing differential exposure to two shows on the same network.) To address concerns about local preferences for watching TV at certain times in the evening correlating with other determinants of COVID-19 trajectories — such as the extent to which people like to socialize in restaurants and bars (in ways which spread the virus) instead of staying home watching TV — we always include the average share of households with TVs turned on during each hourly slot between 8pm and 11pm Eastern Time (three variables, each capturing one hour): $s_{mc,8-9pm}$, $s_{mc,9-10pm}$, $s_{mc,10-11pm}$. These controls hold fixed many potential confounders related to *timing preferences.*

Given this approach, the remaining (residual) variation in exposure effectively comes from the difference in the two interaction terms of Equation (2), *holding constant* local preferences for watching TV in general and watching Fox News in general. We also include additional observable characteristics as controls. For example, since we study the early stages of the COVID-19 pandemic and initial outbreaks occurred around metropolitan hotspots, one concern may be that viewership patterns across the two shows correlate with such hot spot locations. For this reason, we show results with and without controls for rurality and population density and transparently show how much the estimate fluctuates as a result. More broadly, in addition to *population* controls, we show results with and without county-level controls for a range of observable characteristics: *race* (the share of the population white, Hispanic, and black); *education* (the share lacking high school degrees and the share lacking college degrees, for women and men separately); *age* (the share over the age of sixty-five); *economic* factors (the share under the federal poverty line, log median household income, and the unemployment rate); *health* factors (the share lacking health insurance and an age-adjusted measure of the average physical health in the county from 2018); *health capacity* (the number of different types of health personnel per capita); *political* factors (Republican vote share and the log total number of votes cast in the 2016 Presidential election).[37] To account for additional unobservable determinants of health outcomes that differ across localities, we show results using (1) no geographical fixed effects, (2) Census division (nine in total) fixed effects, and (3) state fixed effects. Since time zones are absorbed by the geographical indicator variables in the latter two cases, the fixed effects imply that we hold constant what time the two shows air locally. Our most extensive OLS specification – which is our preferred in that it helps rule out a host of concerns beyond the ones explicitly outlined above – includes state fixed effects and a full set of control variables.

To capture the effects in a transparent manner over time, we run separate cross-sectional regressions each day; in specifications including state fixed effects, this implicitly controls for state-level policies varying at the day level, such as shelter-in-place orders and closures of nonessential businesses. Because our viewership data is at the DMA level and to allow for within-market correlation in the error term, we cluster standard errors at the DMA level (m), resulting in a total of 204 clusters.[38]

Results We report day-by-day results for cases and deaths in Figure 4, including all controls and state fixed effects. The association between relative viewership and both cases and deaths becomes

[37] In Appendix Figure A4, we report regressions of an extensive range of county-level demographic characteristics, scaled to standard deviation one to facilitate interpretation, on our measure of relative viewership of the two shows. For none of these outcomes is our coefficient estimate statistically distinguishable from zero at the 5% level, though counties with a relative preference for *Hannity* are slightly more Hispanic and less black, have slightly fewer residents over the age of 65, and are somewhat less educated.

[38] Our results are also statistically significant if we instead cluster at the state level, as we show in Appendix Figure A5.

stronger over time until the coefficient on cases peaks in late March and then begins to decline; the coefficient on deaths follows with a two-week lag, consistent with the approximately two-to-three week lag between the appearance of COVID-19 symptoms and deaths (Wu et al., 2020). Effects on cases are statistically significant at the 5 percent level throughout the majority of the period, while effects on deaths are only statistically significant at the 5 percent level in late March and April. Effects on cases start to rise in late February and peak in mid-to-late March before starting to decline, consistent with the convergence in coronavirus coverage between Hannity and Carlson. A one standard deviation greater viewership difference is associated with 2 percent more deaths on March 21, 4 percent more deaths on March 28, and 9 percent more deaths on April 11. We report these results at weekly intervals in regression table form in Table 3.

Robustness To probe the robustness of our estimates, we choose a single day for cases — March 14, two weeks into March — and a single day for deaths — March 28, two weeks after our chosen date for cases (given the lag between cases and deaths). We then run our specifications under every possible combination of our nine sets of county-level controls (population density and rurality, race, age, economic, education, health, health capacity, politics) and three levels of fixed effects (no fixed effects, census division fixed effects, state fixed effects). Appendix Figure A6 reports coefficient estimates for each of these 768 models for cases as of March 14 as well as deaths as of March 28. The majority of coefficient estimates on cases and deaths are statistically significant at the 1 percent level. Almost all coefficient estimates from specifications including state fixed effects, our most demanding and most precisely estimated specifications, are significant at the 1 percent level. Moreover, our coefficient estimates are relatively stable.[39] Appendix Figure A7 shows a generally positive correlation between the R^2 of each model and the coefficient estimate, suggestive evidence that omitted variables downward bias our estimates. Indeed, a simple exercise to estimate omitted variables bias, following best practice recommendations from Oster (2019), suggests that the true effect may be several times larger.[40]

3.4 Instrumental Variables Estimates on Health Outcomes

We may remain concerned about factors driving both viewership preferences for *Hannity* over *Tucker Carlson Tonight* and COVID-19 outcomes. In this section, we describe our approach to gen-

[39] We repeat this exercise for every date between February 24 and April 15 for cases and between March 1 and April 15 for deaths. The resulting coefficient stability plots for each day are accessible at `https://raw.githubusercontent.com/AakaashRao/aakaashrao.github.io/master/files/ols-cases.gif` (cases) and `https://raw.githubusercontent.com/AakaashRao/aakaashrao.github.io/master/files/ols-deaths.gif` (deaths).

[40] The method requires assuming a maximum amount of variation that a hypothetical regression including all observable and unobservable covariates could explain; we follow the recommendation provided in Oster (2019) of using 1.3 times the R^2 value of the most extensive specification. The method also requires specifying the relative importance of observables and unobservables in explaining variation in the outcome variable; we again follow the guidance in Oster (2019) and assume observables and unobservables are equally important.

erate plausibly exogenous variation in relative viewership of *Hannity* over *Tucker Carlson Tonight*. As Equation (2) makes clear, the underlying variation in D_{mc} is driven by the combination of timing preferences and channel preferences. A lingering concern may be that these preferences are correlated with other unobservable determinants of COVID-19 outcomes. In particular, while the political slant of different shows on Fox News are similar and arguably cater the content towards viewers with similar beliefs and political viewpoints, the shows are not identical. Therefore, it could be that counties that favor *Hannity* over *Tucker Carlson Tonight* are somehow fundamentally different along dimensions that matter for health outcomes. Here, we alleviate some of these concerns by employing a leave-out approach, isolating cleaner variation that is less subject to confounders.

Leave-out IV The logic of the instrument is as follows. The OLS specification already flexibly controls for the tendency to watch TV at certain hours in the evening. If timing preferences are homogeneous across Fox and non-Fox viewers, timing preferences that determine health outcomes do not bias the OLS estimates. However, if timing preferences are heterogeneous across people that regularly watch Fox compared to those that prefer other channels, estimates may be biased. For example, if during the time *Hannity* airs, regular Fox viewers tend to prefer to stay home and watch TV while non-Fox viewers like to socialize in restaurants and bars (facilitating the spread of the virus), the OLS estimates would be (negatively) biased. To purge the treatment variable D_{mc} from any such variation, we isolate variation in timing preferences among only *non-Fox* viewers: $\tilde{s}_{mc,H}$, the average share of households that watch TV when *Hannity* airs, leaving out households that watch Fox News.

We use an analogous approach for channel preferences. The OLS estimations already control for the market share of Fox News, which may correlate with other determinants of health outcomes. Under the assumption that these other determinants do not also correlate with the *interaction* between channel preferences and timing preferences, the OLS estimates are unbiased. However, if regular Fox viewers that like to socialize in restaurants and bars prefer to watch TV slightly later in the evening when *Hannity* airs, whereas regular Fox viewers that seldom go to restaurants and bars stay home and watch TV earlier while *Tucker Carlson Tonight* is on, the OLS estimates would be (negatively) biased. To address this concern, we isolate variation in channel preferences during other timeslots outside of when *Hannity* and *Tucker Carlson Tonight* is live on air: $\tilde{f}_{mc,-HT}$, the average market share of Fox News, leaving out ratings during the 8-10pm Eastern Time.

Based on this logic, our leave-out instrument, Z_{mc}, consists of the interaction $\tilde{s}_{mc,H} \times \tilde{f}_{mc,-HT}$. The resulting first-stage regression is:

$$D_{mc} = \alpha + \beta_1 Z_{mc} + \beta_2 \tilde{s}_{mc,H} + \beta_3 \tilde{f}_{mc,-HT} + \Pi_t X_{mc} + \varepsilon_{mc} \tag{3}$$

Our identification strategy leverages distinct sources of identifying variation depending on the set of fixed effects that we include. In specifications without any geographic fixed effects, we exploit variation across time zones, thus exploiting variation in local airing time of the shows relative to the local "prime time" — the period in the evening where the number of TVs turned on peaks. For example, *Hannity* airs one hour after the prime time in EST, while it airs two hours before the prime time in PST. On the other hand, specifications with Census division and state fixed effects only exploit variation within a given time zone. Reassuringly, our coefficient estimates are relatively similar in magnitude across different choices of controls and fixed effects.

Correlation with pre-determined characteristics To illustrate the spatial distribution of the induced variation, Figure 5 plots our instrument values, residualized by the baseline controls in specification 3. In Appendix Figure A8, we report regressions using each county-level covariate as an outcome, scaled to a standard normal distribution to facilitate interpretation, on our instrument. No coefficients are significantly different from zero at the 5 percent level, and coefficient magnitudes are generally small.[41] Similarly, in Appendix Figure A9, we use data from the Cooperative Congressional Elections Study (Schaffner et al., 2021) to show that our instrument is not correlated with a wide range of policy preferences, including belief in anthropogenic climate change (a measure of trust in science) and attitudes toward women and minorities. None of the estimates are statistically significant, and the coefficient magnitudes are small. Taken together, this evidence helps mitigate concerns that the differences in the content of the two programs reflect a choice by Fox News or the individual anchors to tailor their programs toward differences in their audiences, as such differences would have to be latent and uncorrelated with any of the observables we examine. Nevertheless, as in the OLS approach, we show in a transparent manner the extent to which results are robust to permutations across all possible combinations of the groups of covariates.

One potential confound is that our effects reflect not differences in exposure to the diverging narratives presented on *Hannity* and *Tucker Carlson Tonight*, but rather heterogeneity in the overall persuasive effect of Fox News on partisanship across counties, which correlates with these counties' timing preferences for TV viewership. In this case, the effects we estimate would be driven not by differences in opinion content, but rather by differences in partisanship prior to the onset of the pandemic. Reassuringly, the instrument is uncorrelated with the 2016 county Republican vote share. A more subtle concern is that there exist differences in the latent persuasive potential of Fox News across counties that became relevant only during the pandemic. We see this contingency as unlikely, especially given the timing of effects we document in Section 3.5.

Exclusion restriction Our approach is motivated by the fact that (1) *Hannity* and *Tucker Carlson Tonight* are the most-viewed cable news programs in the United States, and by the fact that (2) the two shows conveyed very different narratives about the threat posed by COVID-19

[41] Due to wide confidence intervals, we cannot rule out high correlations with percentage Hispanic, percentage black, and log median household income.

at the early stages of the pandemic. In this sense, the instrumental variable approach is designed to shift exposure to different opinions through its effects on the relative viewership of these two programs. However, the assumption that all of the effects of the instrument on COVID-19 outcomes operate exclusively through differential exposure to *Hannity* over *Tucker Carlson Tonight* — the outcome variable in the first-stage regressions — requires that the instrument does not have any spillovers, negative or positive, onto viewership of other shows. This assumption would be violated if, for example, our instrument's effects on relative viewership of *Hannity* and *Tucker Carlson Tonight* induces viewers to change their consumption of other Fox News shows. Such spillovers could be very complex and may violate a narrow exclusion restriction, complicating interpretation of the two-stage least squares regressions. For these reasons, while we proceed in this section under the assumption that the exclusion restriction described above holds, in Appendix Section D.1, we relax this assumption to employ a more general approach allowing for arbitrary spillovers across evening Fox News programs, while still allowing us to investigate the hypothesized mechanism of exposure to differential narratives about COVID-19 crisis.

Instrument relevance As we show in Appendix Table A4, our instrument strongly predicts viewership of *Hannity* relative to *Tucker Carlson Tonight*. The first-stage coefficient estimates remain relatively constant over Census division and state fixed effects and as we include controls for population and population density, MSNBC's share of cable, and our rich set of county-level covariates: a one standard deviation higher value of the instrument is associated with approximately a one standard deviation higher viewership of *Hannity* relative to *Tucker Carlson Tonight* ($p < 0.001$), with somewhat tighter confidence intervals when fixed effects are included. As in the OLS specification, we cluster standard errors at the DMA level.[42]

Results on COVID-19 cases and deaths Figure 6, which for consistency and ease of comparison mirrors the OLS specification of Figure 4 (that is, the specification with the most extensive set of controls and fixed effects), shows the day-by-day 2SLS estimates of the effects of the standardized Hannity-Carlson viewership difference on cases and deaths. Effects on cases start to rise in early March and peak in mid-March before gradually declining, consistent with Hannity's changing position on COVID-19. Consistent with estimated lags between case and death reporting, effects on deaths start emerging approximately three weeks after cases.[43] A one standard deviation greater viewership of *Hannity* relative to *Tucker Carlson Tonight* is associated with 26 percent more deaths on March 28 ($p < 0.01$), 38 percent more deaths on April 4 ($p < 0.05$), and 33 percent more deaths on April 11 ($p < 0.10$).

The initial divergence and eventual plateauing of effects on COVID-19 outcomes are consistent with our proposed mechanism that differential reporting between Hannity and Carlson about

[42]The analogous results with standard errors clustered at the state level are reported in Appendix Figure A10.
[43]See, e.g., "A Second Coronavirus Death Surge is Coming." *The Atlantic*, July 15, 2020.

COVID-19 throughout February and early March are driving our results, as we explore in Section 3.5. We report 2SLS results at weekly intervals in regression table form in Table 3.[44]

Robustness to choice of specification As in Section 3.3, we run our specifications under every possible combination of our eight sets of county-level controls and our three levels of geographical fixed effects. Appendix Figure A6 reports coefficient estimates for each of these 768 models for cases as of March 14 and deaths as of March 28. Confidence intervals for models without any geographical fixed effects are wider due to unobservable variation in the outcome; once division or state fixed effects are included, the coefficients are relatively stable and tightly estimated. The majority of coefficient estimates on cases and deaths are statistically significant at the 1 percent level, as are all estimates drawn from specifications with state fixed effects included.[45]

The estimated OLS coefficients are generally increasing as we control for more observables, suggesting that unobservables generate a negative bias. In contrast, the 2SLS coefficient estimates are relatively stable across these same permutations of controls, suggesting less of a bias. The OLS estimates can thus be interpreted as a plausible lower bound on the true causal effect of differential viewership on COVID-19 trajectories.

One potential concern is that COVID-19 hotspots with large numbers of cases or deaths may skew our results. We probe robustness to outliers by residualizing our outcome variables and the instrument by our controls and fixed effects, then plotting the residuals of our outcome variables against the residuals of the instrument in Appendix Figure D1. As in the OLS estimates, neither plot gives cause for concern that our estimates are driven by outliers. To further ensure that counties with large number of cases or deaths are not driving our results, in Appendix Figure A11, we estimate our time series figures leaving out entire states containing prominent COVID-19 hotspots, leaving out the top 1% of counties by cases, and leaving out the highest-case county in each state. Point estimates decrease slightly, but remain statistically significant at the 5% level in all but the most demanding one (leaving out the entire states of CA, MA, NJ, NY, and WA), in which they are significant at the 10% level. We next replicate our estimates in Appendix Table A5 using Poisson and zero-inflated negative binomial regression. While these coefficients and associated standard errors must be interpreted with caution due to the high number of controls and the high-

[44]Consistent with the pattern that including richer controls increases OLS coefficient estimates, our estimated IV coefficients are larger than the estimated OLS coefficients. In addition to the possibility of measurement error in Nielsen's (sample-based) measure of viewership, this also may be due to the fact that our instrument, while uncorrelated with observables, varies across counties in the strength of its first stage. For example, the instrument has a larger first stage in counties with above-median age. Thus, the induced variation may lead our instrumental variables estimate to place greater weight (relative to the OLS estimate) on counties with higher case and death loads.

[45]We repeat this exercise for every date between February 24 and April 15 for cases and between March 1 and April 15 for deaths. Animations of the resulting coefficient stability plots for each day are accessible at `https://raw.githubusercontent.com/AakaashRao/aakaashrao.github.io/master/files/iv-cases.gif` (cases) and `https://raw.githubusercontent.com/AakaashRao/aakaashrao.github.io/master/files/iv-deaths.gif` (deaths).

dimensional fixed effects (Cameron and Trivedi, 2015), we again estimate statistically significant and relatively stable effects of our instrument on both cases and deaths.

In Appendix D, we carry out a number of exercises to probe the robustness of our results. In particular, we demonstrate that our estimates are not driven by zero values, we conduct several randomization exercises to assess the validity of our inference, and we show our estimates remain stable under two alternative instrumental variable strategies. In Appendix E, we assess the plausibility of our estimated magnitudes through the lens of an epidemiological model.

3.5 Mechanisms

Stay-at-home behavior Based on our survey results and the timing of differences between *Tucker Carlson Tonight* and *Hannity* — which had largely converged in their coverage by mid-March — we do not expect stay-at-home behavior to be a primary mechanism driving our results. Indeed, very few Americans had begun staying at home in response to the pandemic before mid-March (Allcott et al., 2020b), and our survey suggests that behavioral changes such as physical distancing (i.e. staying more than six feet apart from others and avoiding large events) were far more prominent. Nonetheless, we investigate the extent to which differential narratives affected stay-at-home behavior. We use smartphone GPS data from the Bureau of Transportation Statistics, which aggregates data "from merged multiple data sources that address the geographic and temporal sample variation issues often observed in a single data source," mitigating concerns about measurement error (Warren and Skillman, 2020). We use this data to create a panel at the day-by-county level tracking the number of devices that remain home throughout the day.[46] We then estimate our primary instrumental variables specification, using the share staying home in each county as the outcome. In addition to the controls above, we also control for the share of devices staying home on the same day in 2019 in order to increase the precision of our estimates, and we report one-week rolling means.

We report results in Appendix Figure A12, focusing on the period before state and local stay-at-home orders were implemented. Throughout most of February, our estimated effects of differential coverage on the fraction staying home are small and statistically indistinguishable from zero. We detect significant negative effects on stay-at-home behavior in the first two weeks of March, consistent with the gap in narratives presented on the two shows. We estimate relatively small effects (peaking at approximately 0.8 percentage points, or approximately 5 percent of the 2019 mean), consistent with stay-at-home behavior not being a primary mechanism driving our estimated treatment effects. Our results are not statistically significant if we use mobility data from the SafeGraph GPS panel rather than the BTS data; although the coefficient estimates are similar, the standard errors are much larger.

[46]See `https://www.bts.gov/browse-statistical-products-and-data/trips-distance/daily-travel-during-covid-19-pandemic`.

Timing of effects We now examine the timing of deaths and cases relative to the timing of differences in content of the two shows more closely. To construct a Carlson-Hannity "pandemic narrative gap," we use our coding results from Section 3.1: for each day, our index is defined as the difference between the average of the five ratings of the *Tucker Carlson Tonight* episode and the average of the five ratings of the *Hannity* episode on that day. Thus, higher values of the index indicate that the *Tucker Carlson Tonight* episode that aired on that day portrayed COVID-19 as a much more serious threat than the *Hannity* episode on the same day, while lower values of the index indicate that the two episodes were similar in their coverage. Second, to construct the Carlson-Hannity "behavioral change gap," we return to our survey results from Section 3.2: for each day, the gap is defined as the associated Hannity coefficient minus the same-day Carlson coefficient from Panel B of Figure 3 — that is, the difference between the marginal effects of viewership of these two shows on the event that the respondent had changed their behavior to act more cautiously in response to COVID-19 by the date in question. Thus, we should expect the behavioral change gap to lag the pandemic narrative gap, since viewers react to the differences in narratives presented on the two shows.

Figure 7 plots the pandemic narrative gap and the behavioral change gap. To facilitate plotting on the same figure, we rescale the pandemic coverage and behavioral change gaps by dividing each series' coefficients by the maximum coefficient value over the series, such that the maximum value is 1. Figure 7 also plots the (rescaled) 2SLS estimates of the effect of the Hannity-Carlson viewership gap (instrumented by Z_{mc}) on the fraction leaving home each day, log one plus cases, and log one plus deaths.

The pandemic narrative gap peaks in mid-February, a period during which there was no discussion of COVID-19 on *Hannity* and during which *Tucker Carlson Tonight* discussed the topic on virtually every episode, before declining to zero by mid-March. The behavioral change gap and gap in the share leaving home follow a similar shape with a two-week lag, peaking in early March before declining. The trend in coefficient estimates on cases closely mirrors the trend in the pandemic narrative gap (with a lag of approximately one month) and the trend on the behavioral change gap (with a lag of approximately two weeks), while the trend in coefficient estimates on deaths follows with an additional two week lag. These findings suggest that the effects of differential exposure to *Hannity* and *Tucker Carlson Tonight* that we document are *not* driven by longer-term past differential exposure to the shows or unobservable factors correlated both with the spread of the virus and preferences for one show over the other, but rather by differences in how the two shows covered the pandemic as it began to spread.

4 Discussion and Conclusion

Opinion programs represent a dominant and growing share of primetime cable television. Because they are less anchored in factual reporting, different opinion programs offer different, and often contradictory, narratives about reality. We examine the role of these narratives in shaping high-stakes decisions. Motivated by an experiment showing that people turn to opinion programs for information about objective facts, even in the presence of large incentives, we study the effects of the two most popular opinion programs in the United States — *Hannity* and *Tucker Carlson Tonight* — which diverged sharply in their narratives about the dangers posed by COVID-19 at the early stages of the pandemic. We validate these differences in content with independent coding of shows' transcripts and present survey evidence that, consistent with these content differences, viewers of *Hannity* changed behavior in response to the virus later than other Fox News viewers, while viewers of *Tucker Carlson Tonight* changed behavior earlier. Using both a selection-on-observables strategy with a rich set of controls and different instrumental variable strategies exploiting variation in the timing of TV viewership, we then document that greater exposure to *Hannity* relative to *Tucker Carlson Tonight* increased COVID-19 cases and deaths in the initial stages of COVID-19 pandemic.

Broader implications The initial period of the COVID-19 pandemic was characterized by substantial objective uncertainty, with little known about how the virus spreads, its medium and long-term effects, and the best measures to contain it. This uncertainty was reflected not only on cable and broadcast news, but also in the rapidly-changing public health recommendations (Rafkin et al., 2021). How might opinion-based news influence beliefs and behavior in contexts characterized by less objective uncertainty? On one hand, we might expect opinion-based news to be less influential in contexts where the facts are largely established. Yet, narratives featuring exaggerations and distortions of the truth — and viewers' propensity to take such narratives literally — appears widespread even in these contexts. For example, Tucker Carlson claimed in September 2020 that climate change is a "liberal invention"[47]; Sean Hannity repeatedly claimed that the ACA established "death panels" to decide whether individuals were worthy of health care[48]; and Rachel Maddow claimed that new abortion restrictions in Ohio would require women seeking abortions to undergo "mandatory vaginal probes"[49]. While causality remains unclear, all three contexts are characterized by substantial disagreement about objective facts.

Perhaps an even more striking such example is the dramatic growth of election conspiracism following the 2020 election and the historic Capitol Riot of January 6, 2021. On November 7, 2020, "straight news" anchors — including those on Fox News — declared Biden the winner of the election

[47]See "Tucker Carlson says climate change is a liberal invention like racism' in shocking on-air rant," *The Independent*, September 13, 2020.

[48]See "The return of ObamaCare's 'death panels"', *Hannity*, July 30, 2013.

[49]See "Rachel Maddow says that Ohio budget includes requirement for transvaginal ultrasound," *Cleveland.com*, July 9, 2013.

and, over the course of the next several weeks, pushed back against the narrative that Democrats had stolen the election.[50] In contrast, opinion anchors continued to question or outright challenge the election's legitimacy, giving outsized weight to conspiracy theories and personal anecdotes from observers at the expense of unambiguous statements from election officials and judges.[51] To shed light on the relationship between the diverging narratives on opinion vs. straight news shows on Fox News, we conducted a large-scale representative survey ($n = 13,744$) in which we elicited rich data on people's news consumption as well as their beliefs about election fraud. The results highlight robust correlations between viewership of opinion programs and beliefs about election fraud (see Figure 8). Table 4 shows that controlling for a rich set of observable characteristics, Fox News viewers who regularly watched opinion shows, relative to Fox News viewers who did not watch opinion shows, were 17 percentage points more likely to believe that Trump had received more votes than Biden, 19 percentage points more likely to believe that voting machines had switched votes from Trump to Biden, and 9 percentage points more likely to believe that Trump would be inaugurated on January 20.[52] Consistent with our findings, the social media feeds of participants in the January 6 Capitol Riot (including Ashli Babbitt, the woman fatally shot while attempting to breach the Speaker's Lobby) disproportionately feature content from these anchors.[53]

Neither the examples of narratives about climate change, the Affordable Care Act, or abortion restrictions nor our survey can establish causality. The evidence suggests, however, that opinion-based news may be important not only in contexts with high objective uncertainty, such as the COVID-19 pandemic, but also in contexts where the facts are largely established.

Open areas for research Our paper suggests several directions for future research. While we study the effects of *short-run*, contemporaneous exposure to diverging narratives on opinion programs, equally important are the effects of *long-run* exposure to opinion programming and other media untethered from factual reality. Such content undermines the role of expertise — e.g. that of climate scientists, election administrators, or public health officials — and promotes subjective commentary and personal reactions over factual reporting. Particularly given the dramatic growth

[50] See "The Moment Fox News Called the Election and Ended the Trump Love Affair." *The Independent*, November 8, 2020. "Bret Baier Corners Josh Hawley About Contesting Election, Makes Senator Squirm." *The Wrap*, January 4, 2021.

[51] See "Tucker Carlson Claims Virtually Every Power Center on Earth Rigged the Election for Joe Biden." *Media Matters for America*, January 4, 2021. "Opinion: Sean Hannity, Americas No. 2 Threat to Democracy: An A-to-Z guide." *The Washington Post*, December 14, 2020. The Fox News Decision Desk was in fact the first major outlet to project Arizona, a state crucial for Trump's reelection chances, for Biden. Faced with criticism from opinion anchors Sean Hannity and Tucker Carlson, Decision Desk director Arnon Mishkin defended his team's call, stating "The primetime schedule at Fox is the opinion part of Fox. And, you know, the great thing about opinion is that everyone can have an opinion." See "The Man Behind the Fox News Decision Desk." *The Dispatch*, November 6, 2020.

[52] Appendix Figure A13 presents a coefficient stability plot demonstrating the robustness of these correlations to different combinations of demographic controls, including the strength of partisan identification.

[53] See "Fox Settled a Lawsuit Over Its Lies. But It Insisted on One Unusual Condition." *The New York Times*, January 17, 2021. "After Deadly Capitol Riot, Fox News Stays Silent On Stars' Incendiary Rhetoric." *National Public Radio*, January 13, 2021.

in opinion programming at the expense of straight news reporting, this trend may fuel conspiracism, disagreement about objective facts, and affective and belief polarization. Empirical work that is able to identify these long-run effects — and more fundamentally, greater behavioral evidence on the determinants of trust in subjective vs. objective statements, including the role of personal experiences and reference points — would be particularly valuable.

References

Acemoglu, Daron, Victor Chernozhukov, Ivan Werning, and Michael D Whinston, "Optimal Targeted Lockdowns in a Multi-Group SIR Model," Working Paper 27102, National Bureau of Economic Research May 2020. Series: Working Paper Series.

Adena, Maja, Ruben Enikolopov, Maria Petrova, Veronica Santarosa, and Ekaterina Zhuravskaya, "Radio and the Rise of The Nazis in Prewar Germany," *The Quarterly Journal of Economics*, November 2015, *130* (4), 1885–1939.

Akbarpour, Mohammad, Cody Cook, Aude Marzuoli, Simon Mongey, Abhishek Nagaraj, Matteo Saccarola, Pietro Tebaldi, Shoshana Vasserman, and Hanbin Yang, "Socioeconomic Network Heterogeneity and Pandemic Policy Response," SSRN Scholarly Paper ID 3623593, Social Science Research Network, Rochester, NY June 2020.

Allcott, Hunt, Levi Boxell, Jacob Conway, Billy Ferguson, Matthew Gentzkow, and Benny Goldman, "Economic and Health Impacts of Social Distancing Policies during the Coronavirus Pandemic," SSRN Scholarly Paper ID 3610422, Social Science Research Network, Rochester, NY May 2020.

_ , _ , _ , Matthew Gentzkow, Michael Thaler, and David Yang, "Polarization and public health: Partisan differences in social distancing during COVID-19," *Working Paper*, 2020.

Ananyev, Maxim, Michael Poyker, and Yuan Tian, "The Safest Time to Fly: Pandemic Response in the Era of Fox News," Technical Report August 2020.

Andrews, Isaiah, James H. Stock, and Liyang Sun, "Weak Instruments in Instrumental Variables Regression: Theory and Practice," *Annual Review of Economics*, 2019, *11* (1), 727–753. _eprint: https://doi.org/10.1146/annurev-economics-080218-025643.

Ash, Elliott, Sergio Galletta, Dominik Hangartner, Yotam Margalit, and Matteo Pinna, "The Effect of Fox News on Health Behavior During COVID-19," SSRN Scholarly Paper ID 3636762, Social Science Research Network, Rochester, NY June 2020.

Athey, S. and G. W. Imbens, "Chapter 3 - The Econometrics of Randomized Experimentsa," in Abhijit Vinayak Banerjee and Esther Duflo, eds., *Handbook of Economic Field Experiments*, Vol. 1 of *Handbook of Field Experiments*, North-Holland, January 2017, pp. 73–140.

Atkeson, Andrew, Karen Kopecky, and Tao Zha, "Estimating and Forecasting Disease Scenarios for COVID-19 with an SIR Model," Working Paper 27335, National Bureau of Economic Research June 2020. Series: Working Paper Series.

Bai, Nina, "The Science Behind How Face Masks Prevent Coronavirus," June 2020.

Banerjee, Abhijit, Eliana La Ferrara, and Victor H Orozco-Olvera, "The Entertaining Way to Behavioral Change: Fighting HIV with MTV," Working Paper 26096, National Bureau of Economic Research July 2019.

__ , **Marcella Alsan, Emily Breza, Arun G Chandrasekhar, Abhijit Chowdhury, Esther Duflo, Paul Goldsmith-Pinkham, and Benjamin A Olken**, "Messages on COVID-19 Prevention in India Increased Symptoms Reporting and Adherence to Preventive Behaviors Among 25 Million Recipients with Similar Effects on Non-recipient Members of Their Communities," Working Paper 27496, National Bureau of Economic Research July 2020. Series: Working Paper Series.

Barnett, Michael L. and David C. Grabowski, "Nursing Homes Are Ground Zero for COVID-19 Pandemic," *JAMA Health Forum*, March 2020, *1* (3), e200369.

Bartik, Timothy J, "Who benefits from state and local economic development policies?," 1991. tex.publisher: WE Upjohn Institute for Employment Research.

Bennett, W. Lance, *News: The Politics of Illusion, Tenth Edition*, Chicago, IL: University of Chicago Press, September 2016.

Berry, Jeffrey M and Sarah Sobieraj, *The outrage industry: Political opinion media and the new incivility*, Oxford University Press, 2013.

Blei, David M., Andrew Y. Ng, and Michael I. Jordan, "Latent Dirichlet Allocation," *Journal of Machine Learning Research*, 2003, *3* (Jan), 993–1022.

Bootsma, Martin CJ and Neil M Ferguson, "The effect of public health measures on the 1918 influenza pandemic in US cities," *Proceedings of the National Academy of Sciences*, 2007, *104* (18), 7588–7593. tex.publisher: National Acad Sciences.

Bordalo, Pedro, Marco Tabellini, and David Y. Yang, "Issue Salience and Political Stereotypes," SSRN Scholarly Paper ID 3581750, Social Science Research Network, Rochester, NY January 2021.

Budak, Ceren, Ashley Muddiman, Yujin Kim, Caroline C. Murray, and Natalie J. Stroud, "COVID-19 Coverage By Cable and Broadcast Networks," *Proceedings of the International AAAI Conference on Web and Social Media*, May 2021, *15*, 952–960.

Bursztyn, Leonardo, Georgy Egorov, Ingar Haaland, Aakaash Rao, and Christopher Roth, "Justifying Dissent," *University of Chicago, Becker Friedman Institute for Economics Working Paper*, 2022, (2020-73).

__ , __ , **Ruben Enikolopov, and Maria Petrova**, "Social media and xenophobia: evidence from Russia," Working Paper 26567, National Bureau of Economic Research December 2019.

Cameron, A. Colin and Pravin K. Trivedi, "Count Panel Data," January 2015. ISBN: 9780199940042.

Davidson, Patricia M. and Sarah L. Szanton, "Nursing homes and COVID19: We can and should do better," *Journal of Clinical Nursing*, May 2020, p. 10.1111/jocn.15297.

DellaVigna, Stefano and Eliana La Ferrara, "Chapter 19 - Economic and Social Impacts of the Media," in Simon P. Anderson, Joel Waldfogel, and David Strömberg, eds., *Handbook of Media Economics*, Vol. 1 of *Handbook of Media Economics*, North-Holland, January 2016, pp. 723–768.

_ **and Ethan Kaplan**, "The fox news effect: Media bias and voting," *The Quarterly Journal of Economics*, 2007, *122* (3), 1187–1234. tex.publisher: MIT Press.

_ **and Matthew Gentzkow**, "Persuasion: Empirical Evidence," *Annual Review of Economics*, 2010, *2* (1), 643–669.

Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," *arXiv:1810.04805 [cs]*, May 2019. arXiv: 1810.04805.

Djourelova, Milena, "Media persuasion through slanted language: Evidence from the coverage of immigration," 2020.

Dong, Ensheng, Hongru Du, and Lauren Gardner, "An interactive web-based dashboard to track COVID-19 in real time," *The Lancet Infectious Diseases*, 2020. tex.booktitle: The lancet infectious diseases tex.date-added: 2020-03-06 15:00:59 +0000 tex.date-modified: 2020-03-06 15:00:59 +0000 tex.isbn: 1473-3099 tex.m3: doi: 10.1016/S1473-3099(20)30120-1 tex.publisher: Elsevier tex.ty: JOUR.

Draca, Mirko and Carlo Schwarz, "How Polarized are Citizens? Measuring Ideology from the Ground-Up," SSRN Scholarly Paper ID 3154431, Social Science Research Network, Rochester, NY May 2021.

Durante, Ruben and Ekaterina Zhuravskaya, "Attack when the world is not watching? US news and the Israeli-Palestinian conflict," *Journal of Political Economy*, 2018, *126* (3), 1085–1133. tex.publisher: University of Chicago Press Chicago, IL.

_ **, Paolo Pinotti, and Andrea Tesei**, "The political legacy of entertainment tv," *American Economic Review*, 2019, *109* (7), 2497–2530.

Dyer, Travis, Mark Lang, and Lorien Stice-Lawrence, "The evolution of 10-K textual disclosure: Evidence from Latent Dirichlet Allocation," *Journal of Accounting and Economics*, November 2017, *64* (2), 221–245.

Edison, Hali and Hector Carcel, "Text data analysis using Latent Dirichlet Allocation: an application to FOMC transcripts," *Applied Economics Letters*, January 2021, *28* (1), 38–42. Publisher: Routledge _eprint: https://doi.org/10.1080/13504851.2020.1730748.

Eisensee, Thomas and David Strömberg, "News droughts, news floods, and US disaster relief," *The Quarterly Journal of Economics*, 2007, *122* (2), 693–728. tex.publisher: MIT Press.

Enikolopov, Ruben, Maria Petrova, and Ekaterina Zhuravskaya, "Media and Political Persuasion: Evidence from Russia," *American Economic Review*, December 2011, *101* (7), 3253–3285.

Ferrara, Eliana La, "Mass media and social change: Can we use television to fight poverty?," *Journal of the European Economic Association*, July 2016, *14* (4), 791–827. tex.publisher: Oxford University Press.

— , **Alberto Chong, and Suzanne Duryea**, "Soap operas and fertility: Evidence from Brazil," *American Economic Journal: Applied Economics*, October 2012, *4* (4), 1–31.

Gallup, "Americans' Trust in Media Dips to Second Lowest on Record," October 2021. Section: Politics.

Gentzkow, Matthew and Jesse M. Shapiro, "What Drives Media Slant? Evidence From U.S. Daily Newspapers," *Econometrica*, 2010, *78* (1), 35–71. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA7195.

Jacobs, Ronald N and Eleanor Townsley, *The space of opinion: Media intellectuals and the public sphere*, Oxford University Press, 2011.

Kavanagh, Jennifer and Michael Rich, *Truth decay: An initial exploration of the diminishing role of facts and analysis in American public life*, Rand Corporation, 2018.

— , **William Marcellino, Jonathan S. Blake, Shawn Smith, Steven Davenport, and Mahlet G. Tebeka**, *News in a Digital Age: Comparing the Presentation of News Information over Time and Across Media Platforms*, Santa Monica, CA: RAND Corporation, 2019.

Kearney, Melissa S and Phillip B Levine, "Media influences on social outcomes: The impact of MTV's 16 and pregnant on teen childbearing," *American Economic Review*, 2015, *105* (12), 3597–3632.

Kuchler, Theresa, Dominic Russel, and Johannes Stroebel, "The Geographic Spread of COVID-19 Correlates with Structure of Social Networks as Measured by Facebook," Technical Report w26990, National Bureau of Economic Research, Cambridge, MA April 2020.

Lachmann, Alexander, "Correcting under-reported COVID-19 case numbers," *medRxiv*, 2020. tex.publisher: Cold Spring Harbor Laboratory Press.

Landis, J. Richard and Gary G. Koch, "The Measurement of Observer Agreement for Categorical Data," *Biometrics*, March 1977, *33* (1), 159.

Levy, Ro'ee, "Social media, news consumption, and polarization: Evidence from a field experiment," *American economic review*, 2021, *111* (3), 831–70.

Markel, Howard, Harvey B Lipman, J Alexander Navarro, Alexandra Sloan, Joseph R Michalsen, Alexandra Minna Stern, and Martin S Cetron, "Nonpharmaceutical interventions implemented by US cities during the 1918-1919 influenza pandemic," *Jama*, 2007, *298* (6), 644–654. tex.publisher: American Medical Association.

Martin, Gregory J. and Ali Yurukoglu, "Bias in Cable News: Persuasion and Polarization," *American Economic Review*, September 2017, *107* (9), 2565–2599.

Martinez-Bravo, Monica and Andreas Stegmann, "In vaccines we trust? The effect of anti-vaccine propaganda on immunization: Evidence from pakistan," *Journal of the European Economic Associations*, 2021.

McGee, Ryan Seamus, "seirsplus," May 2020. original-date: 2020-03-20T02:06:45Z.

Monod, Mlodie, Alexandra Blenkinsop, Xiaoyue Xi, Daniel Hebert, Sivan Bershan, Simon Tietze, Marc Baguelin, Valerie C. Bradley, Yu Chen, Helen Coupland, Sarah Filippi, Jonathan Ish-Horowicz, Martin McManus, Thomas Mellan, Axel Gandy, Michael Hutchinson, H. Juliette T. Unwin, Sabine L. van Elsland, Michaela A. C. Vollmer, Sebastian Weber, Harrison Zhu, Anne Bezancon, Neil M. Ferguson, Swapnil Mishra, Seth Flaxman, Samir Bhatt, Oliver Ratmann, and on behalf of the Imperial College COVID-19 Response Team, "Age groups that sustain resurging COVID-19 epidemics in the United States," *Science*, March 2021, *371* (6536), eabe8372.

Muller, Karsten and Carlo Schwarz, "Making America Hate Again? Twitter and Hate Crime Under Trump," Working Paper 3149103, Social Science Research Network March 2018.

Oster, Emily, "Unobservable Selection and Coefficient Stability: Theory and Evidence," *Journal of Business & Economic Statistics*, April 2019, *37* (2), 187–204.

Pei, Sen, Sasikiran Kandula, and Jeffrey Shaman, "Differential Effects of Intervention Timing on COVID-19 Spread in the United States," *medRxiv*, May 2020, p. 2020.05.15.20103655. Publisher: Cold Spring Harbor Laboratory Press.

Pew, "Demographics and political views of news audiences," *Pew Research Center's Global Attitudes Project*, March 2019.

Rafkin, Charlie, Advik Shreekumar, and Pierre-Luc Vautrey, "When guidance changes: Government stances and public beliefs," *Journal of Public Economics*, April 2021, *196*, 104319.

Rogers, Anna, Olga Kovaleva, and Anna Rumshisky, "A Primer in BERTology: What We Know About How BERT Works," *Transactions of the Association for Computational Linguistics*, 2020, *8*, 842–866. Place: Cambridge, MA Publisher: MIT Press.

Schaffner, Brian, Stephen Ansolabehere, and Sam Luks, "Cooperative Election Study Common Content, 2020," August 2021. Type: dataset.

Schwarz, Carlo, "Ldagibbs: A Command for Topic Modeling in Stata Using Latent Dirichlet Allocation," *The Stata Journal: Promoting communications on statistics and Stata*, March 2018, *18* (1), 101–117.

Shadmehr, Mehdi and Ethan Bueno de Mesquita, "Coordination and Social Distancing: Inertia in the Aggregate Response to COVID-19," SSRN Scholarly Paper ID 3586839, Social Science Research Network, Rochester, NY April 2020.

Simonov, Andrey, Szymon Sacher, Jean-Pierre Dubé, and Shirsho Biswas, "The Persuasive Effect of Fox News: Non-Compliance with Social Distancing During the Covid-19 Pandemic," SSRN Scholarly Paper ID 3600088, Social Science Research Network, Rochester, NY May 2020.

Stock, James H., Karl M. Aspelund, Michael Droste, and Christopher D. Walker, "Estimates of the Undetected Rate among the SARS-CoV-2 Infected using Testing Data from Iceland," *medRxiv*, April 2020, p. 2020.04.06.20055582. Publisher: Cold Spring Harbor Laboratory Press.

Unwin, H, S Mishra, VC Bradley, A Gandy, M Vollmer, T Mellan, H Coupland, K Ainslie, C Whittaker, J Ish-Horowicz, S Filippi, X Xi, M Monod, O Ratmann, M Hutchinson, F Valka, H Zhu, I Hawryluk, P Milton, M Baguelin, A Boonyasiri, N Brazeau, L Cattarino, G Charles, L Cooper, Z Cucunuba Perez, G Cuomo-Dannenburg, A Djaafara, I Dorigatti, O Eales, J Eaton, S Van Elsland, R Fitzjohn, K Gaythorpe, W Green, T Hallett, W Hinsley, N Imai, B Jeffrey, E Knock, D Laydon, J Lees, G Nedjati Gilani, P Nouvellet, L Okell, A Ower, K Parag, I Siveroni, H Thompson, R Verity, P Walker, C Walters, Y Wang, O Watson, L Whittles, A Ghani, N Ferguson, S Riley, C Donnelly, S Bhatt, and S Flaxman, "Report 23: State-level tracking of COVID-19 in the United States," Technical Report, Imperial College London May 2020.

Warren, Michael S. and Samuel W. Skillman, "Mobility Changes in Response to COVID-19," *arXiv:2003.14228 [cs]*, March 2020. arXiv: 2003.14228.

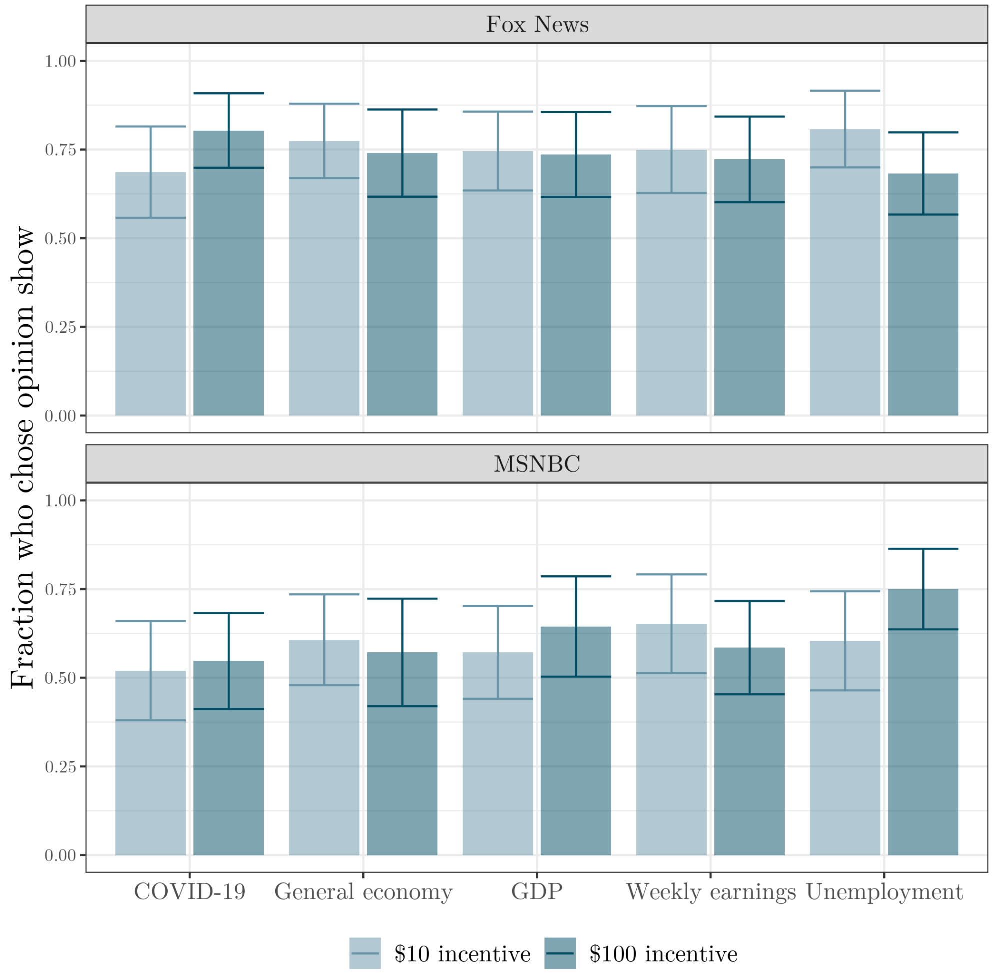
Wu, Joseph T., Kathy Leung, Mary Bushman, Nishant Kishore, Rene Niehus, Pablo M. de Salazar, Benjamin J. Cowling, Marc Lipsitch, and Gabriel M. Leung, "Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China," *Nature Medicine*, 2020.

Yan, Li-Meng, Shu Kang, Jie Guan, and Shanchang Hu, "SARS-CoV-2 Is an Unrestricted Bioweapon: A Truth Revealed through Uncovering a Large-Scale, Organized Scientific Fraud," Working Paper October 2020.

Yanagizawa-Drott, David, "Propaganda and Conflict: Evidence from the Rwandan Genocide," *The Quarterly Journal of Economics*, November 2014, *129* (4), 1947–1994.

Figures

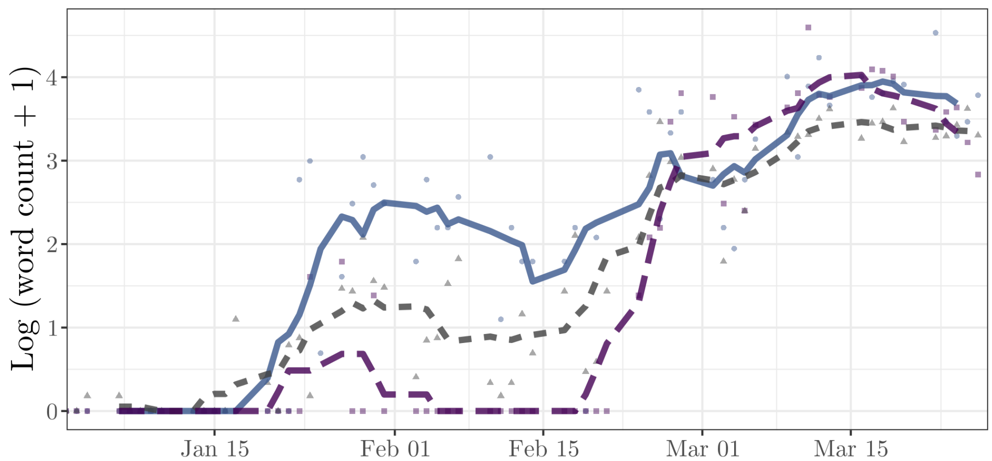
Figure 1: Trust in opinions



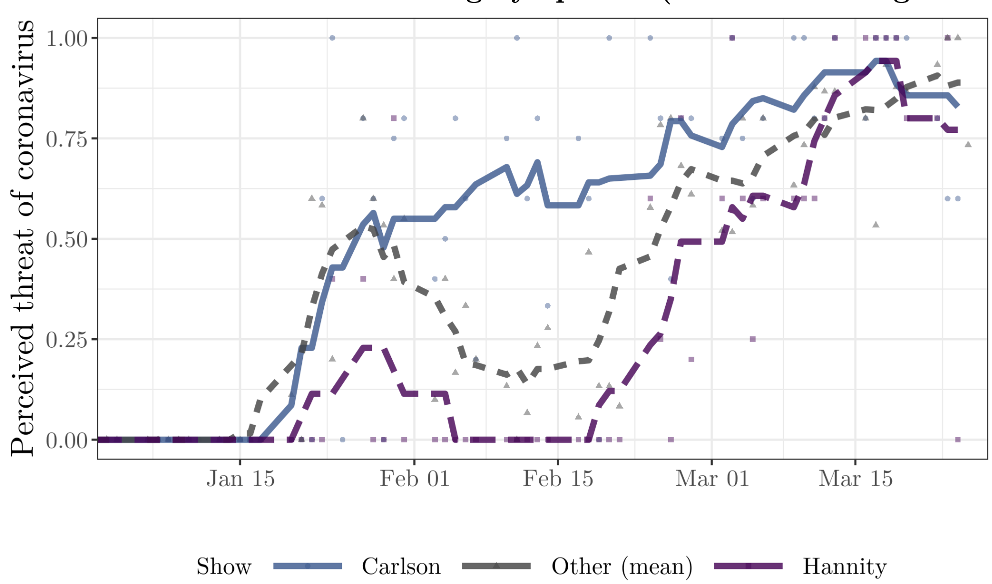
Notes: Figure 1 plots the fraction of respondents who choose an opinion program over a straight news program, separately by network, domain of question, and level of incentive for accuracy. 95% confidence intervals based upon robust standard errors are reported.

Figure 2: Show content validation

Panel A: Counts of coronavirus-related terms by episode (one-week rolling means)

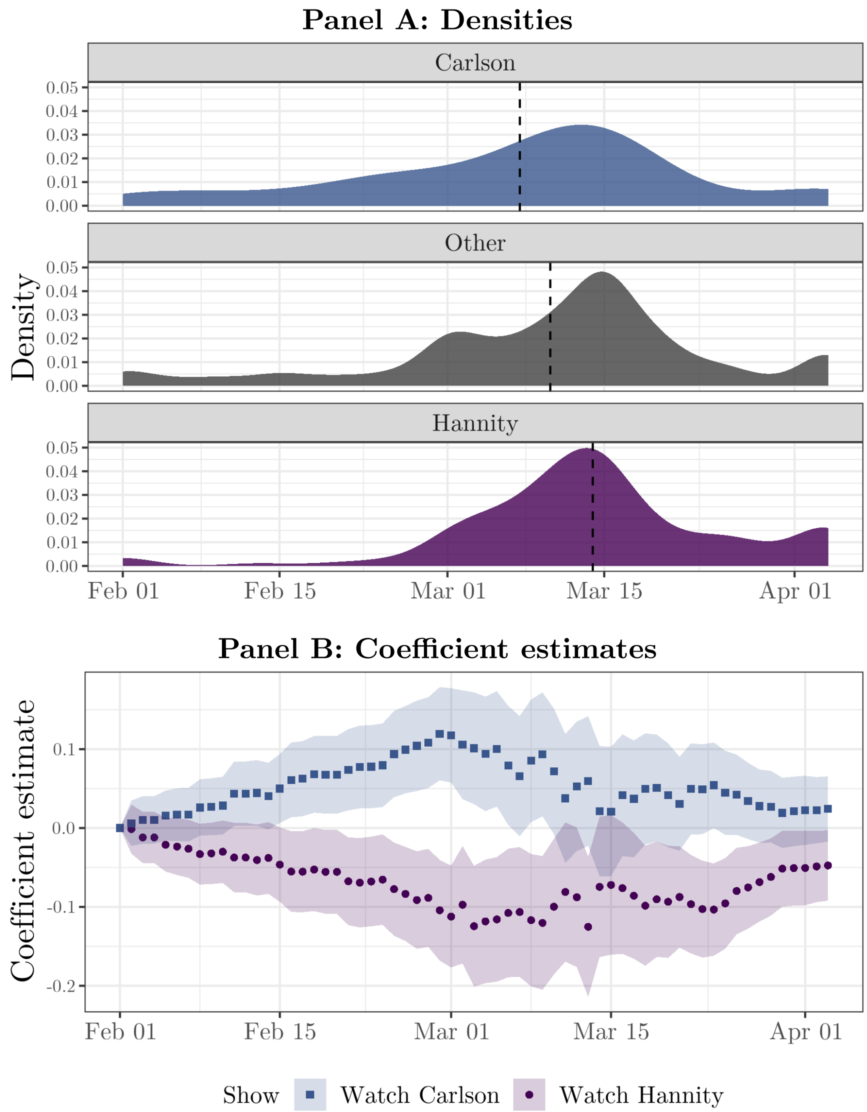


Panel B: MTurk seriousness rating by episode (one-week rolling means)



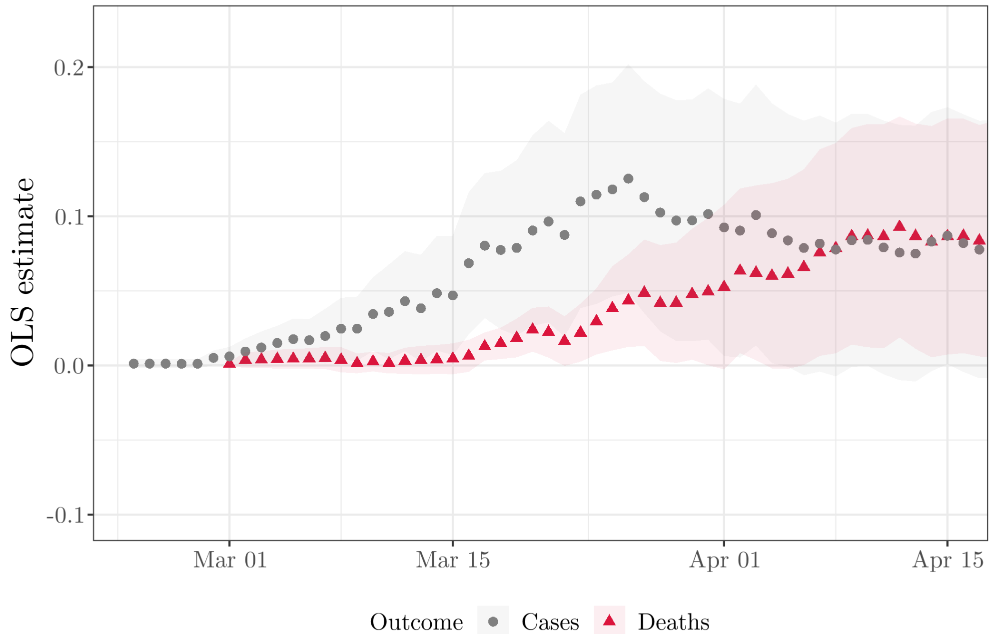
Notes: Panel A shows counts of coronavirus-related terms (coronavirus, COVID, virus, influenza, and flu) separately for *Hannity*, *Tucker Carlson Tonight*, and the other Fox News shows aired on Fox News between 5pm and 11pm local time across all four major time zones in the continental US (*The Five*, *Special Report with Bret Baier*, *The Story with Martha MacCallum*, *Fox News at Night*, and *The Ingraham Angle*). Panel B shows the seriousness rating for each episode, constructed as an average of Amazon Mechanical Turk ratings. For each show containing at least one coronavirus-related term, five MTurk workers read the entire script and answered "Yes" or "No" to the following question: "Did [the show] indicate that the virus is likely to infect many people in the US, causing many deaths or serious illnesses, or that many have already become infected and have died or become seriously ill?" We impute "No" for each episode that does not mention any coronavirus-related terms and recode "Yes" to 1 and "No" to 0.

Figure 3: Timing of behavioral change by show viewership



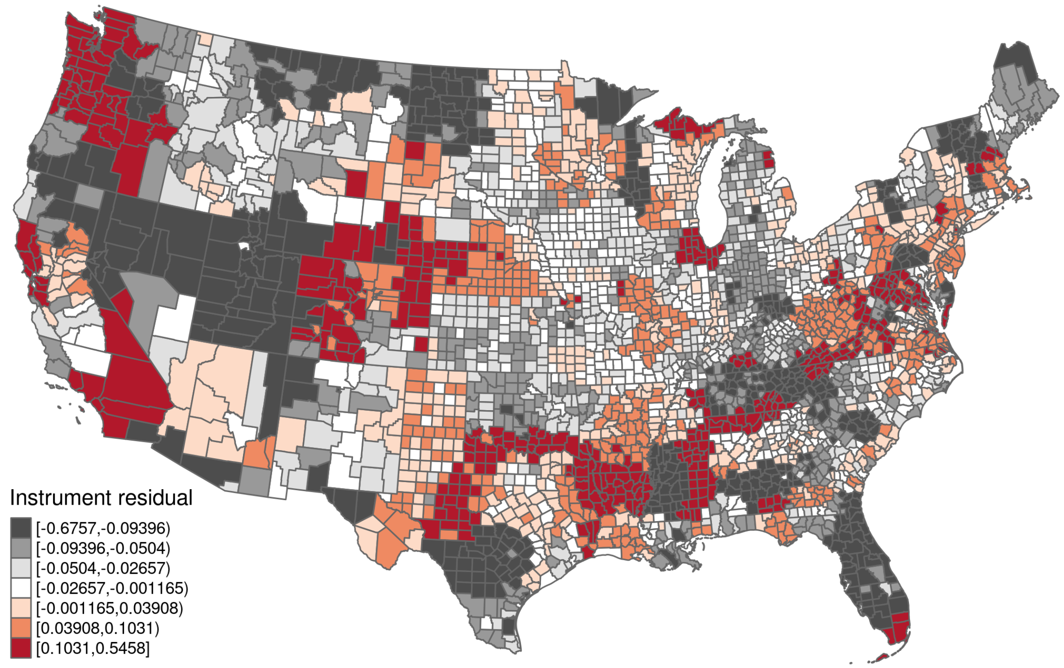
Notes: Panel A of Figure 3 displays the density function of viewers' reported day of behavior change in response to the coronavirus. For respondents who report that they have not changed any of their behaviors by the date of the survey, we impute the date of the survey (April 3). The dashed line indicates the mean date of behavior change among viewers of each show. To mirror our regressions, the top pane includes only *Tucker Carlson Tonight* viewers that do not watch *Hannity*, while the bottom pane includes only *Hannity* viewers that do not watch *Tucker Carlson Tonight*. Panel B reports coefficient estimates from linear probability models in which the dependent variable is an indicator for whether the respondent reported changing behavior before the date in question and the explanatory variables include an indicator for whether the respondent watches *Tucker Carlson Tonight*, an indicator for whether the respondent watches *Hannity*, an indicator for whether the respondent watches any other Fox News shows, and controls for gender, employment status, income, race, education, and viewership of CNN and MSNBC. We report 95% confidence intervals.

Figure 4: OLS estimates of effect of differential viewership on cases and deaths



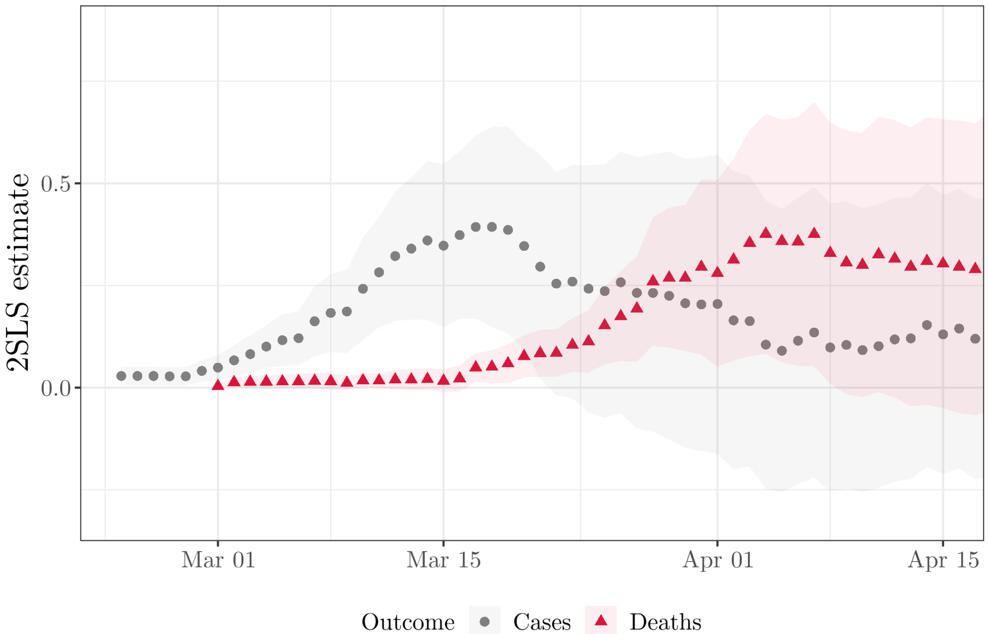
Notes: Figure 4 displays effects of differential viewership of *Hannity* and *Tucker Carlson Tonight* on log one plus cases and log one plus deaths. We report day-by-day results for the correlation between log deaths and log cases with the standardized viewership difference between *Hannity* and *Tucker Carlson Tonight*. All regressions are conditional on state fixed effects and a large set of controls: the November 2018 and January 2020 market share of Fox News, the November 2018 market share of MSNBC, log total population, population density, the share of TVs turned to non-Fox channels during *Hannity*, *Tucker Carlson Tonight*, and *The Ingraham Angle*, the population-weighted latitude and longitude, the percent in the county living in rural areas, the percent white, Hispanic, and black, the percent over the age of sixty-five, the share of men and women lacking high school degrees, the share of men and women lacking college degrees, the fraction of the population lacking health insurance, an age-adjusted measure of the average physical health in the county from 2018, the percent under the federal poverty line, log median household income, the unemployment rate, the 2016 Republican vote share, and the log total number of votes cast in 2016. We cluster standard errors at the DMA level and report 95 percent confidence intervals.

Figure 5: Residualized Hannity-Carlson instrument values



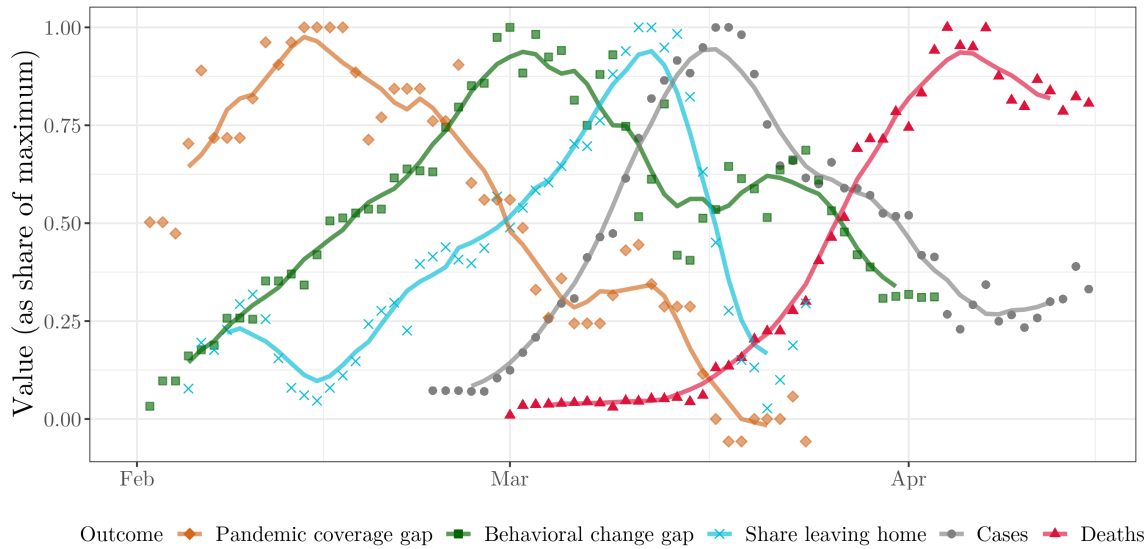
Notes: Figure 5 plots the values of our instrument, $\tilde{s}_{mc,H} \times \tilde{f}_{mc,-HT}$, residualized by our full set of controls: the November 2018 and January 2020 market share of Fox News, the November 2018 market share of MSNBC, log total population, population density, the predicted share of TVs turned to non-Fox channels during *Hannity*, *Tucker Carlson Tonight*, and *The Ingraham Angle*, the population-weighted latitude and longitude, the percent in the county living in rural areas, the percent white, Hispanic, and black, the percent over the age of sixty-five, the share of men and women lacking high school degrees, the share of men and women lacking college degrees, the fraction of the population lacking health insurance, an age-adjusted measure of the average physical health in the county from 2018, the percent under the federal poverty line, log median household income, the unemployment rate, the 2016 Republican vote share, and the log total number of votes cast in 2016.

Figure 6: 2SLS estimates of effect of differential viewership on cases and deaths



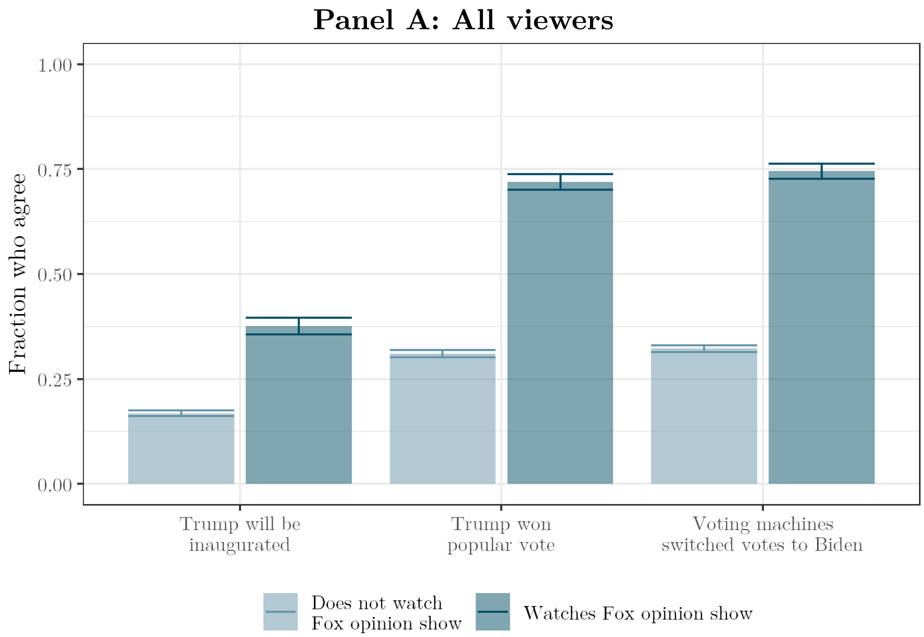
Notes: Figure 6 shows day-by-day 2SLS estimates on log one plus cases and log one plus deaths. We report day-by-day effects of the standardized difference in viewership of *Hannity* vs. *Tucker Carlson Tonight*, instrumented by $\tilde{s}_{mc,H} \times \tilde{f}_{mc,-HT}$ and controlling for state fixed effects and a large set of controls: Fox News' and MSNBC's share of cable in January 2018, Fox News' share of television in January 2020, the population density of the county, the log of the county's total population, the predicted share of TVs turned to non-Fox channels during *Hannity*, *Tucker Carlson Tonight*, and *The Ingraham Angle*, the population-weighted latitude and longitude, the percent in the county living in rural areas, the percent white, Hispanic, and black, the percent over the age of sixty-five, the share of men and women lacking high school degrees, the share of men and women lacking college degrees, the fraction of the population lacking health insurance, an age-adjusted measure of the average physical health in the county from 2018, the percent under the federal poverty line, log median household income, the unemployment rate, the 2016 Republican vote share, and the log total number of votes cast in 2016. We cluster standard errors at the DMA level and report 95 percent confidence intervals.

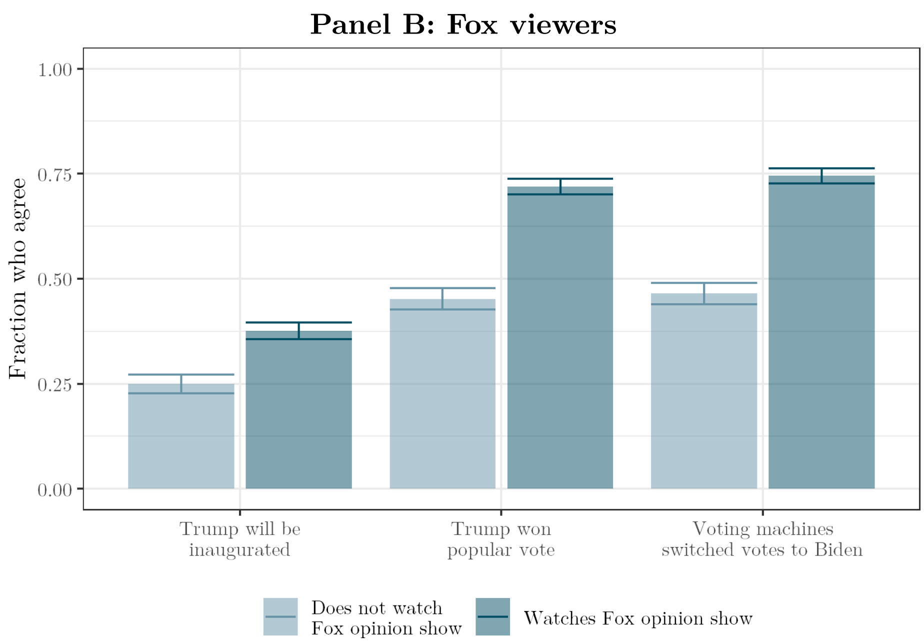
Figure 7: Carlson-Hannity content gaps and effects on cases and deaths



Notes: Figure 7 shows five time series. First, in tan diamonds, we plot the "pandemic coverage gap": the difference in portrayed seriousness of the coronavirus threat on *Tucker Carlson Tonight* vs. *Hannity*. Second, in green squares, we plot the "behavioral change gap": the difference between the *Hannity* and *Tucker Carlson Tonight* coefficients in regressions of an indicator variable for whether the respondent has changed their behavior by the date in question on indicators for viewership of different Fox News shows. In blue crosses, gray circles, and red triangles, we plot the 2SLS estimates of the Hannity-Carlson viewership gap (instrumented by $\tilde{s}_{mc,H} \times \tilde{f}_{mc,-HT}$) on the share leaving home, log one plus cases, and log one plus deaths, respectively. These latter three specifications control for state fixed effects, Fox News' and MSNBC's share of cable in January 2018, Fox News' share of television in January 2020, the population density of the county, the log of the county's total population, the predicted share of TVs tuned to non-Fox channels during *Hannity*, *Tucker Carlson Tonight*, and *The Ingraham Angle*, the population-weighted latitude and longitude, the percent in the county living in rural areas, the percent white, Hispanic, and black, the percent over the age of sixty-five, the share of men and women lacking high school degrees, the share of men and women lacking college degrees, the fraction of the population lacking health insurance, an age-adjusted measure of the average physical health in the county from 2018, the percent under the federal poverty line, log median household income, the unemployment rate, the 2016 Republican vote share, and the log total number of votes cast in 2016. We show one-week moving averages for each time series. All coefficients are rescaled to facilitate plotting on the same figure.

Figure 8: Opinion show viewership and election conspiracism





Notes: Figure 8 reports data from a election survey conducted between December 30 and January 2. The figure plots the mean level of agreement with statements indicating beliefs in various election-related conspiracy theories, separately for respondents who watch Fox News opinion shows and respondents who do not. Panel A presents estimates using the full sample; Panel B restricts to Fox News viewers. 95% confidence intervals based upon robust standard errors are reported.

Tables

Table 1: Trust in opinion shows

	Dependent variable:					
	Respondent chose opinion show					
	COVID-19 (1)	Economy (2)	GDP (3)	Earnings (4)	Unemployment (5)	Pooled (6)
Panel A: Fox News viewers						
$100 incentive	−0.022	−0.010	0.006	−0.049	−0.124	−0.027
	(0.104)	(0.098)	(0.096)	(0.101)	(0.093)	(0.039)
Dep. var. mean	0.748	0.759	0.741	0.735	0.739	0.745
Observations	107	112	112	102	115	548
Panel B: MSNBC viewers						
$100 incentive	0.107	−0.072	0.091	−0.014	0.095	0.058
	(0.111)	(0.113)	(0.117)	(0.109)	(0.115)	(0.045)
Dep. var. mean	0.534	0.592	0.604	0.616	0.683	0.606
Observations	103	98	101	99	104	505

Notes: The dependent variable in all columns of both panels is an indicator for whether the respondent chose to watch an opinion show (*Tucker Carlson Tonight* or *Hannity* for Panel A, and *The Rachel Maddow Show* or *The Last Word with Lawrence O'Donnell* for Panel B) rather than a straight news show. Columns 1–5 limit the sample to respondents assigned to the designated outcome; Column 6 pools all respondents. The explanatory variable is an indicator taking value one if the respondent was assigned to a $100 incentive and zero if the respondent was assigned to a $10 incentive. All specifications control for age, a set of race indicators, a Hispanic indicator, a male indicator, a set of education indicators, a set of household income indicators, and a set of employment indicators. Robust standard errors are reported.

Table 2: Correlation between show viewership and timing of behavior change

	Dependent variable:			
	—	Changed before...		
	Change day	March 1	March 15	April 1
	(1)	(2)	(3)	(4)
Watches Hannity	4.452^{***}	-0.112^{***}	-0.076^{*}	-0.051^{**}
	(1.282)	(0.033)	(0.043)	(0.024)
Watches Carlson	-3.362^{***}	0.117^{***}	0.042	0.021
	(1.188)	(0.031)	(0.039)	(0.022)
p-value (Hannity=Carlson)	< 0.001	< 0.001	0.097	0.076
DV mean	39.016	0.163	0.680	0.922
R^2	0.058	0.063	0.022	0.043

Notes: The dependent variable in Column 1 is the number of days after February 1, 2020 on which the respondent reported having significantly changed any of their behaviors in response to the coronavirus. For respondents who report not changing behavior by the date of the survey, we recode the dependent variable to the date of the survey (April 3). The dependent variables in Columns 2-4 are indicators for whether the respondent reported having significantly changed their behaviors before the date specified in the column header. Demographic controls include age, a white/not Hispanic indicator, a male indicator, a set of education indicators, a set of household income indicators, and a set of employment indicators. Other viewership controls include indicators for whether the respondent watches CNN or MSNBC at least once a week. Robust standard errors are reported.

Table 3: Effect of differential viewership on COVID-19 outcomes

	Dependent variable:						
	COVID-19 outcomes						
	Feb 29	Mar 07	Mar 14	Mar 21	Mar 28	Apr 04	Apr 11
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Estimates on cases							
Subpanel A.1: OLS							
Hannity-Carlson viewership difference	0.005**	0.020**	0.048**	0.096***	0.102**	0.089**	0.079*
	(0.003)	(0.009)	(0.020)	(0.034)	(0.041)	(0.044)	(0.043)
Subpanel A.2: Two-stage least squares							
H-C viewership difference (predicted)	0.041***	0.162***	0.360***	0.296**	0.232	0.105	0.101
	(0.013)	(0.044)	(0.099)	(0.140)	(0.173)	(0.180)	(0.177)
Panel B: Estimates on deaths							
Subpanel B.1: OLS							
Hannity-Carlson viewership difference	0.001	0.005	0.004	0.022**	0.042**	0.060*	0.086**
	(0.001)	(0.004)	(0.005)	(0.009)	(0.020)	(0.032)	(0.038)
Subpanel B.2: Two-stage least squares							
H-C viewership difference (predicted)	0.004**	0.017*	0.021	0.084***	0.260***	0.376**	0.326*
	(0.002)	(0.010)	(0.013)	(0.029)	(0.080)	(0.150)	(0.172)
Full controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,100	3,100	3,100	3,100	3,100	3,100	3,100

Notes: The dependent variable is the log of one plus the cumulative number of COVID-19 cases in the county as of the date referenced in the column. Panel A.1 reports OLS estimates of the log of one plus cases upon the standardized difference in Hannity-Carlson viewership; Panel B.1 replicates for deaths. Panel A.2 reports two-stage least squares estimates of the log of one plus cases upon the standardized difference in Hannity-Carlson viewership, instrumented by $\tilde{s}_{mc,H} \times \tilde{f}_{mc,-HT}$ — that is, the number of TVs on during Hannity's timeslot based on other DMAs in the same time zone, excluding TVs watching *Hannity*, multiplied by Fox News' viewership share, excluding *Hannity* and *Tucker Carlson Tonight*.; Panel B.2 replicates for deaths. All specifications include controls for the predicted share of TVs tuned to non-Fox channels during *Hannity*, *Tucker Carlson Tonight*, and *The Ingraham Angle*, Fox News' and MSNBC's share of cable in January 2018, Fox News' share of television in January 2020, and the full set of county-level controls. Standard errors are clustered at the DMA level.

Table 4: Correlation between opinion show viewership and election conspiracism

	Dependent variable:		
	Voting machines	Popular vote	Inauguration
	(1)	(2)	(3)
Panel A: All respondents			
Watches Fox opinion show	0.198***	0.179***	0.081***
	(0.014)	(0.014)	(0.012)
Watches MSNBC opinion show	−0.090***	−0.091***	−0.003
	(0.026)	(0.025)	(0.022)
Watches Fox News	0.045***	0.042***	−0.002
	(0.012)	(0.012)	(0.011)
Watches MSNBC	−0.068***	−0.039*	−0.058***
	(0.023)	(0.023)	(0.020)
Watches CNN	−0.060***	−0.064***	−0.046***
	(0.010)	(0.010)	(0.008)
Dep. var. mean	0.390	0.376	0.201
Observations	13,744	13,744	13,744
Panel B: Fox News viewers only			
Watches Fox opinion show	0.191***	0.167***	0.085***
	(0.015)	(0.015)	(0.015)
Dep. var. mean	0.632	0.612	0.325
Observations	3,681	3,681	3,681

Notes: The dependent variable in Column 1 is an indicator taking value 1 if the respondent believed that Trump would be inaugurated on January 20, 2021. The dependent variable in Column 2 is an indicator taking value 1 if the respondent believes that Trump won the popular vote in the 2020 US Presidential election. The dependent variable in Column 3 is an indicator taking value 1 if the respondent believes that voting machines switched votes from Trump to Biden. Panel A presents estimates on the full sample; Panel B restricts to Fox News viewers. All specifications control for age, a white indicator, a Hispanic indicator, a male indicator, a set of education indicators, a set of household income indicators, a set of employment indicators, a married indicator, and the respondent's political party. Robust standard errors are reported.

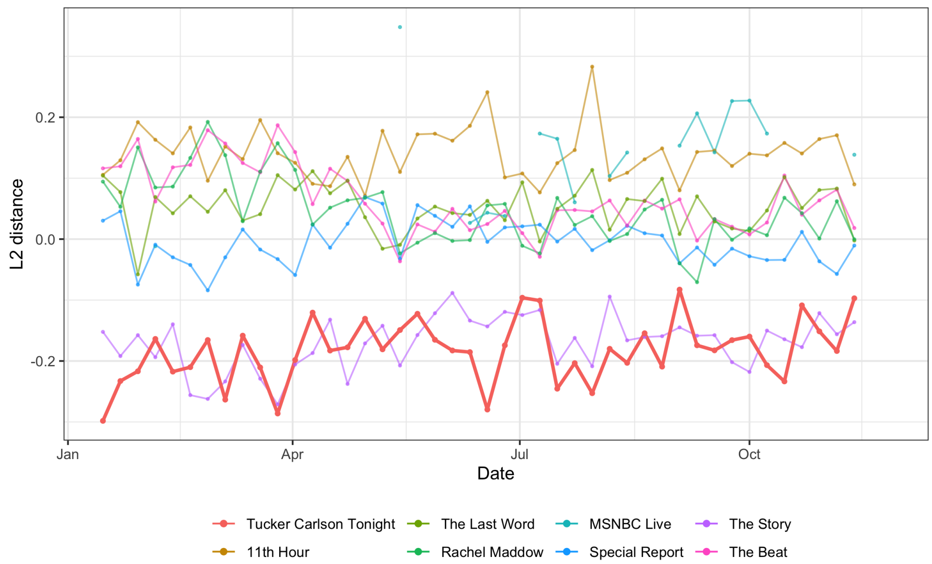
Supplementary Appendix

Our supplementary material is organized as follows. In Appendix A, we present the supplementary tables and figures referenced in the main body text. In Appendix B, we describe the data sources used in our observational analysis. In Appendix C, we present a natural language processing exercise designed to quantify differences in the content of *Hannity* and *Tucker Carlson Tonight* prior to and during early 2020. In Appendix D, we carry out several exercises to probe the robustness of our estimates and inference. In Section E, we calibrate an epidemiological model to assess our estimated effects. In Appendix F, we include copies of the instruments for our survey and experiment.

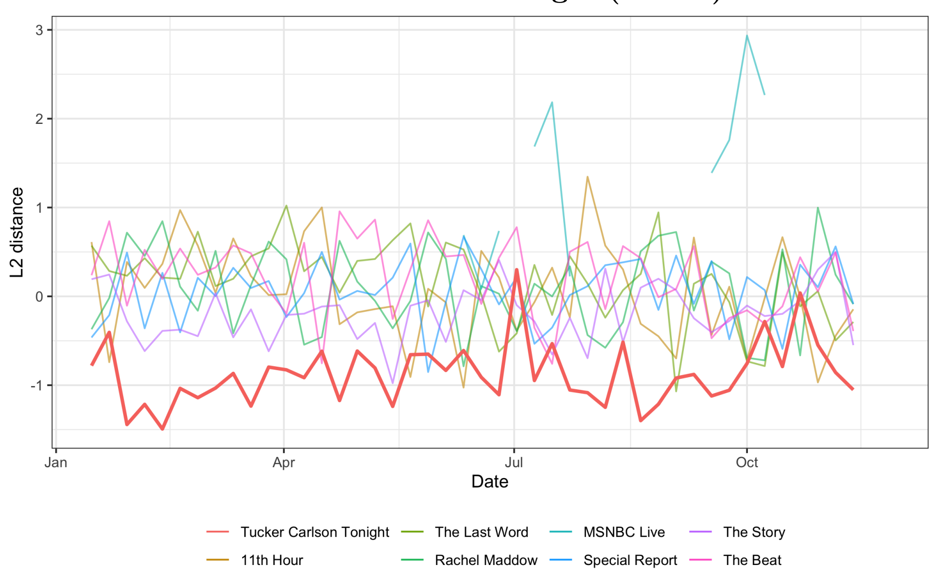
A Appendix Tables and Figures

Figure A1: Similarity of show content in 2019

Panel A: Extensive margin (LDA)



Panel B: Intensive margin (BERT)



Notes: For each week's transcripts, Figure A1 plots the Euclidean distance between the vector for *Hannity* and the topic vector for each other show. In Panel A, we compare the topic vectors for pairs of shows, as generated by Latent Dirichlet Allocation (LDA). In Panel B, we consider the contextual embeddings for pairs of shows, as generated by the Bidirectional Encoder Representation from Transformers (BERT), and average across all LDA topics. Each point is normalized such that the mean similarity across all shows is zero.

Table A1: Sample representativeness

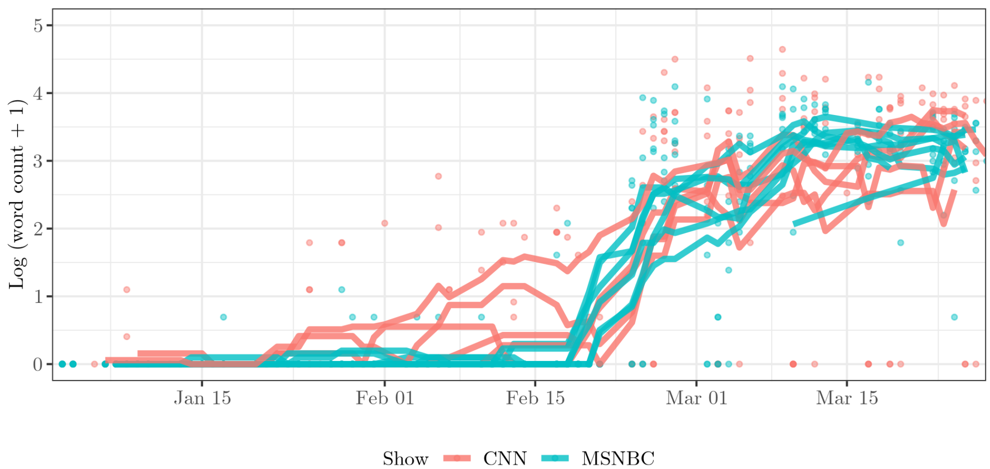
Variables:	Survey	Gallup
Male	0.61	0.50
Age	65.34	67.31
Race: White	0.95	0.93
At least high school degree	0.99	0.93
Bachelor degree or above	0.38	0.30
Employed full-time	0.26	0.29
Annual household income (USD)	71758.37	60115.93
Observations	1045	12932

Table A2: Demographics of Tucker Carlson Tonight vs. Hannity viewers

Demographic	***Tucker Carlson Tonight***	***Hannity***
Age	65.41	64.9
Male	0.52	0.56
Retired	0.57	0.49
Works full time	0.2	0.27
Household income ($)	75982.14	71816.41
White	0.89	0.96
Years of education	14.71	14.44
Watches CNN	0.16	0.24
Watches MSNBC	0.07	0.15
Watches broadcast news	0.04	0.07

Notes: Table presents mean values of each demographic characteristic among exclusive viewers of *Hannity* and *Tucker Carlson Tonight*, based on our survey of 1,045 Republicans over the age of 65 who watch Fox News. We code a respondent as watching broadcast news if they mention watching NBC, CBS, ABC, BBC, or any form of “local” or “broadcast” news.

Figure A2: Show content: CNN and MSNBC



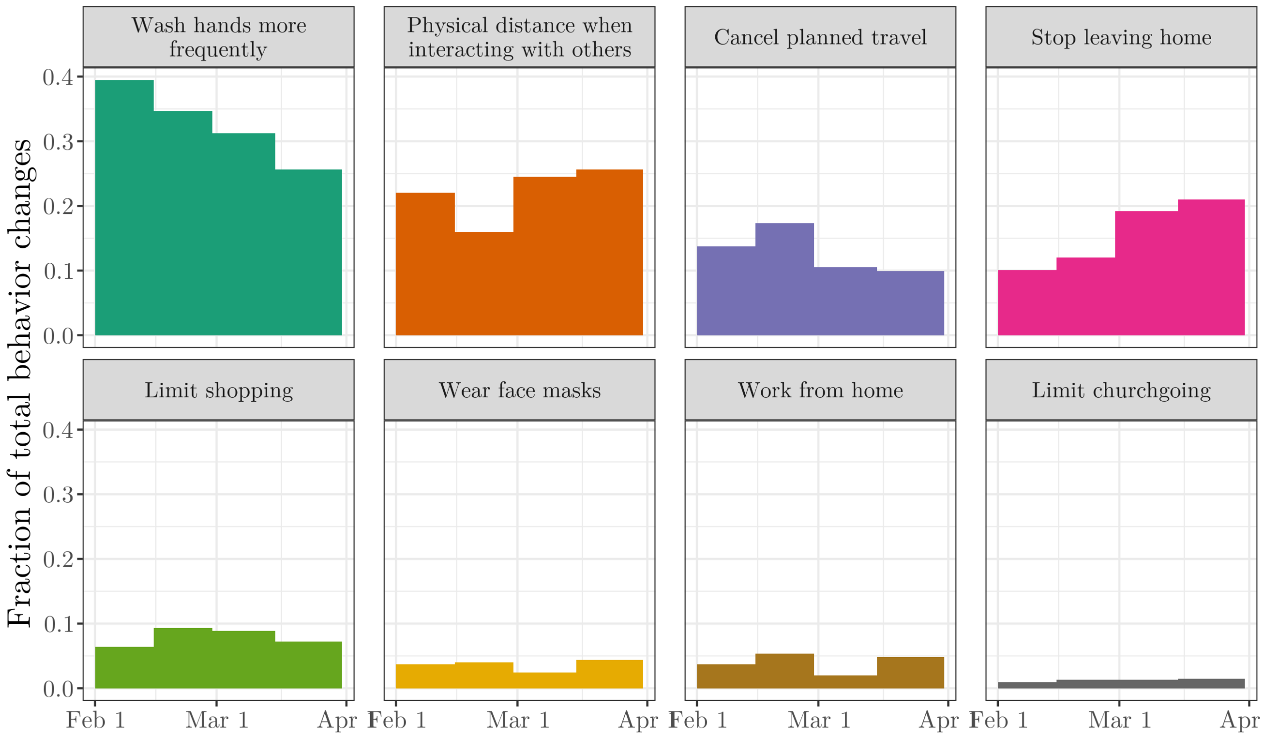
Notes: Figure displays counts of coronavirus-related terms (coronavirus, COVID, virus, influenza, and flu) separately for all shows aired on CNN and MSNBC between 5pm and 11pm local time across all four major time zones in the continental US. We display one-week rolling means.

Table A3: Demographics of Tucker Carlson Tonight vs. Hannity viewers (Dec 2020 survey)

Demographic	***Tucker Carlson Tonight***	***Hannity***
Age	50.22	52.58
Household income ($)	75797.72	70764.02
Household size	3.1	3.02
White	0.76	0.8
Hispanic	0.11	0.1
Years of education	14.92	15.03
Married	0.61	0.64
Republican	0.49	0.54
2016 Trump voter	0.61	0.64
2020 Trump voter	0.65	0.6

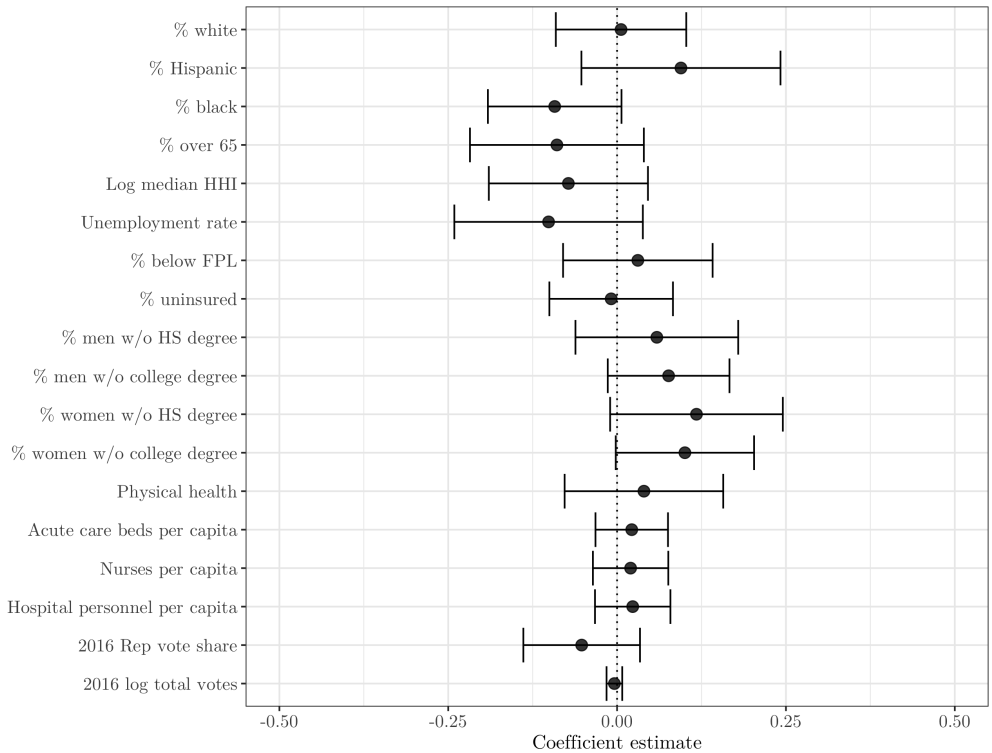
Notes: Table presents mean values of each demographic characteristic among exclusive viewers of *Hannity* and *Tucker Carlson Tonight*, based on our December 2020 survey. Sample consists of 3,694 respondents who report watching Fox News.

Figure A3: Margins of behavioral adjustment



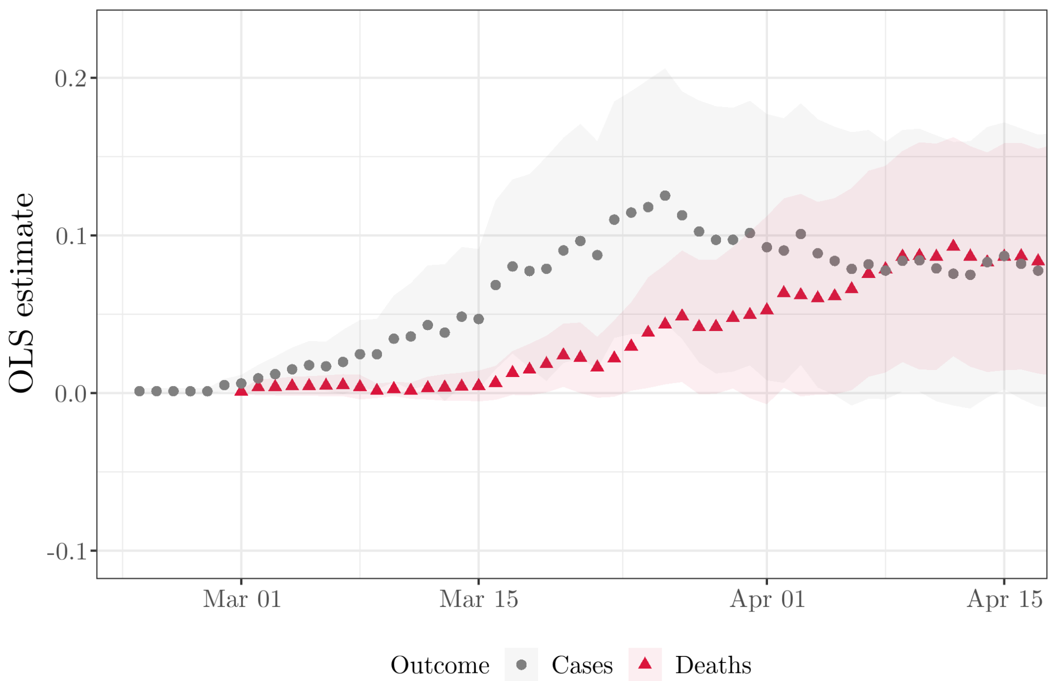
Notes: For each two-week interval between February 1 and April 1, Figure A3 shows the fraction of reported behavioral changes falling under each category. Behaviors were coded based upon responses to the following open-ended question from our survey: "When did you first significantly change any of your behaviors (for example, cancelling travel plans, washing hands or disinfecting significantly more than often, staying six feet away from others, asking to work from home, etc.) in response to the coronavirus? How did you change your behavior? Why did you change your behavior?"

Figure A4: Selection into viewership



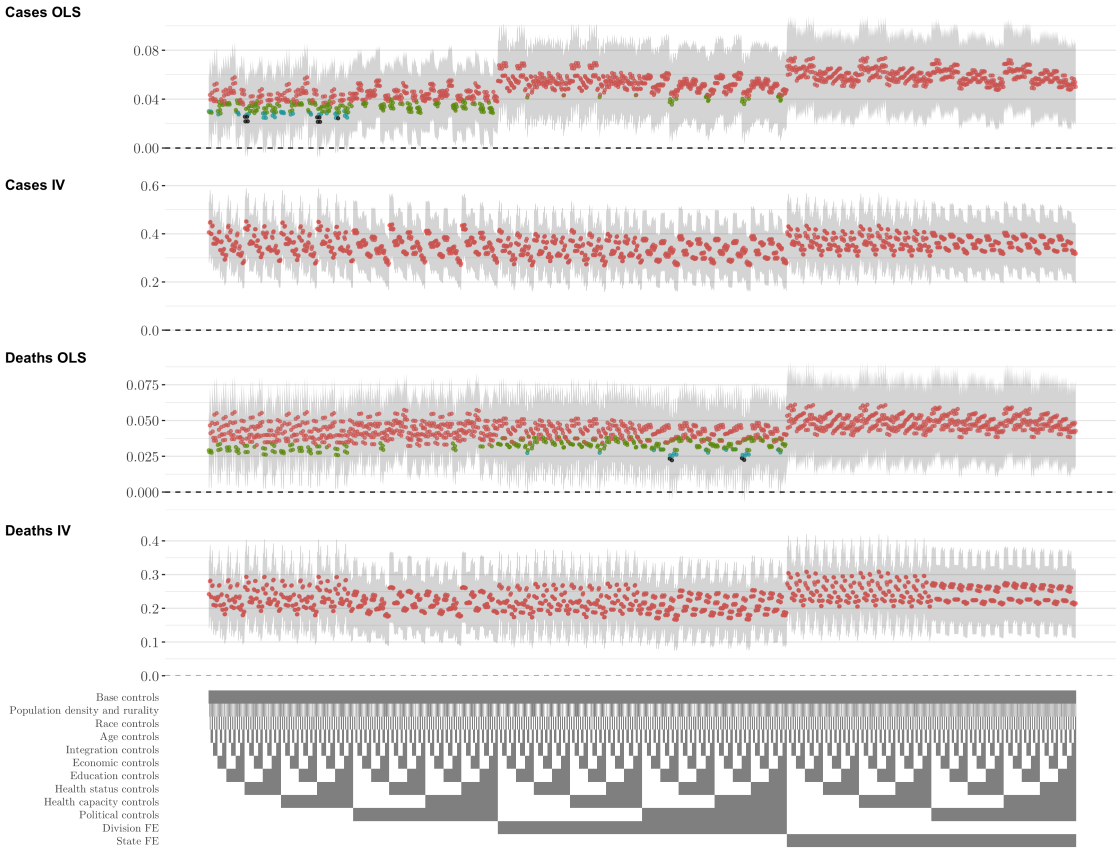
Notes: Figure A4 shows the coefficients from a series of regressions of each demographic characteristic on our measure of relative viewership of *Hannity* compared to *Tucker Carlson Tonight*, conditional on the two interactants and a small set of other controls accounting for local viewership patterns (the predicted share of TVs tuned to non-Fox channels during *Hannity*, *Tucker Carlson Tonight*, and *The Ingraham Angle*, the local viewership share of MSNBC, log population and population density, and population-weighted latitude and longitude). All dependent variables are standardized to mean zero and standard deviation one. We cluster standard errors at the DMA level and report 95 percent confidence intervals.

Figure A5: OLS estimates of effect of differential viewership on cases and deaths (state clustering)



Notes: Figure A5 displays OLS estimates of the effect of differential viewership of *Hannity* and *Tucker Carlson Tonight* on log one plus cases and log one plus deaths. We report day-by-day results for the correlation between log deaths and log cases with the standardized viewership difference between *Hannity* and *Tucker Carlson Tonight*. All regressions are conditional on state fixed effects and a large set of controls: the November 2018 and January 2020 market share of Fox News, the November 2018 market share of MSNBC, log total population, population density, the share of TVs turned to non-Fox channels during *Hannity*, *Tucker Carlson Tonight*, and *The Ingraham Angle*, the population-weighted latitude and longitude, the percent in the county living in rural areas, the percent white, Hispanic, and black, the percent over the age of sixty-five, the share of men and women lacking high school degrees, the share of men and women lacking college degrees, the fraction of the population lacking health insurance, an age-adjusted measure of the average physical health in the county from 2018, the percent under the federal poverty line, log median household income, the unemployment rate, the 2016 Republican vote share, and the log total number of votes cast in 2016. We cluster standard errors at the state level and report 95 percent confidence intervals.

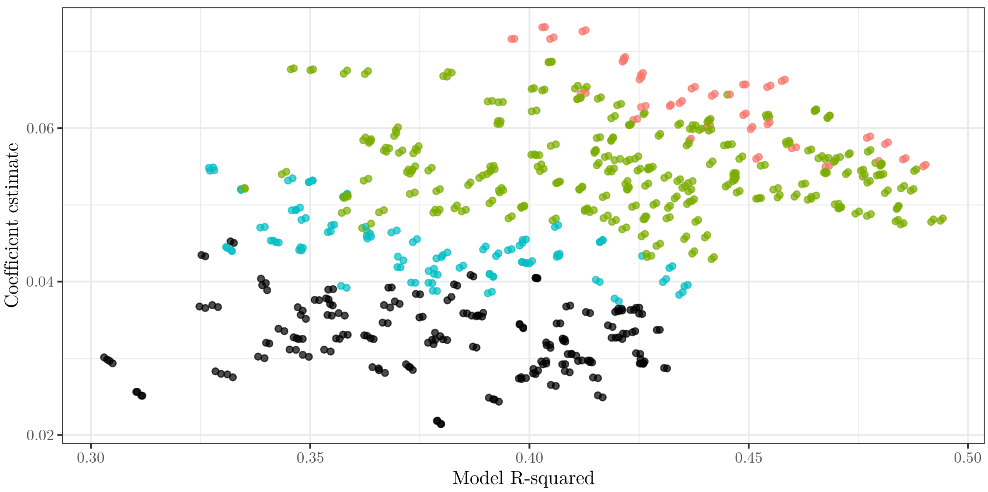
Figure A6: Stability of coefficient estimates on cases and deaths



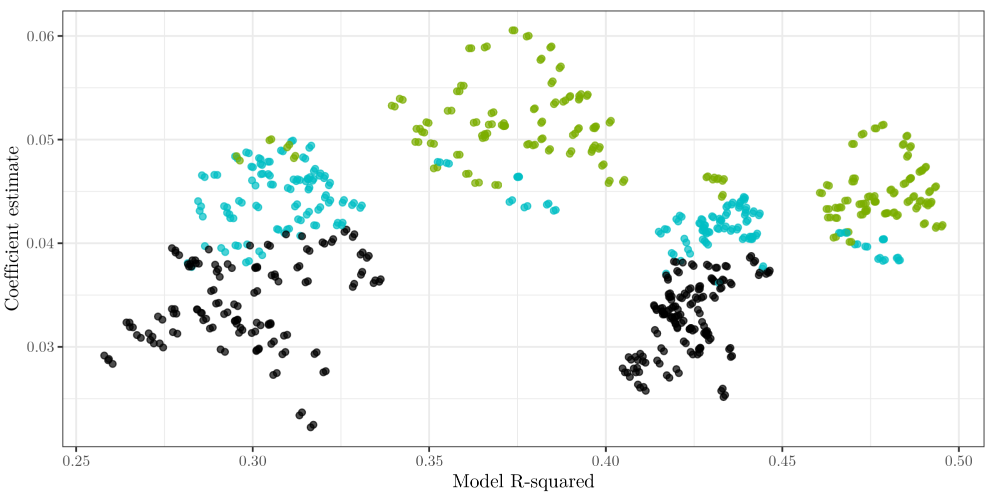
Notes: Figure A6 shows robustness of our OLS and IV estimates for the specifications for log one plus cases on March 14 and for log one plus deaths on March 28 under every possible combination of nine sets of county-level controls — population density and rurality, race, age, economic, education, health status, health capacity, politics, and market integration (presence of an airport, miles of highways, distance to a city with a population of greater than 500,000) — and our three levels of fixed effects (no fixed effects, Census division fixed effects, and state fixed effects). We cluster standard errors at the DMA level and report 95 percent confidence intervals. Black points are not significant at the ten percent level; blue points are significant at the ten percent level; green points are significant at the five percent level, and red points are significant at the one percent level.

Figure A7: OLS: R^2 vs. coefficient estimates under combinations of controls

Panel A: Estimates on log cases (March 14, 2020)

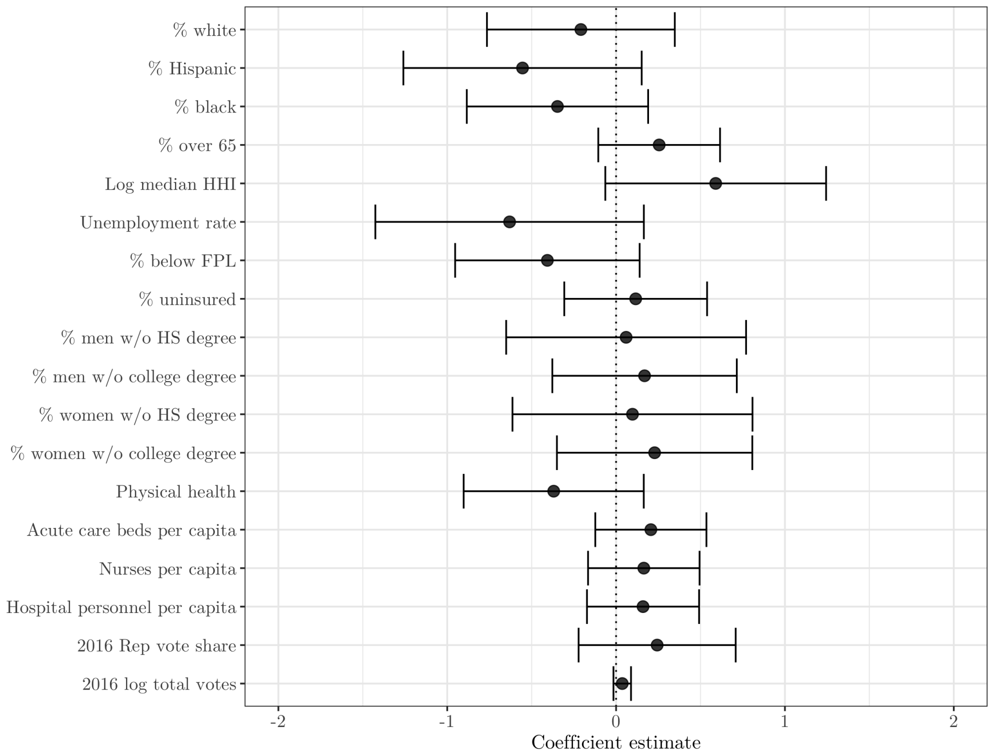


Panel B: Estimates on log deaths (March 28, 2020)



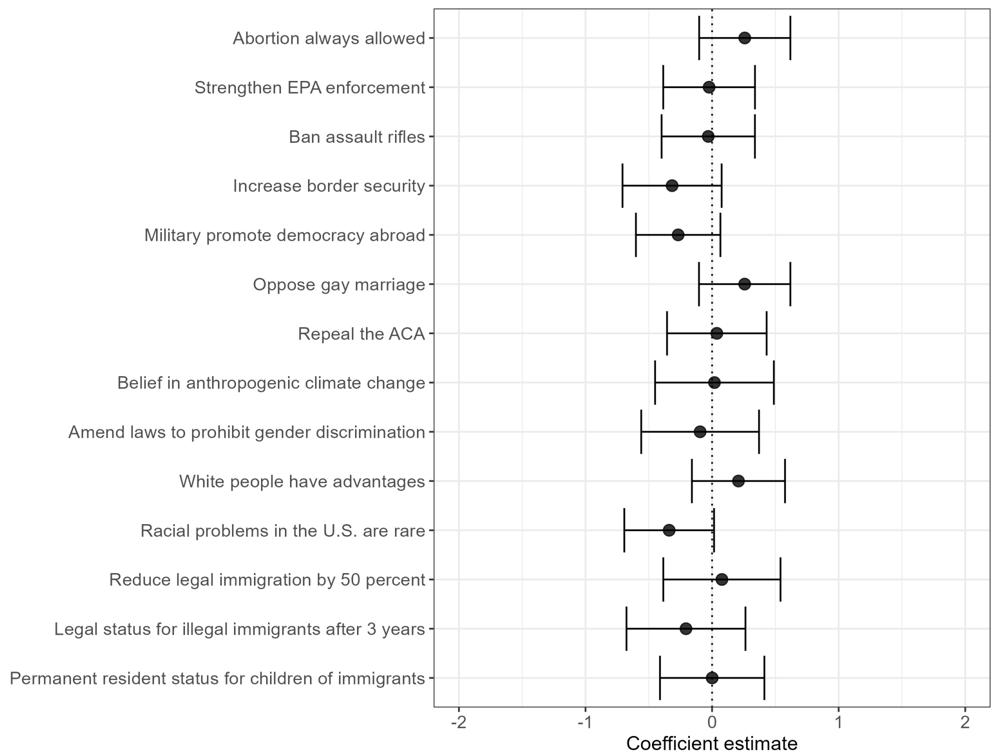
Notes: Figure A7 shows the relationship between the OLS coefficient estimates (y-axis) and the model R^2 (x-axis) for log cases on March 14 (Panel A) and for log deaths on March 28 (Panel B) from specifications with every possible combination of our eight sets of county-level controls (population density and rurality, race, age, economic, education, health status, health capacity, politics) and our three levels of fixed effects (no fixed effects, Census division fixed effects, and state fixed effects). We cluster standard errors at the DMA level. Black points are not significant at the $p < 0.1$ level; blue points are significant at the $p < 0.1$ level; green points are significant at the $p < 0.05$ level, and red points are significant at the $p < 0.01$ level.

Figure A8: Instrument correlation with county-level demographics



Notes: Figure A8 shows the coefficients from a series of regressions of each demographic characteristic on our instrument, $\tilde{s}_{mc,H} \times \tilde{f}_{mc,-HT}$, conditional on the two interactants and a small set of other controls accounting for local viewership patterns (the predicted share of TVs tuned to non-Fox channels during *Hannity*, *Tucker Carlson Tonight*, and *The Ingraham Angle*, the local viewership share of MSNBC, log population and population density, and population-weighted latitude and longitude). All dependent variables are standardized to mean zero and standard deviation one. We cluster standard errors at the DMA level and report 95 percent confidence intervals.

Figure A9: Instrument correlation with county-level political attitudes



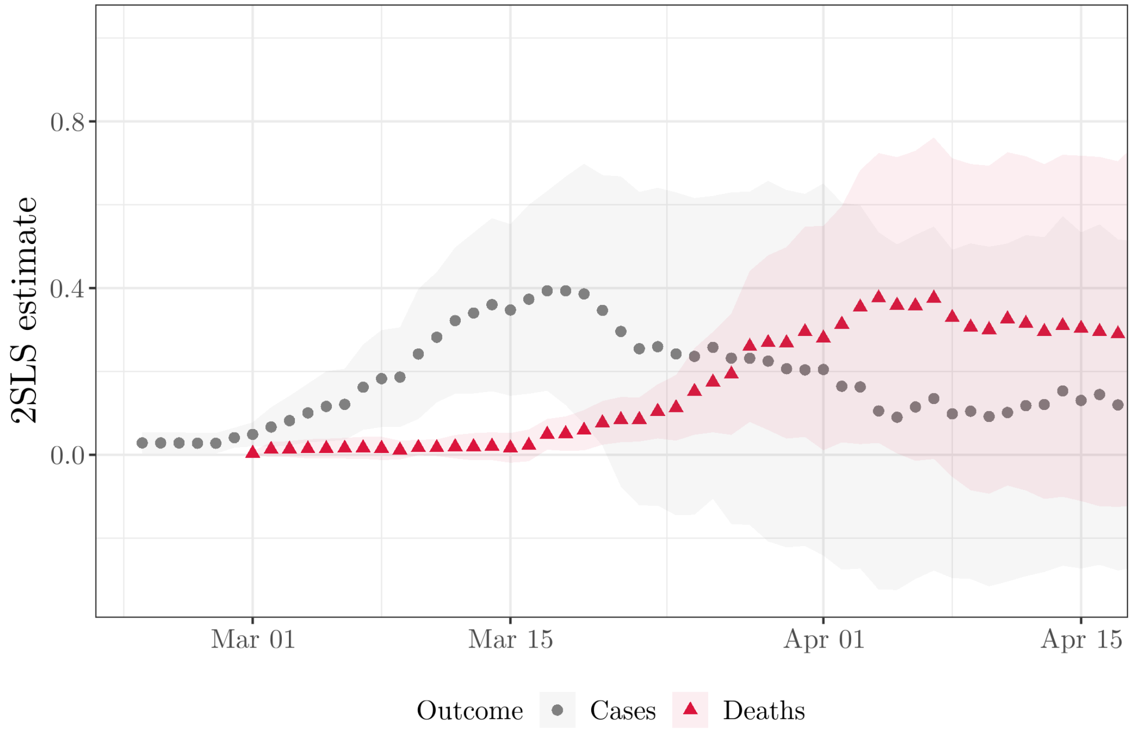
Notes: Figure A9 shows the coefficients from a series of regressions of each measure of political attitude on our instrument, $\tilde{s}_{mc,H} \times \tilde{f}_{mc,-HT}$, conditional on the two interactants and a small set of other controls accounting for local viewership patterns (the predicted share of TVs tuned to non-Fox channels during *Hannity*, *Tucker Carlson Tonight*, and *The Ingraham Angle*, the local viewership share of MSNBC, log population and population density, and population-weighted latitude and longitude). All dependent variables are standardized to mean zero and standard deviation one. We cluster standard errors at the DMA level and report 95 percent confidence intervals.

Table A4: First-stage regressions

	Dependent variable:					
	Difference in Hannity-Carlson viewership					
Non-Fox TVs on × Fox share	1.122***	1.088***	1.184***	1.123***	1.127***	1.117***
	(0.331)	(0.314)	(0.275)	(0.264)	(0.260)	(0.258)
Controls	Base	Full	Base	Full	Base	Full
Fixed effects	None	None	Division	Division	State	State
Observations	3,103	3,100	3,103	3,100	3,103	3,100
R^2	0.733	0.752	0.804	0.811	0.835	0.838
F-statistic	11.48	12.02	18.49	18.08	18.74	18.74

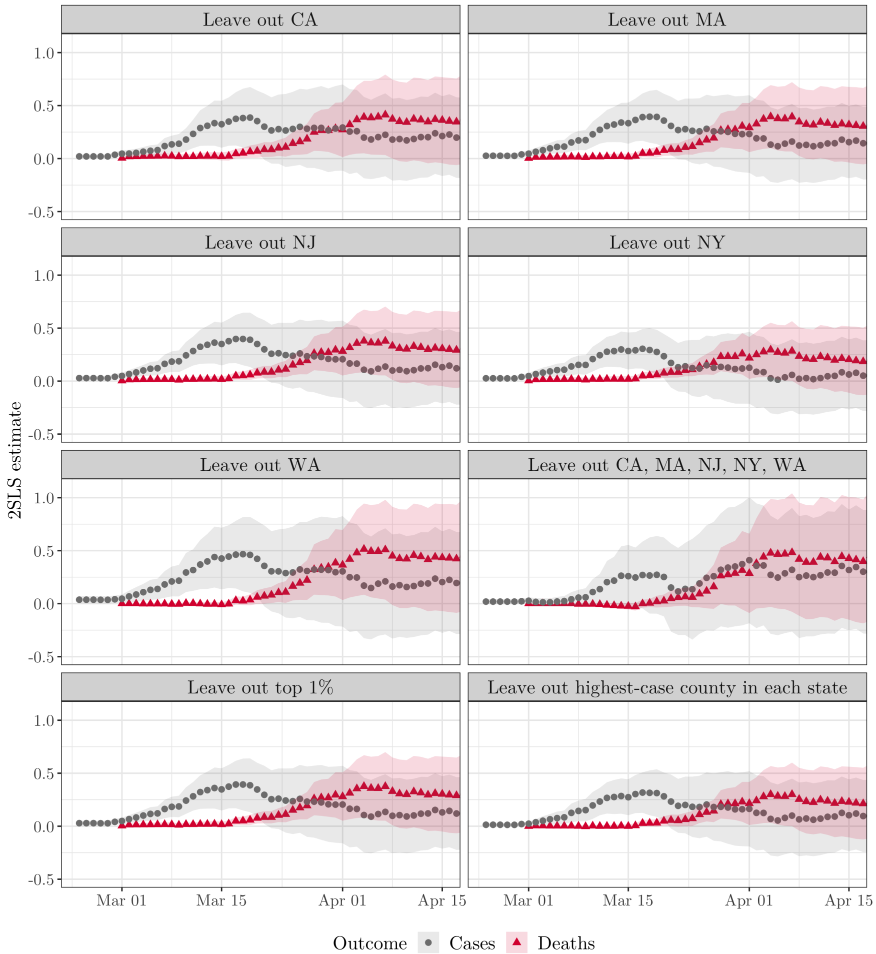
Notes: Table reports regressions of the standardized difference between viewership of *Hannity* and *Tucker Carlson Tonight* on our instrument, $\tilde{s}_{mc,H} \times \tilde{f}_{mc,-HT}$ — that is, the number of TVs on during Hannity's timeslot based on other DMAs in the same time zone, excluding TVs watching *Hannity*, multiplied by Fox News' viewership share, excluding *Hannity* and *Tucker Carlson Tonight*. Fox share and predicted viewership include the predicted share of TVs tuned to non-Fox channels during *Hannity* and during the show immediately before and immediately afterward, as well as Fox News' share of cable, leaving out *Hannity* and *Tucker Carlson Tonight*. "Base controls" include the predicted share of TVs tuned to non-Fox channels during *Hannity*, *Tucker Carlson Tonight*, and *The Ingraham Angle*, Fox News' and MSNBC's share of cable in January 2018, Fox News' share of television in January population density and log population, and population-weighted latitude and longitude. "Full controls" additionally include the percent of the population living in a rural area, the population over the age of 65, the percent male with no high school degree, the percent female with no high school degree, the percent male with no college degree, the percent female with no college degree, an age-adjusted measure of the average physical health in the county, the percent uninsured, the percent below the federal poverty line, the log of the median household income, the unemployment rate, the Republican vote share in 2016, and the log of the total number of votes in the county in 2016. Standard errors are clustered at the DMA level.

Figure A10: 2SLS estimates of effect of differential viewership on cases and deaths (state clustering)



Notes: Figure A10 shows day-by-day 2SLS estimates on log one plus cases and log one plus deaths. We report day-by-day effects of the standardized difference in viewership of *Hannity* vs. *Tucker Carlson Tonight*, instrumented by $\tilde{s}_{mc,H} \times \tilde{f}_{mc,-HT}$ and controlling for state fixed effects and a large set of controls: Fox News' and MSNBC's share of cable in January 2018, Fox News' share of television in January 2020, the population density of the county, the log of the county's total population, the predicted share of TVs turned to non-Fox channels during *Hannity*, *Tucker Carlson Tonight*, and *The Ingraham Angle*, the population-weighted latitude and longitude, the percent in the county living in rural areas, the percent white, Hispanic, and black, the percent over the age of sixty-five, the share of men and women lacking high school degrees, the share of men and women lacking college degrees, the fraction of the population lacking health insurance, an age-adjusted measure of the average physical health in the county from 2018, the percent under the federal poverty line, log median household income, the unemployment rate, the 2016 Republican vote share, and the log total number of votes cast in 2016. We cluster standard errors at the state level and report 95 percent confidence intervals.

Figure A11: Leave-out IV estimates of effect of differential viewership on cases and deaths



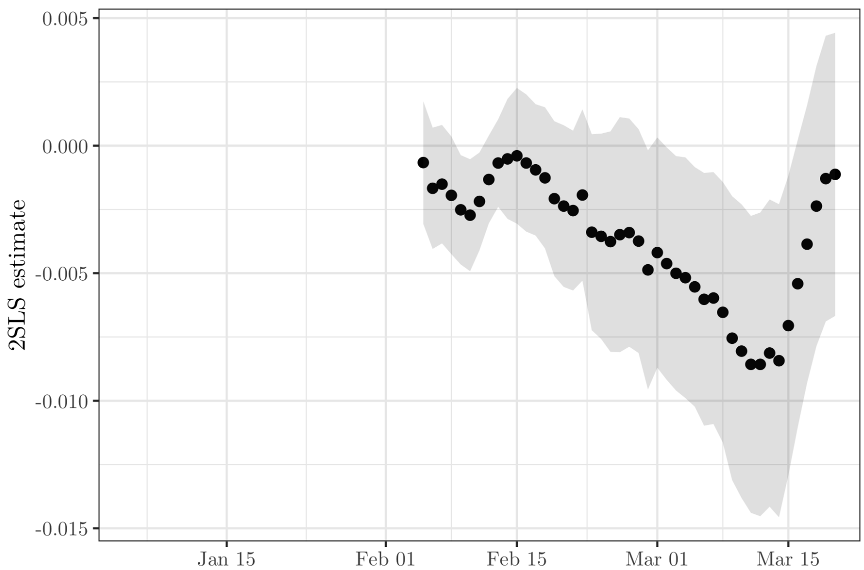
Notes: Figure A11 displays effects of differential viewership of *Hannity* and *Tucker Carlson Tonight* on log one plus cases and log one plus deaths, leaving out states containing known COVID-19 hotspots or high-COVID counties. We report day-by-day results for the correlation between log deaths and log cases with the standardized viewership difference between *Hannity* and *Tucker Carlson Tonight*. All regressions are conditional on state fixed effects and a large set of controls: the November 2018 and January 2020 market share of Fox News, the November 2018 market share of MSNBC, log total population, population density, the predicted share of TVs turned to non-Fox channels during *Hannity*, *Tucker Carlson Tonight*, and *The Ingraham Angle*, the population-weighted latitude and longitude, the percent in the county living in rural areas, the percent white, Hispanic, and black, the percent over the age of sixty-five, the share of men and women lacking high school degrees, the share of men and women lacking college degrees, the fraction of the population lacking health insurance, an age-adjusted measure of the average physical health in the county from 2018, the percent under the federal poverty line, log median household income, the unemployment rate, the 2016 Republican vote share, and the log total number of votes cast in 2016. We cluster standard errors at the DMA level and report 95 percent confidence intervals.

Table A5: Effects of differential viewership on COVID-19 outcomes (count models)

	Poisson			Zero-inflated NB		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Estimates on cases						
Non-Fox TVs on × Fox share	1.410***	1.425***	1.283***	2.106***	1.644**	2.039***
	(0.292)	(0.317)	(0.288)	(0.706)	(0.749)	(0.726)
Panel B: Estimates on deaths						
Non-Fox TVs on × Fox share	1.040***	0.531	1.009***	2.268***	2.538***	2.052***
	(0.298)	(0.369)	(0.297)	(0.691)	(0.762)	(0.759)
Fixed effects	None	Division	State	None	Division	State
Observations	3,100	3,100	3,100	3,103	3,103	3,103

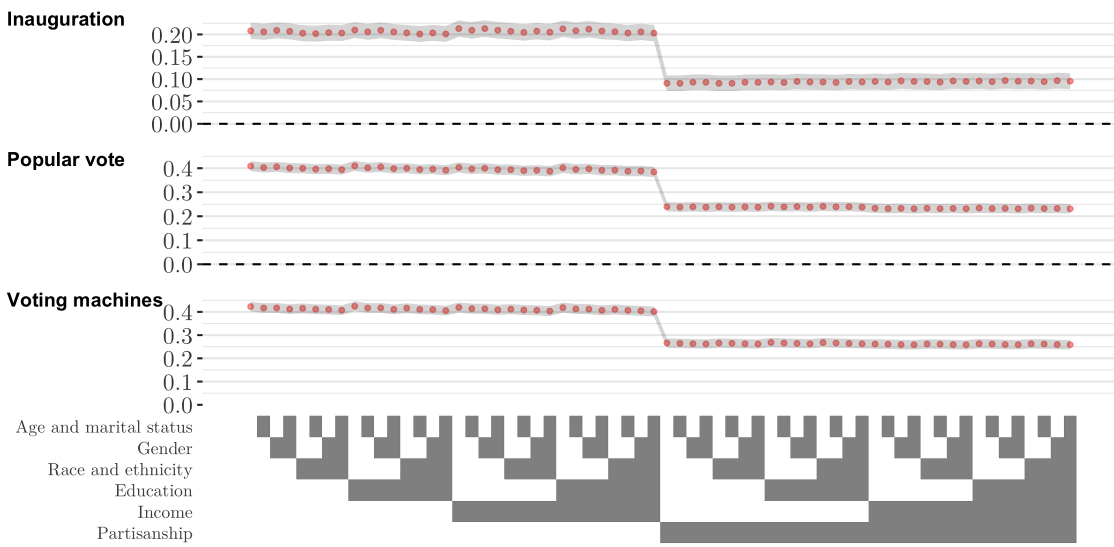
Notes: Table reports reduced-form regressions of the log of one plus cases (Panel A) and the log of one plus deaths (Panel B) on our instrument, $\tilde{s}_{mc,H} \times \tilde{f}_{mc,-HT}$ — that is, the number of TVs on during Hannity's timeslot based on other DMAs in the same time zone, excluding TVs watching *Hannity*, multiplied by Fox News' viewership share, excluding *Hannity* and *Tucker Carlson Tonight.*. Columns 1–3 report Poisson regressions; Columns 4–6 report zero-inflated negative binomial regressions, using the same set of regressors for both the hurdle and the intensive margin. All specifications control for the share of TVs tuned to non-Fox channels during *Hannity*, *Tucker Carlson Tonight*, and *The Ingraham Angle*, Fox News' and MSNBC's share of cable in January 2018, Fox News' share of television in January population density and log population, population-weighted latitude and longitude, the percent of the population living in a rural area, the population over the age of 65, the percent male with no high school degree, the percent female with no high school degree, the percent male with no college degree, the percent female with no college degree, an age-adjusted measure of the average physical health in the county, the percent uninsured, the percent below the federal poverty line, the log of the median household income, the unemployment rate, the Republican vote share in 2016, and the log of the total number of votes in the county in 2016. Standard errors are clustered at the DMA level.

Figure A12: 2SLS estimates of effect of differential viewership on stay-at-home behavior



Notes: Figure A12 shows day-by-day 2SLS estimates on the fraction of people staying home on each day. We report day-by-day effects of the standardized difference in viewership of *Hannity* vs. *Tucker Carlson Tonight*, instrumented by $\tilde{s}_{mc,H} \times \tilde{f}_{mc,-HT}$ and controlling for state fixed effects and a large set of controls: Fox News' and MSNBC's share of cable in January 2018, Fox News' share of television in January 2020, the population density of the county, the log of the county's total population, the predicted share of TVs turned to non-Fox channels during *Hannity*, *Tucker Carlson Tonight*, and *The Ingraham Angle*, the population-weighted latitude and longitude, the percent in the county living in rural areas, the percent white, Hispanic, and black, the percent over the age of sixty-five, the share of men and women lacking high school degrees, the share of men and women lacking college degrees, the fraction of the population lacking health insurance, an age-adjusted measure of the average physical health in the county from 2018, the percent under the federal poverty line, log median household income, the unemployment rate, the 2016 Republican vote share, and the log total number of votes cast in 2016. We cluster standard errors at the DMA level and report 95 percent confidence intervals.

Figure A13: Stability of coefficient estimates



Notes: Figure A13 shows robustness of our OLS estimates of the correlation between opinion show viewership and belief in three elections-related conspiracy theories: that Trump will be inaugurated as president in 2021, that Trump won the popular vote, and that voting machines switched votes from Biden to Trump. All estimates are significant at the 1% level. We report 95% confidence intervals.

B Overview of Data Sources

Aside from our survey and the show transcripts we use in our previously-described content validation, we employ six primary categories of data in our observational analysis, which we describe in detail below.

Viewership data Our show viewership data is provided by Nielsen. Nielsen reports viewership at the Designated Market Area (DMA) level, of which there are 210 in the US.[54] We focus on the continental United States, excluding the two DMAs in Alaska (Anchorage and Fairbanks) and the single DMA in Hawaii (Honolulu).[55] Our dataset contains viewership data between 5pm and 11pm (local time) at the DMA-by-timeslot-by-day level (i.e. hourly ratings). In addition to the fraction of TVs watching Fox News, we observe the fraction of TVs turned on during each timeslot. We supplement this dataset with 2018 data, previously acquired, on the local market share of each of the "Big Three" networks: CNN, MSNBC, and Fox News. To avoid using variation based on *Hannity* and *Tucker Carlson Tonight*, these market shares are calculated based on evening time slots outside of those two shows. Our primary analysis uses January and February viewership data; however, given the high degree of persistence in show viewership, our results are quantitatively extremely similar and qualitatively identical if we instead use only January data (to rule out concerns about reverse causality in our OLS estimates) or if we use data from January 1 through March 8 (the beginning of Daylight Savings Time, a natural stopping point given the structure of our identification strategy).

COVID-19 cases and deaths data We use publicly-available county-level data on *confirmed* COVID-19 cases and deaths from Johns Hopkins University (Dong et al., 2020). The data is a panel at the day-by-county level, with data sourced from a variety of agencies, including the World Health Organization, the Centers for Disease Control, state health departments, and local media reports. Throughout our main analyses, we take the logarithm of one plus the cumulative number of cases and deaths, both to correct for outliers with a large number of cases and because the exponential nature by which a virus spreads makes the logarithm normalization natural. However,

[54] Comprehensive viewership data is not available at more granular levels after 2015. It is possible to approximate ZIP-level (and thus county-level) viewership in 2015 or earlier, as in Simonov et al. (2020). This approximation involves aggregating "headends," or cable systems, to ZIP codes, a procedure that requires discarding all but the largest headend in each ZIP code; (Simonov et al. 2020 find that 47% of ZIP codes have more than one headend, though the largest headend accounts for at least half of subscribers in the vast majority). Aside from this measurement error and the possibility that the change in viewership between 2015 and 2020 is endogenous, we use 2020 DMA-level data for two reasons: first, because we are interested in the effects of *contemporaneous* exposure to misinformation on pandemic outcomes and thus require viewership data from the period of interest; and second, because *Tucker Carlson Tonight* first aired in 2016, and thus constructing accurate ZIP-code level estimates of differential viewership is not feasible using currently-available data.

[55] We also exclude Palm Springs, CA; this DMA is so small that it does not contain a county centroid, and thus we are unable to consistently map any counties to Palm Springs.

our results are qualitatively identical and quantitatively extremely similar if we instead transform cases and deaths by the inverse hyperbolic sine (IHS) rather than the natural logarithm.

In our primary analysis, we focus on outcomes during the early stages of the pandemic — from late February to April 15 — given that stay-at-home orders were widely enacted in late March and the estimated 1-3 week lag between infections and deaths.[56]

Demographics We collect demographic data at the county level from a wide variety of sources. Our data on age, racial composition, and household income and educational attainment is drawn from the 2018 round of the American Community Survey. We use data on county rurality from the 2010 Census and data on population drawing from the Annual Estimates of the Resident Population for Counties in the United States. Our measures of poverty and health insurance are provided by the Census under the 2018 Small Area Income and Poverty Estimates (SAIPE) and 2018 Small Area Health Insurance Estimates (SAHIE) programs. Our data on unemployment is from the US Bureau of Labor Statistics' 2019 Local Area Unemployment Statistics (LAUS). Finally, our data on physical health is from the CDC's Behavioral Risk Factor Surveillance System (BRFSS).

2016 Republican vote share We obtain county-level voting data for the 2016 US Presidential election from the MIT Election Lab, which contains the total number of votes cast and the number of votes cast for each of the major parties.

Health system capacity We use standard measures of health capacity from the Dartmouth Atlas of Health Care's Hospital and Physician Capacity dataset. Data are at the Hospital Referral Region level, defined by the Atlas as "regional health care markets for tertiary care"; we use the most recent version of the dataset (2012). We include all three measures included in the data — the number of nurses, hospital personnel, and hospital beds — and divide by population to construct per capita measures.

Sunset timing Our data on sunset timing is drawn from `www.timeanddate.com`. We extract sunset times for every day from January 1, 2020 to March 1, 2020 for all counties based on their centroids, and we construct the sunset time of each DMA for each day as the population-weighted mean sunset time on that day of all counties in that DMA.

[56] The earliest stay-at-home order was enacted in California on March 19; other states followed suit between March 20 and April 7. While our primary specification is estimated separately for each day and employs state fixed effects, thus controlling for any state-specific policies, it is possible that the timing of regional stay-at-home orders (e.g. at the municipal, county, or DMA level) are directly influenced by coverage of the pandemic on Fox News, though such effects are likely of limited quantitative significance. It is, however, likely that the timing of regional stay-at-home orders were affected by the trajectories of cases and deaths in the county, which, as we show, are themselves affected by Fox News coverage; we view this as a mechanism.

C Evaluating Differences in Program Content using Natural Language Processing

In this section, we describe in greater depth how we evaluate differences in content between *Hannity* and *Tucker Carlson Tonight* prior to and during early 2020. We measure content differences on both the extensive and the intensive margin. On both margins, we measure differences not only between *Hannity* and *Tucker Carlson Tonight*, but also between *Hannity* and all other shows considered in Section 2 — namely, *The Story*, *Special Report*, *The Rachel Maddow Show*, *The Last Word*, *MSNBC Live*, *The 11th Hour*, and *The Beat*. This affords us a natural benchmark for our comparison of interest. In order to avoid picking up mechanical differences, we remove all anchor names and references to the program or network from the transcripts before proceeding with our exercise.

C.1 Extensive Margin of Coverage

To quantify divergence along the extensive margin — that is, differences in the topics upon which programs focus — we employ Latent Dirichlet Allocation (Blei et al., 2003; Schwarz, 2018). LDA is an unsupervised topic model that, given a sufficiently large corpus of text, outputs "topics" of words that appear in similar contexts. It then uses this word-to-topic mapping to measure the extent to which a given document consists of each topic. LDA has been used in a wide variety of social science contexts, including measuring ideological polarization (Draca and Schwarz, 2021), analyzing trends over time in the content of FOMC transcripts, (Edison and Carcel, 2021) and evaluating the effects of regulation on 10-K disclosures (Dyer et al., 2017).

The econometrician specifies the number of topics to be found; we set 10 topics and train the model using all transcripts from 2019.[57] The resulting topics are intuitively coherent and correspond to the most salient political issues of 2019. We display the most characteristic words of each topic, and our subjectively-assigned topic labels, in Table C1.

With this word-to-topic mapping in hand, we calculate the L2 distance between the topic vector for *Hannity* and that of each other program for each week to construct our measure of program similarity on the extensive margin. We plot these distances in Panel A of Appendix Figure A1. We plot distances separately for each topic in Appendix Figure C1. To help validate our measure, we also repeat this exercise for 2020 and plot distances in Appendix Figure C1.

C.2 Intensive Margin of Coverage

We next turn to measuring differences in intensive margin of coverage — that is, differences in how programs discuss a given topic, conditional on discussing it at all. To do so, we use BERT (Devlin

[57]We verify that our results are also robust to choosing more or fewer topics.

Table C1: Topics generated by Latent Dirichlet Allocation (2019)

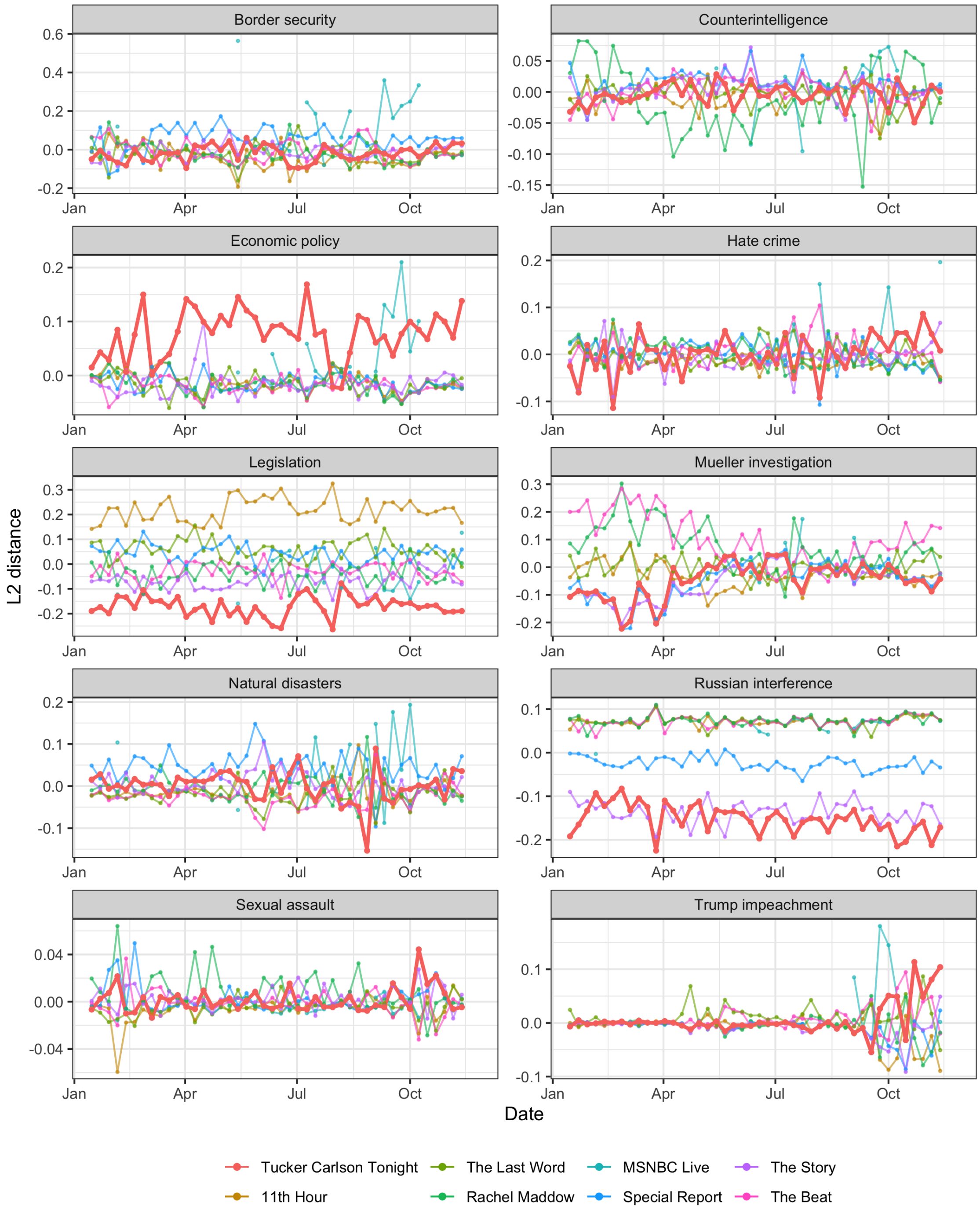
Topic	**Most characteristic words**	**Subjective label**
1	mueller, prosecutor, obstruction, judge, court	Mueller investigation
2	storm, hurricane, water, bahama, farmer	Hurricanes/natural disasters
3	collusion, lie, destroy, america, dossier	Russian interference/campaign
4	tax, college, rich, wealth, income	Economic policy
5	intelligence, russian, director, foreign, agency	Counterintelligence
6	allegation, photo, kavanaugh, christie, ronan farrow	Sexual assault/Me Too
7	racist, el paso, woman, racism, hate	Hate crime
8	border, debate, child, wall, immigration	Border security
9	secretary, poll, meeting, decision, senator	Legislation
10	ukraine, impeachment, whistleblower, giuliani, inquiry	Trump impeachment

et al., 2019), a machine learning technique that creates high-dimensional vector representations, or contextual embeddings, capturing documents' semantic meaning. The implementation we use was pre-trained on the Google Books corpus and English-language Wikipedia, and "has become a ubiquitous baseline in NLP experiments" (Rogers et al., 2020).

We begin by using BERT to create a vector representation of each 512-token segment from our dataset of transcripts (the maximum length that BERT can process, corresponding to approximately 300 words) using BERT. We then classify the topic of each segment as the topic with the highest measured probability in the segment, as measured by the LDA procedure outlined above. For each topic, we calculate the L2 distances between all embeddings from that topic and average to collapse to a program-by-topic-by-week measure of distance. We plot distances averaged across topics in Panel B of Appendix Figure A1. We plot distances separately for each topic in Panel B of Appendix Figure C2.

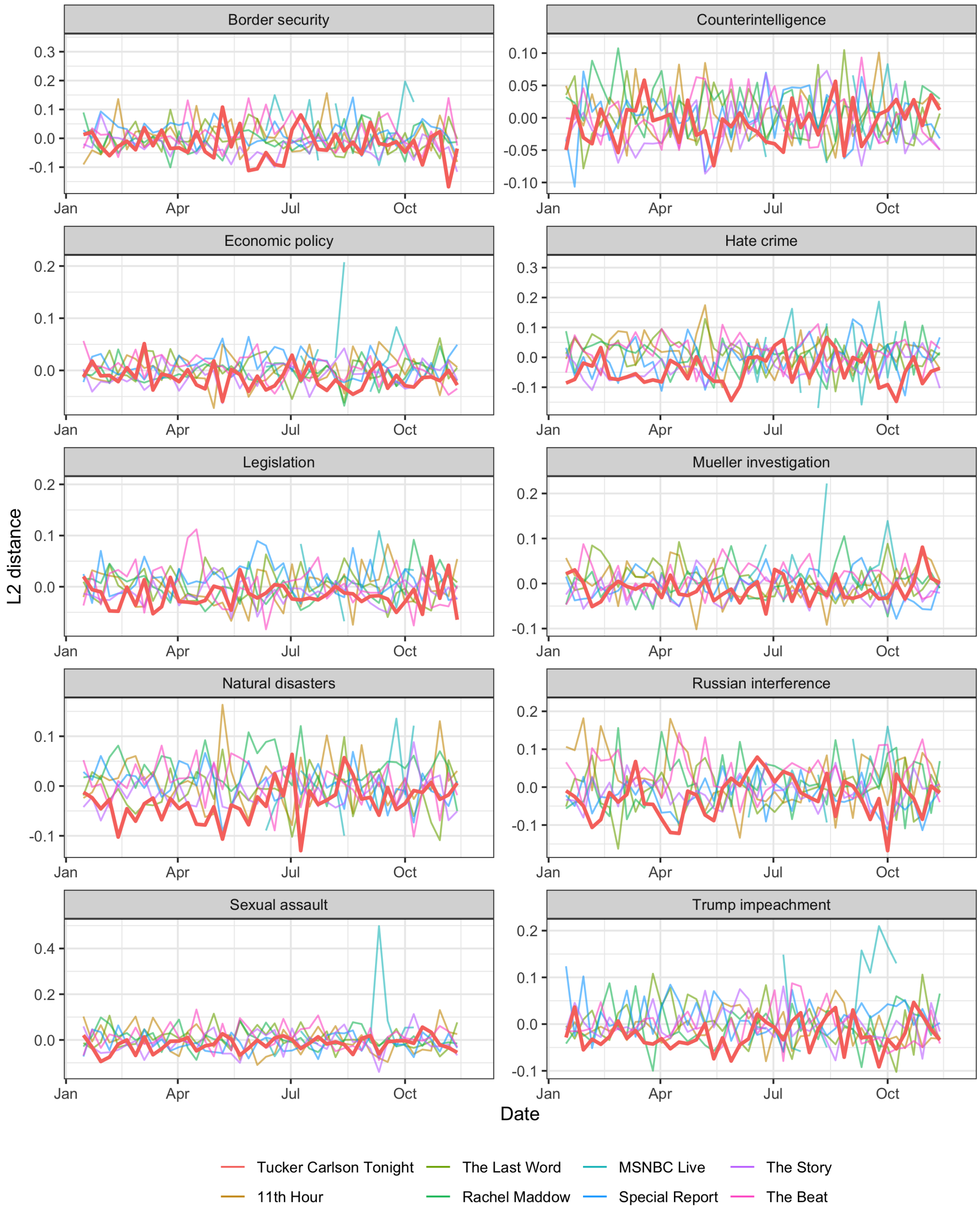
One important caveat of this approach is that the reliability of comparisons in how two programs discuss a given topic relies on the existence of a sufficiently large set of segments on that topic from both programs. Thus, comparisons are likely to be uninterpretable in weeks where one or both shows spend relatively little airtime on the topic under consideration.

Figure C1: Similarity of show topics in 2019



Notes: For each week's transcripts, and for each topic, Figure C1 plots the difference between the fraction of *Hannity* devoted to that topic and the fraction of each other show devoted to that topic, as measured by Latent Dirichlet Allocation (LDA). Each point is normalized such that the mean similarity across all shows is zero.

Figure C2: Similarity in how topics are discussed in 2019



Notes: For each week's transcripts, and for each topic, Figure C2 plots the average L2 distance between the contextual embeddings from *Hannity* classified by LDA as pertaining to that topic and the analogous contextual embeddings from each other show. Contextual embeddings are generated using BERT. Each point is normalized such that the mean similarity across all shows is zero.

Figure C3: Similarity in show topics in 2020

Notes: For each week's transcripts, and for each topic, Figure C3 plots the difference between the fraction of *Hannity* devoted to that topic and the fraction of each other show devoted to that topic, as measured by Latent Dirichlet Allocation (LDA). Each point is normalized such that the mean similarity across all shows is zero.

D Robustness

Robustness to zero values To ensure that our results are not driven by zero values, we construct an unbalanced panel wherein a county only enters the panel once it has a COVID-19 case. In Appendix Figure D2, we report 2SLS estimates. Because relatively few counties had a non-zero number of cases during early March, our main specification (which includes a rich set of county-level controls, along with state fixed effects) results in a singular or close-to-singular matrix until mid-March, and even afterward, confidence intervals are relatively large. Nonetheless, our estimates are qualitatively similar (though quantitatively smaller), and our estimates on deaths are statistically significant at the five percent level between mid-March and early April. The somewhat smaller effect sizes are consistent with an important role of movements in both the intensive and extensive margins in shaping our results. Estimates on cases are not statistically significant at the five percent level.

Bootstrap In Appendix Figure D3, we calculate our standard errors via a block bootstrap procedure, randomly sampling DMAs with replacement and estimating counterfactual treatment effects for each day. We employ a conservative approach to calculating standard errors: rather than ex ante fixing the set of counties between the 0.025-quantile and the 0.975-quantile of *average* treatment effects, we compute confidence intervals separately by day, using the 0.025-quantile and the 0.975-quantile of the estimated treatments effects *on each day* as the upper and lower bounds on our confidence intervals, respectively. Our bootstrapped standard errors are larger and thus our effects are statistically significant for a somewhat shorter period of time: effects on cases are statistically significant from early to mid March, while effects on deaths are statistically significant from mid-March to late April. However, our findings remain qualitatively unchanged.

Randomization inference To address error arising from treatment variation (including spatial autocorrelation), in Appendix Figure D4, we employ a randomization inference approach (Athey and Imbens, 2017), permuting the plausibly exogenous "shift" ($\tilde{s}_{mc,H}$) across DMAs while leaving the "shares" ($\tilde{f}_{mc,-HT}$), the county-level covariates, and cases and deaths unchanged. For each repetition, we then regenerate our instrument as the interaction of the placebo $\tilde{s}_{mc,H}$ with $\tilde{f}_{mc,-HT}$, then estimate placebo treatment effects as before. Under this approach, we find that our effects on cases and deaths are statistically significant at the 5% level throughout essentially the same period as described above.

Permutation test To ensure that our results are not driven by statistical artifacts, in Appendix Figure D5 we randomly permute the joint tuple of case and death counts across counties and estimate counterfactual treatment effects. The resulting distribution of estimates is centered around zero; and once more, our true estimates for cases exceed the 0.975-quantile of counterfactual estimates from early to mid March, while our true estimates for deaths exceed the 0.975-quantile of counterfactual estimates from late March to mid-April.

Predicted DMA level viewership curve A key source of variation driving variation in our main leave-out instrument, Z_{mc}, is differing preferences across localities for when to watch TV. The use of leave-outs to generate cleaner and plausibly exogenous variation in differential exposure to the two shows has the limitation that it is somewhat unclear what remaining underlying factors are driving the residual variation in timing preferences. In particular, the concern would be some confounding determinant of health outcomes still covarying with preferences for the time slot of the respective shows, in ways which interact with the market share of Fox News. While this possibility seems somewhat remote, it cannot be ruled out. By contrast, in an ideal experiment, one would randomly assign Fox viewers to different timeslots, exposing some areas more to *Hannity* and other areas more to *Tucker Carlson Tonight*. To get closer to this ideal, we now consider an extension of the instrument which more explicitly exploits variation in timing preferences.

Specifically, we show – and empirically exploit – important systematic patterns that drive TV viewership over the course of the evening, in ways that are highly unlikely to interact with the leave-out Fox News market share to drive health outcomes. In particular, DMAs across the country exhibit a relatively consistent *inverse-U shaped* relationship between the time since sunset and total TV viewership. Panel A of Figure D6 plots a non-parametric local polynomial fitting the relationship between time since sunset and the fraction of TVs tuned to non-Fox channels. On average across the country, TV viewership peaks 2.5 hours after sunset and then declines smoothly. Panel A also shows a histogram depicting, at each twelve-minute interval relative to sunset, the number of DMAs in which *Tucker Carlson Tonight* begins in that interval (blue) and in which *Hannity* begins in that interval (purple). Because both shows are broadcast live — *Tucker Carlson Tonight* at 8pm Eastern Time and *Hannity* at 9pm Eastern Time — both shows are aired much earlier and closer to sunset in more Western time zones (e.g. 5pm and 6pm Pacific Time, respectively). Yet as Panel B of Figure D6 highlights, even holding constant what (clock) time shows air, there remains substantial variation in start time relative to sunset. While DMAs differ in the precise shape of their viewership curve over the course of the evening, the vast majority exhibit a clear inverted-U pattern.[58] For example, on February 1, 2020, the sun set at 6:05pm in Louisville, KY, whereas it set at 5:19pm in Philadelphia, PA — nearly an hour earlier. Thus, predicted viewership during *Hannity*'s timeslot is larger in Louisville, as "prime time" is at approximately

[58] Episodes of *Tucker Carlson Tonight* and *Hannity* are generally re-run three hours after they first air, and because our data spans 5pm to 11pm, we observe repeats in more western time zones but not in Eastern Time. In order to avoid making assumptions about viewership patterns in western time zones relative to Eastern Time by failing to include Eastern Time viewership that falls outside of the window covered by our data, we simply set viewership to the average viewership across both airings in DMAs in which we observe re-runs. However, our results are robust to only using viewership of the live broadcasts.

8:30pm, only 30 minutes before *Hannity* airs. Predicted viewership during *Hannity*'s timeslot is lower in Philadelphia, where the local prime time of TV consumption is forty five minutes earlier.

Our identification strategy exploits cross-DMA variation in sunset timing and viewership preferences alongside timezone-specific variation in local airtimes of *Hannity* and *Tucker Carlson Tonight*, such that cross-DMA variation in the predicted amount of total TV viewership during Hannity's timeslot — or more precisely, total non-Fox TV viewership during this timeslot — generates variation in relative viewership of *Tucker Carlson Tonight* vs. *Hannity.*

Let $\widehat{\mathrm{s}}_{mc,H}$ denote the *predicted* fraction of TVs turned on in DMA d at the time slot of *Hannity*, leaving out TVs watching Fox News (i.e. leaving out TVs watching *Hannity*).[59] We predict $\mathrm{s}_{mc,H}$ parametrically for each DMA using a second-degree polynomial. Denoting by n_{mt} the sunset time in DMA m on day t, we have:

$$\mathrm{s}_{mc,H} = \alpha_m + \delta_{m1}(s - n_{mt}) + \delta_{m2}(s - n_{mt})^2 + \epsilon_{dst} \tag{4}$$

As before, letting f_{mc} denote the viewership share of Fox News in DMA m, leaving out *Hannity* and *Tucker Carlson Tonight*, the modified instrument is given by $\widehat{\mathrm{s}}_{mc,H} \times \tilde{f}_{mc,-HT}$. The underlying logic for this modified version is the instrument is simple: if people like to turn on their TVs to watch *something* when *Hannity* happens to be on rather than when another Fox show happens to be on, simply as a function of when shows air relative to when it gets dark locally (and not just what official time it is locally), the number of viewers shifted into watching *Hannity* is disproportionately large in areas where Fox News is popular in general, for arguably exogenous reasons. As before, conditional upon the small set of controls accounting for local viewership patterns, this instrument is not significantly correlated with demographic characteristics (Appendix Figure D7) and has a strong first stage on viewership (Columns 3-4 of Appendix Table D1). Our resulting 2SLS estimates are statistically significant and are quantitatively extremely similar to those derived from our primary instrumental variables approach.

Division-level viewership curve One possible concern with both our main instrument and our sunset instrument is that they might rely excessively on local preferences (that is, DMA-specific preferences) for watching TV over the course of the evening. We now consider a prediction of the share of TVs turned on during *Hannity* and *Tucker Carlson Tonight* using *Census division-wide*, rather than DMA-specific, preferences for TV viewership over the course of the evening. Thus, our identifying variation is driven by the interaction of the viewership curve *at the division level* with DMA-specific market shares of Fox News, controlling for the main effects at the DMA level. To allow DMAs to differ in their *absolute* preference for TV viewership while keeping our identifying

[59] As mentioned above, we leave out TVs watching Fox News in order to capture a general DMA preference for TV viewership at a given time rather than specific preferences for Fox News. The logic is analogous to the logic of the leave-one-out estimator used in Bartik instruments (Bartik, 1991).

variation — the viewership curve over the course of the evening — constant, we allow the level and scale of the viewership curve to differ between DMAs within a division but hold the shape of the curve fixed. In particular, we estimate the following first-stage regression separately for each of the nine Census divisions in the United States:

$$\log(\mathrm{s}_{mc,H}) = \alpha_m + \delta_1(s - n_m) + \delta_2(s - n_m)^2 + \epsilon_{ms},$$

where the DMA-specific fixed effect α_m allows the level of the curve to vary between DMAs and the log transformation of $\mathrm{s}_{mc,H}$ allows the scale of the curve to vary between DMAs. We re-define $\widehat{\mathrm{s}_{mc,H}} = \exp(\widehat{\log \mathrm{s}_{mc,H}})$ and, as before, construct our instrument based on the interaction of $\widehat{\mathrm{s}_{mc,H}}$ with the viewership share of Fox News in DMA m, leaving out *Hannity* and *Tucker Carlson Tonight*. Our first-stage specifications are otherwise identical to those in Section 3.4.

Like our main instrument, conditional upon the small set of controls accounting for local viewership patterns, this alternative instrument is not significantly correlated with demographic characteristics (Appendix Figure D8), and it has a first stage on viewership (Columns 5-6 of Appendix Table D1), though the relationship is weaker than that which we find with our main instrument or the DMA-based sunset prediction. The 2SLS estimates should therefore be interpreted with caution. Nonetheless, we again estimate positive and significant effects of differential viewership on cases and deaths.

Two instruments Table D1 shows robustness of the 2SLS estimates when two instruments are included, the one for *Hannity* as specified in Section 3.4 and an analogously constructed instrument for *Tucker Carlson Tonight*, $\tilde{s}_{mc,T} * \tilde{f}_{mc,-HT}$. Two-stage least squares estimates are similar in magnitude and statistical significance, but — as might be expected given the correlation between the two instruments — the first stage F-statistic is smaller and below the generally-accepted threshold of 10, suggesting that including both instruments may induce a weak instruments problem and bias both our coefficients and standard errors. As there remains uncertainty in how to test for and overcome weak instruments in over-identified models, as opposed to in simpler just-identified settings (see Andrews et al. 2019 for a discussion), our primary specification uses only the instrument for *Hannity*'s timeslot.

D.1 Generalized Pandemic Coverage Index

Our previous estimates focused on the effects of our instrument on differential viewership of *Hannity* and *Tucker Carlson Tonight*. These two shows were the largest outliers on Fox News in their coverage of COVID-19 (in opposite directions), and are the most widely-watched programs on the network and in the United States, suggesting that the viewership gap between the two shows alone had effects on cases and deaths. Yet as we discuss in Section 3.4, differences in viewership across those two Fox News shows may, through various spillovers, also correlate with viewership of many other shows. Specifically, for any given DMA, regular viewership of *Tucker Carlson Tonight* (airing 8pm-9pm ET) and *Hannity* (airing 9pm-10pm ET) could lead to positive or negative selection into various combinations of: *The Five* (5pm-6pm ET); *Special Report with Bret Baier* (6pm-7pm ET); *The Story with Martha MacCallum* (7pm-8pm ET); *The Ingraham Angle* (10pm-11pm ET); and *Fox News at Night* (11pm-12pm ET).[60] Despite the fact that the other evening shows are neither as widely watched as *Hannity* and *Tucker Carlson Tonight* nor as extreme in their coverage, their content may also have influenced COVID-19 outcomes. In this case, the narrow exclusion restriction, which requires that effects operate through viewership of *Hannity* or *Tucker Carlson Tonight*, would be violated. Thus, we now turn to a more general approach to capture viewers' (predicted) exposure to misinformation on Fox News.

Specifically, for each DMA, we first calculate a measure of local exposure to information about the pandemic across *all* evening-time shows on Fox News, allowing us to consider the broad information set to which Fox News viewers were exposed. We combine our data on viewership shares of the different shows at the DMA-by-day level with our Mechanical Turk episode coding results to construct a measure of information exposure, the *pandemic coverage index*, as the average of the degree to which each episode portrayed COVID-19 as a serious threat to the United States, weighted by viewership of that episode within the DMA. More formally, we define r_{st} to be the average seriousness rating of show s on day t and m_{sdt} to be the average viewership share of episde s in DMA d among all Fox News evening-time episodes on day t. Then the *daily exposure* e_{dt} of a DMA is given by:

$$e_{dt} := \frac{1}{|S_d|} \sum_{s \in S_d} r_{st} m_{sdt}.$$

where S_d is the menu of shows between 5pm and 11pm in DMA d. We then construct the pandemic coverage index for DMA d as the sum of $\tilde{e}_{dt}$ throughout the months of January and February:

$$PCI_d := \sum_{t \in \text{Jan, Feb}} \tilde{e}_{dt}.$$

The index therefore captures an (inverse) local "stock" of exposure to narratives on Fox News downplaying the pandemic threat throughout February relative to the mean exposure across DMAs

[60]Of course, there might also be spillovers to day-time Fox News shows, but such selection would arguably be less significant given that TV is primarily viewed between 5pm and 11pm. Cross-network spillovers are also possible. Such spillovers are likely minor given that viewers tend to favor shows within the same network; indeed, in the survey discussed in Section 3.2, 73 percent of respondents report that Fox News is the only cable TV network they watch at least once a week. Moreover, as we show in Appendix Figure A2, the other two dominant cable TV networks, CNN and MSNBC, featured far less variation between shows in their coverage of COVID-19, limiting the extent that spillovers might bias our results.

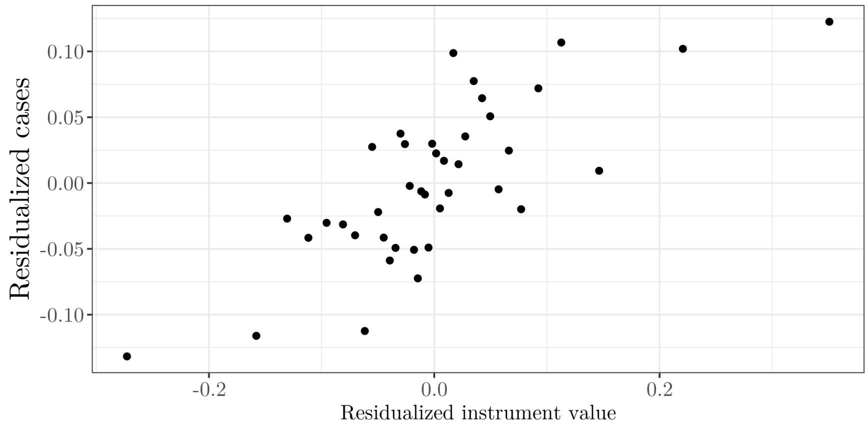
in the same period. For ease of interpretation, we scale the index to a standard normal distribution. For consistency with our previous figures, we use the inverse of our pandemic coverage index, $-1 \times PCI_d$ throughout the rest of this section.

Columns 1 and 2 of Table D2 highlight that our measure of viewership of *Hannity* relative to *Tucker Carlson Tonight* strongly predicts the pandemic coverage index ($p < 0.001$), whether we include only the minimum set of controls to capture local viewership patterns or we condition on the full set of controls employed in Section 3.4. Next, we examine the extent to which our instrument, Z_{mc}, is associated with the pandemic coverage index. Columns 3 and 4 of Table D2 show that our instrument is strongly and significantly associated with the pandemic coverage index, again whether we include only the minimum set of controls or we condition on the full set of county characteristics. Finally, in Columns 5 and 6 of Table D2, we examine the relationship between the pandemic coverage index and COVID-19 cases and deaths through 2SLS. We follow the approach from Section 3.4, but we use the pandemic coverage gap as the endogenous variable instead of the standardized difference in viewership of *Hannity* versus *Tucker Carlson Tonight*, allowing us to fully capture spillovers between shows on Fox News evening shows. Our results suggest that a one percentage point increase in the inverse of the pandemic coverage index increases the number of cases by 3.96 percent on March 14 ($p < 0.001$) and the number of deaths by 2.83 percent by March 28 ($p < 0.001$).

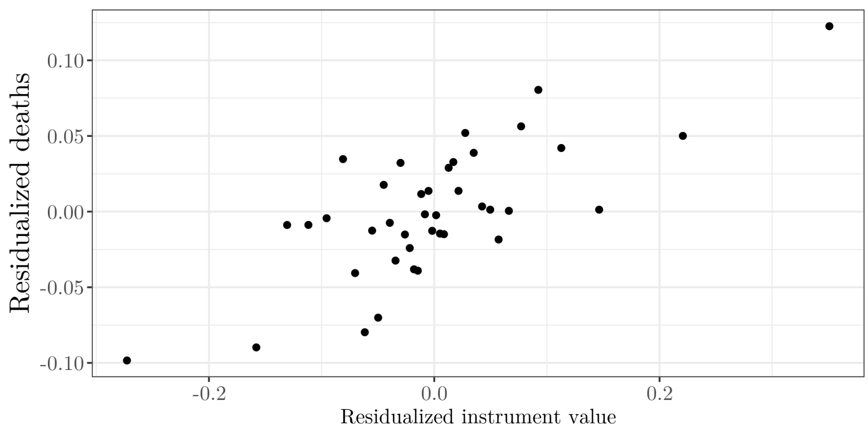
D.2 Figures and Tables

Figure D1: IV: residual-residual plot

Panel A: Estimates on log cases (March 14, 2020)

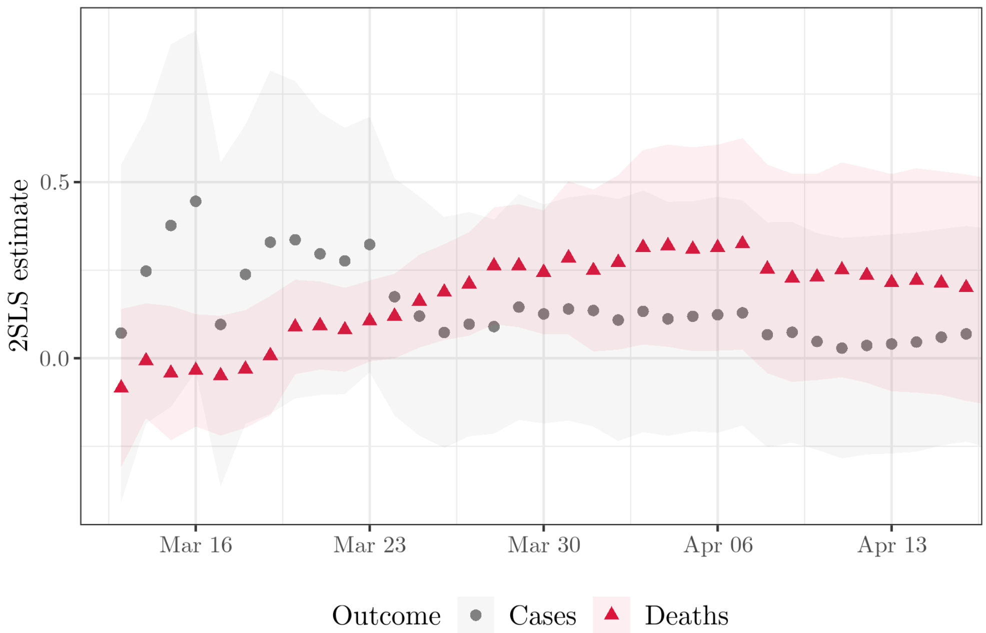


Panel B: Estimates on log deaths (March 28, 2020)



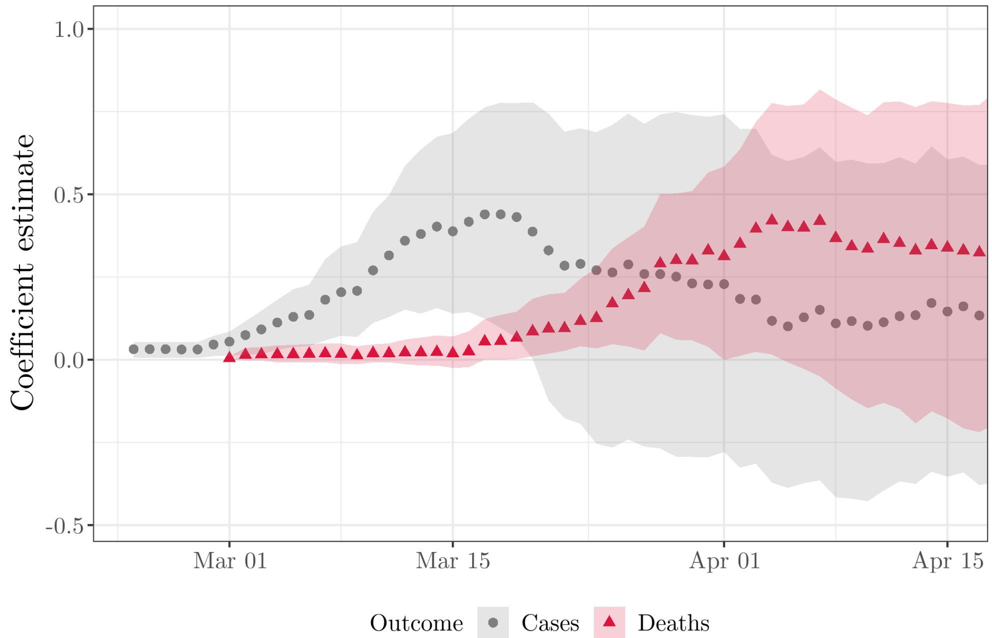
Notes: Figure D1 displays a binscatter of the residuals of log one plus cases (Panel A) and log one plus deaths (Panel B) on the residuals of $\tilde{s}_{mc,H} \times \tilde{f}_{mc,-HT}$, where both outcome variables and the instrument are residualized by state fixed effects and our full set of controls: Fox News' and MSNBC's share of cable in January 2018, Fox News' share of television in January 2020, the population density of the county, the log of the county's total population, the predicted share of TVs tuned to non-Fox channels during *Hannity*, *Tucker Carlson Tonight*, and *The Ingraham Angle*, the population-weighted latitude and longitude, the percent in the county living in rural areas, the percent white, Hispanic, and black, the percent over the age of sixty-five, the share of men and women lacking high school degrees, the share of men and women lacking college degrees, the fraction of the population lacking health insurance, an age-adjusted measure of the average physical health in the county from 2018, the percent under the federal poverty line, log median household income, the unemployment rate, the 2016 Republican vote share, and the log total number of votes cast in 2016.

Figure D2: 2SLS estimates on cases and deaths: unbalanced panel approach



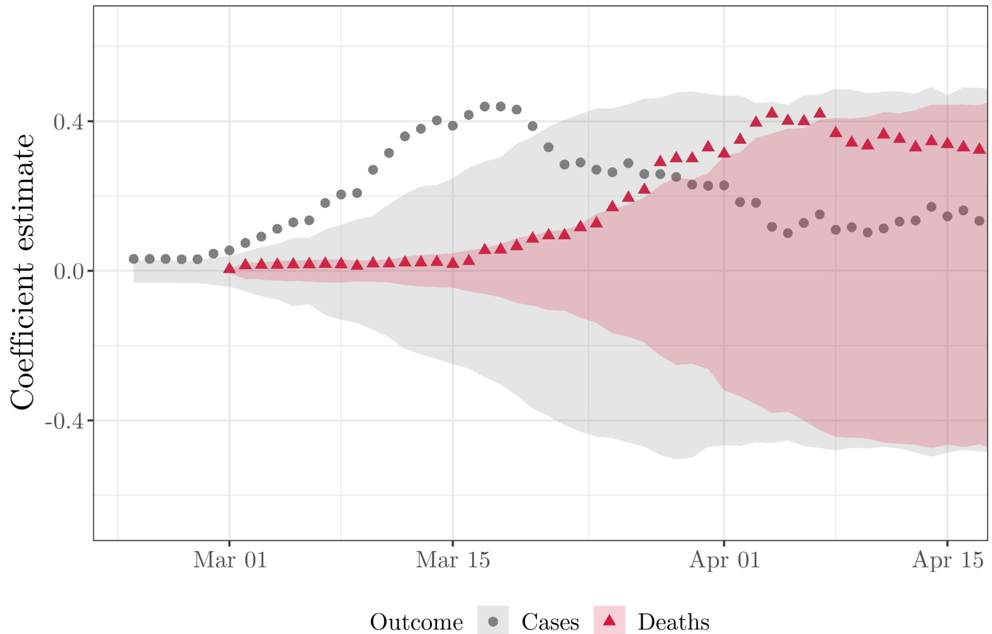
Notes: Figure D2 shows day-by-day 2SLS estimates on log one plus cases and log one plus deaths, using an unbalanced panel approach in which we drop observations with zero values of the dependent variable. We report day-by-day effects of the standardized difference in viewership of *Hannity* vs. *Tucker Carlson Tonight*, instrumented by $\tilde{s}_{mc,H} \times \tilde{f}_{mc,-HT}$ and controlling for state fixed effects and a large set of controls: Fox News' and MSNBC's share of cable in January 2018, Fox News' share of television in January 2020, the population density of the county, the log of the county's total population, the predicted share of TVs turned to non-Fox channels during *Hannity*, *Tucker Carlson Tonight*, and *The Ingraham Angle*, the population-weighted latitude and longitude, the percent in the county living in rural areas, the percent white, Hispanic, and black, the percent over the age of sixty-five, the share of men and women lacking high school degrees, the share of men and women lacking college degrees, the fraction of the population lacking health insurance, an age-adjusted measure of the average physical health in the county from 2018, the percent under the federal poverty line, log median household income, the unemployment rate, the 2016 Republican vote share, and the log total number of votes cast in 2016. We cluster standard errors at the DMA level and report 95 percent confidence intervals.

Figure D3: DMA-level block bootstrap



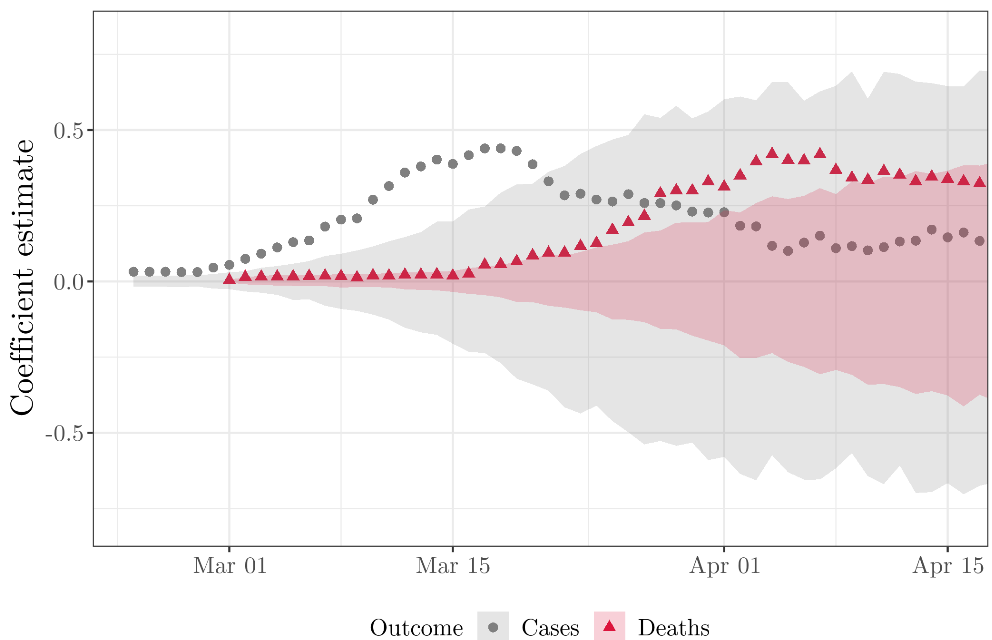
Notes: Figure D3 presents confidence intervals derived from a block bootstrapping procedure. We randomly sample DMAs with replacement and estimate counterfactual treatment effects for each day. We repeat 1000 times to calculate a distribution of counterfactual treatment effects for each day. Confidence intervals are calculated separately for each day: the upper boundary of the confidence interval corresponds to the 0.975-quantile of treatment effects on that day, while the lower boundary corresponds to the 0.025-quantile.

Figure D4: Randomization inference



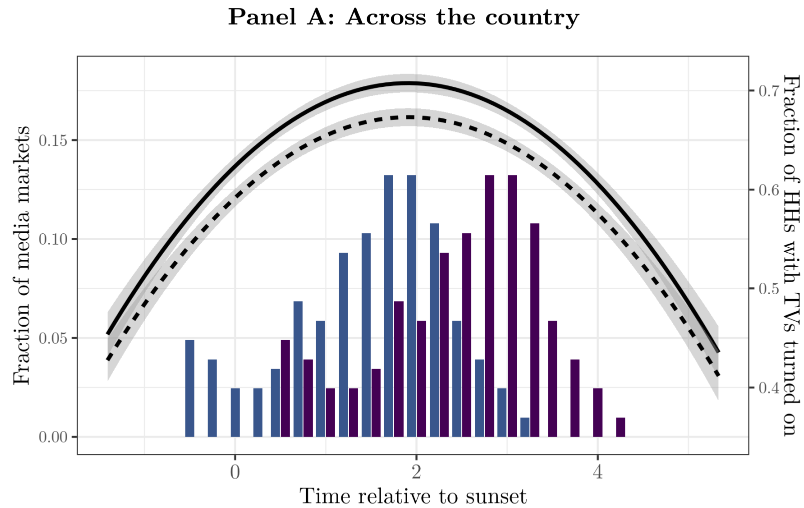
Notes: Figure D4 presents placebo treatment effects derived from a randomization inference procedure. We permute the plausibly exogenous "shift" ($\tilde{s}_{mc,H}$) across DMAs while leaving the "shares" (FoxShare$_d$), the county-level covariates, and cases and deaths unchanged. For each repetition, we then regenerate our instrument as the interaction of the placebo $\tilde{s}_{mc,H}$ with FoxShare$_d$, then calculate placebo treatment effects. We repeat 1000 times to calculate a distribution of counterfactual treatment effects for each day. The upper boundary of the shaded region corresponds to the 0.975-quantile of treatment effects on that day, while the lower boundary corresponds to the 0.025-quantile.

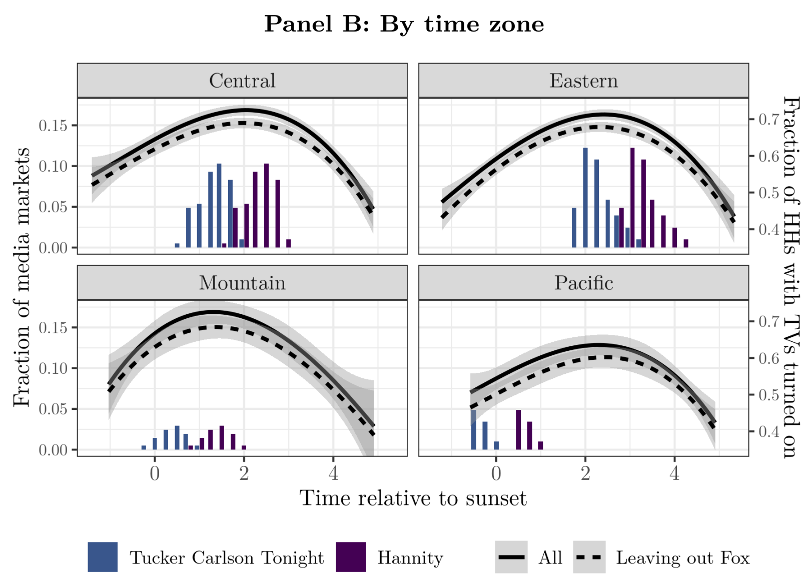
Figure D5: Permutation test



Notes: Figure D5 presents placebo treatment effects derived from a permutation test. We permute the joint tuple of cases and deaths across counties, leaving all other covariates unchanged, then estimate placebo treatment effects. We repeat 1000 times to calculate a distribution of counterfactual treatment effects for each day. The upper boundary of the shaded region corresponds to the 0.975-quantile of treatment effects on that day, while the lower boundary corresponds to the 0.025-quantile.

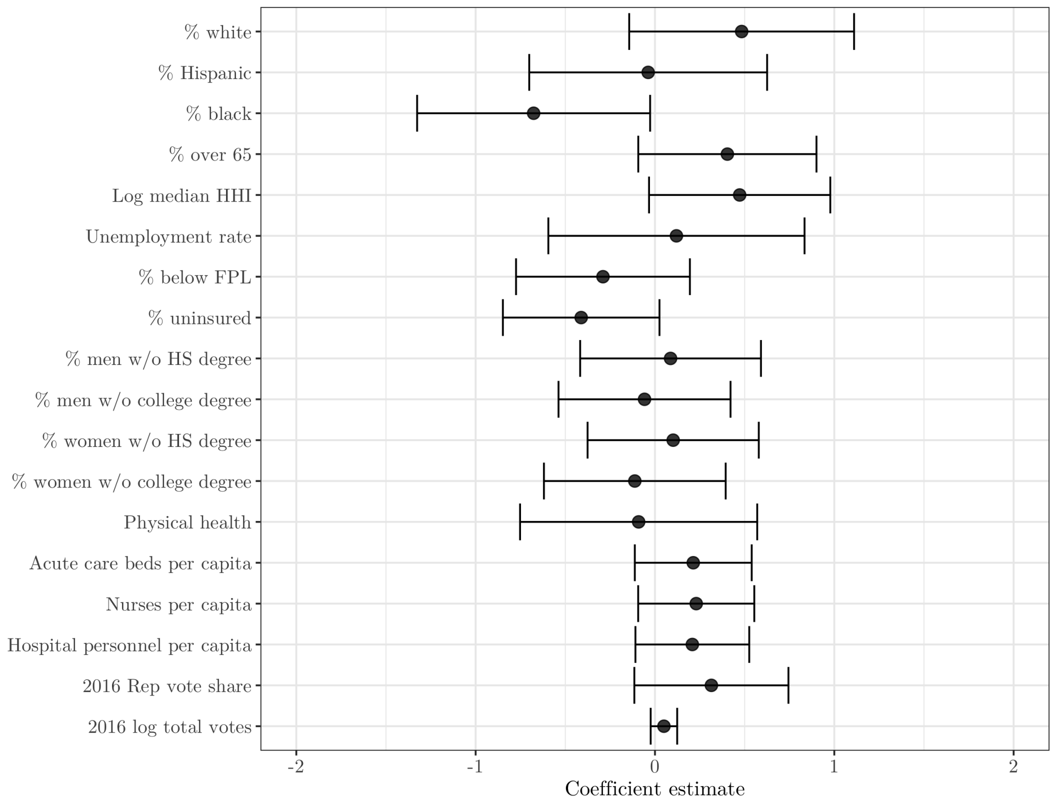
Figure D6: Viewership and program start relative to sunset





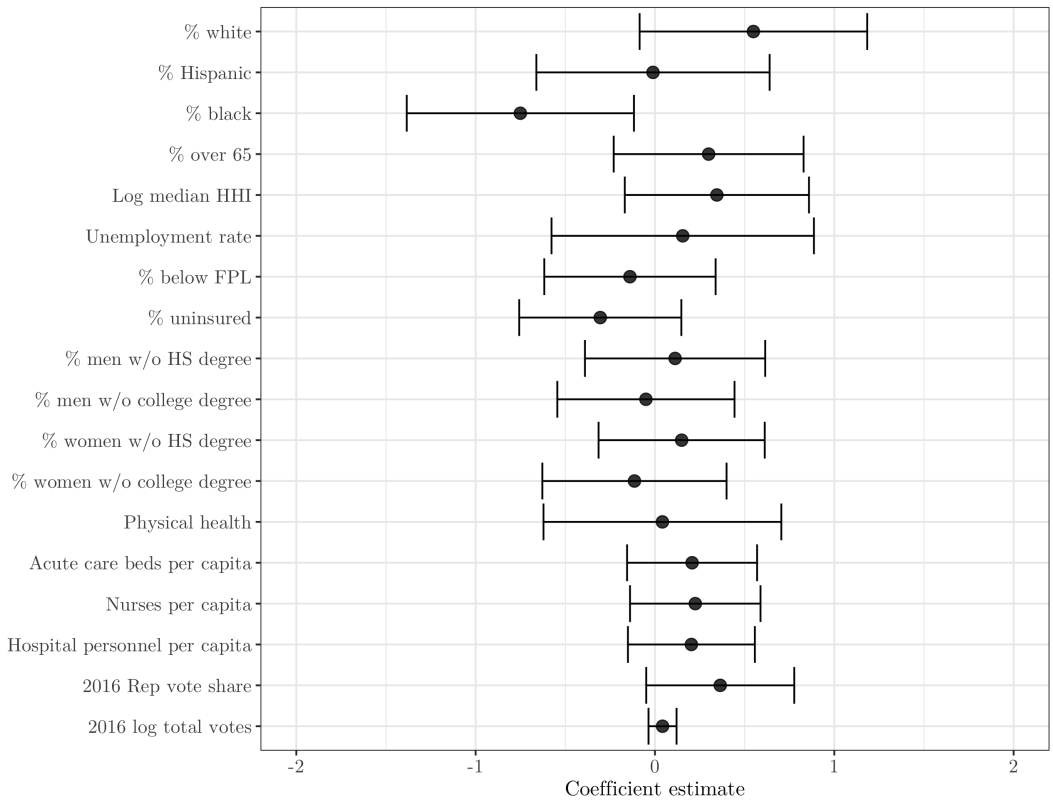
Notes: Panel A of Figure D6 plots a third-degree polynomial fitting the relationship between time since sunset in a DMA and the fraction of households in that DMA with TVs turned on (solid line) and the relationship between time since sunset and the fraction of households with TVs turned on and tuned to non-Fox channels (dashed line). 95% confidence intervals are reported. Panel A also shows a histogram depicting, at each fifteen-minute interval relative to sunset, the number of DMAs in which *Tucker Carlson Tonight* begins in that interval (blue) and in which *Hannity* begins in that interval (purple). Episodes of *Tucker Carlson Tonight* and *Hannity* are generally re-run three hours after they first air, and because our data spans 5pm to 11pm, we observe repeats in more western time zones but not in Eastern Time. Panel B is similar, but plots the relationship and histogram separately for each of the four major time zones in the continental United States.

Figure D7: Predicted viewership curve: correlation with county-level demographics



Notes: Figure D7 shows the coefficients from a series of regressions of each demographic characteristic on our instrument, $\widehat{\mathrm{s}}_{mc,H} \times \tilde{f}_{mc,-HT}$, conditional on the two interactants, $\widehat{\mathrm{s}}_{mc,H}$ and $\mathrm{FoxShare}_d$, and a small set of other controls accounting for local viewership patterns (the predicted share of TVs tuned to non-Fox channels during *Hannity*, *Tucker Carlson Tonight*, and *The Ingraham Angle*, the local viewership share of MSNBC, log population and population density, and population-weighted latitude and longitude). All dependent variables are scaled to a standard normal distribution. We cluster standard errors at the DMA level and report 95 percent confidence intervals.

Figure D8: Division viewership curve: correlation with county-level demographics



Notes: Figure D8 shows the coefficients from a series of regressions of each demographic characteristic on our instrument, $\widehat{s}_{mc,H} \times \tilde{f}_{mc,-HT}$, conditional on the two interactants and a small set of other controls accounting for local viewership patterns (the predicted share of TVs tuned to non-Fox channels during *Hannity*, *Tucker Carlson Tonight*, and *The Ingraham Angle*, the local viewership share of MSNBC, and population size and density). All dependent variables are scaled to a standard normal distribution. We cluster standard errors at the DMA level and report 95 percent confidence intervals.

Table D1: 2SLS estimates: robustness to choice of controls and instrument variations

	Dependent variable:					
	COVID-19 outcomes					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: *COVID-19 cases on March 14*						
H-C viewership difference (predicted)	0.360***	0.328***	0.373***	0.362***	0.786**	0.632**
	(0.099)	(0.093)	(0.106)	(0.104)	(0.336)	(0.270)
Panel B: *COVID-19 deaths on March 28*						
H-C viewership difference (predicted)	0.260***	0.243***	0.294***	0.288***	0.575**	0.539**
	(0.080)	(0.077)	(0.089)	(0.088)	(0.265)	(0.251)
F-statistic (Kleibergen-Paap)	19.22	10.04	20.45	10.35	6.60	3.33
Controls	Full	Full	Full	Full	Full	Full
Instruments	H	H&T	H	H&T	H	H&T
Instrument	Leave-out	Leave-out	Sunset	Sunset	Division sunset	Division sunset
State FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,100	3,100	3,100	3,100	3,100	3,100

Notes: Table reports 2SLS regressions of the log of one plus the number of cases on March 14 (Panel A) and the log of one plus the number of deaths on March 28 (Panel B) on the standardized difference between viewership of *Hannity* and *Tucker Carlson Tonight*. In Column 1, we instrument this difference by $\tilde{s}_{mc,H} \times \tilde{f}_{mc,-HT}$; in Column 2, we additionally instrument by $\tilde{s}_{mc,T} \times \tilde{f}_{mc,-HT}$ — that is, an analogous instrument for viewership during the *Tucker Carlson Tonight* timeslot. Columns 3-4 are identical to Columns 1-2, except that we use fitted rather than actual values of $\tilde{s}_{mc,H}$ (fitted based on sunset time, where the viewership curve is estimated at the DMA level): that is, the instruments are $\widehat{\tilde{s}_{mc,H}}_d \times \text{FoxShare}_d$ and $\widehat{\tilde{s}_{mc,T}}_d \times \text{FoxShare}_d$. Columns 5-6 are identical to Columns 1-2, except that we use fitted rather than actual values of $\tilde{s}_{mc,H}$ (fitted based on sunset time, where the viewership curve is estimated at the Census division level): that is, the instruments are $\widehat{\tilde{s}_{mc,H}}_d \times \text{FoxShare}_d$ and $\widehat{\tilde{s}_{mc,T}}_d \times \text{FoxShare}_d$. "Full controls" include the predicted share of TVs tuned to non-Fox channels during *Hannity*, *Tucker Carlson Tonight*, and *The Ingraham Angle*, Fox News' and MSNBC's share of cable in January 2018, Fox News' share of television in January population density and log population, population-weighted latitude and longitude, the percent of the population living in a rural area, the population over the age of 65, the percent male with no high school degree, the percent female with no high school degree, the percent male with no college degree, the percent female with no college degree, an age-adjusted measure of the average physical health in the county, the percent uninsured, the percent below the federal poverty line, the log of the median household income, the unemployment rate, the Republican vote share in 2016, and the log of the total number of votes in the county in 2016. As a test for weak instruments, we report first-stage Kleibergen-Paap F-statistics. Standard errors are clustered at the DMA level.

Table D2: Differential coverage and COVID-19 outcomes across all Fox News evening shows

	Dependent variable:					
	Inverse pandemic coverage index				Cases Mar 14	Deaths Mar 28
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: *OLS: inverse pandemic coverage index on relative viewership*						
H-C viewership difference	0.548***	0.545***				
	(0.053)	(0.052)				
Panel B: *RF: inverse pandemic coverage index on instrument*						
Non-Fox TVs on × Fox share			0.502**	0.490**		
			(0.230)	(0.227)		
Panel C: *2SLS: cases and deaths on inverse predicted pandemic coverage index*						
−1× coverage index (predicted)					0.776**	0.538*
					(0.364)	(0.281)
Controls	Base	Full	Base	Full	Full	Full
State FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,102	3,102	3,102	3,102	3,102	3,102

Notes: Panel A reports OLS estimates of the (inverse of the) pandemic coverage index on the standardized difference between viewership of *Hannity* and *Tucker Carlson Tonight*. Panel B reports reduced-form estimates of the inverse pandemic coverage index on our instrument, $\tilde{s}_{mc,H} \times \tilde{f}_{mc,-HT}$ — that is, the number of TVs on during Hannity's timeslot based on other DMAs in the same time zone, excluding TVs watching *Hannity*, multiplied by Fox News' viewership share, excluding *Hannity* and *Tucker Carlson Tonight.*. Columns (5) and (6) in Panel C report 2SLS estimates of the log of one plus the number of cases on March 14 and the log of one plus the number of deaths on March 28, respectively, on the standardized difference between viewership of *Hannity* and *Tucker Carlson Tonight*, instrumented by $\tilde{s}_{mc,H} \times \tilde{f}_{mc,-HT}$. Base OLS controls include the share of TVs tuned to non-Fox channels during *Hannity*, *Tucker Carlson Tonight*, and *The Ingraham Angle*, Fox News' and MSNBC's share of cable in January 2018, Fox News' share of television in January 2020, the population density of the county, and the log of the county's total population. Base controls for the reduced form and the two-stage least squares are identical, except the share of TVs tuned to non-Fox channels during *Hannity*, *Tucker Carlson Tonight*, and *The Ingraham Angle* are replaced with the predicted share of TVs tuned to non-Fox channels during these timeslots. 'Full controls" additionally include all controls described in Section 3. Standard errors are clustered at the DMA level.

E Assessing Effect Sizes

We now assess our estimated effect sizes through a simple epidemiological model. The key behavioral foundation is that *Hannity* and *Tucker Carlson Tonight* influence the behavior of viewers by changing their beliefs about the threat posed by COVID-19, thus influencing the extent to which they take precautionary measures (such as washing hands or disinfecting more frequently) and in turn affect the disease transmission rate among viewers.[61]

Our model allows us to estimate the extent to which the shows would need to affect transmissibility among viewers in order to generate treatment effects similar in magnitude to those we estimate. Our goal is not to point-identify structural parameters of the model: estimating models of the COVID-19's spread is notoriously difficult (as evidenced by the wide variance in model predictions from different sources over the course of the pandemic) and there may be several sets of parameters that fit the data; and moreover, our identification strategy does not allow us to account for inter-county externalities, a crucial element in explaining the virus' spread (Kuchler et al., 2020). Instead, we view our exercise as a back-of-the-envelope calculation to evaluate whether our observed treatment effects on deaths are consistent with reasonable changes in disease transmissibility.

Basic SIR (Susceptible-Infected-Removed) models, or most standard variants thereof, do not allow for heterogeneous groups that differ in their mortality or transmission rates. We wish to account for heterogeneity in age, since the elderly both have elevated COVID-19 fatality rates and are disproportionately likely to watch Fox News. Indeed, former CDC director Tom Frieden described nurinsg homes as "ground zero" of the pandemic, pointing out that nursing homes were likely to be hotbeds of COVID-19 contagion not only among residents but also among staff and visitors.[62] Academic work has similarly emphasized the role of the elderly in spreading COVID-19 (Davidson and Szanton, 2020; Barnett and Grabowski, 2020).

We also wish to account for heterogeneity in viewership of *Tucker Carlson Tonight* and *Hannity*, since only a fraction of the population are exposed to these shows and an even smaller fraction are "treated" (in the sense of being shifted into watching more *Hannity* relative to *Tucker Carlson Tonight* by our instrument inducing a one standard deviation increase in relative viewership).

We thus adapt the multi-group SIR model introduced in Acemoglu et al. (2020) to model four groups: the "untreated" population between 25 and 64 (of size N_{yu}); the "treated" population between 25 and 64 (of size N_{yt}); the "untreated" population aged 65 and older (of size N_{ou}); and the "treated" population aged 65 and older (of size N_{ot}). We calibrate N_j using ACS data on the age distribution of the US population alongside our Nielsen data on daily viewership and our

[61] Viewership of *Hannity* and *Tucker Carlson Tonight* may also affect transmissibility through indirect channels. For example, these shows might change social norms associated with behavior such as wearing masks, temporarily closing businesses, and providing employees with sick leave (Shadmehr and de Mesquita, 2020), or, relatedly, viewers might share the information they learned on the shows with others. For simplicity, we do not model these channels.

[62] See "Former CDC director: It's time to restrict visits to nursing homes," *CNN*, March 8, 2020.

survey data on viewership frequency.[63] Following Acemoglu et al. (2020), we normalize the total population size $N = \sum_j N_j$ to 1. [64] We assume that death and recovery rates are invariant to time and the number of patients. To capture differential interaction patterns — the fact that young agents are more likely to interact with other young agents (e.g. through the workplace) while old agents are more likely to interact with old agents (e.g. in nursing homes), we calibrate the interaction matrix ρ using the intergenerational interaction matrix from Akbarpour et al. (2020).[65] While age affects the probability of interaction between groups, treatment status does not: conditional on age, a treated person is equally likely to interact with another treated person as with an untreated person. Following Allcott et al. (2020a), we model the effect of cautious behaviors such as washing hands, wearing face masks, or social distancing — and thus, the effect of differential viewership of *Hannity* and *Tucker Carlson Tonight* — by assuming that they directly affect the transmission rate β_j.[66]

Denoting the susceptible, infected, recovered, and dead populations by S, I, R, and D, respectively, the model is characterized by the following system of differential equations:

$$\begin{aligned}
\dot{I}_j &= S_j \left(\sum_k c(\beta_j, \beta_k) \rho_{jk} I_k \right) - \gamma_j I_j - \delta_j I_j \\
\dot{R}_j &= \gamma_j I_j \\
\dot{D}_j &= \delta_j I_j \\
\dot{S}_j &= -\dot{I}_j - \dot{R}_j - \dot{D}_j
\end{aligned}$$

To fix notation, let $\bar{X}$ denote the value of variable X in a representative county with a mean viewership of *Hannity* relative to *Tucker Carlson Tonight*, and let X^+ denote the value of X in a representative county with a one standard deviation higher viewership of *Hannity* relative to *Tucker Carlson Tonight*. By construction, there is no "treated" population in the county with mean

[63] As in our survey analysis, we include "occasional" viewers (those who watch the shows between one and three times per week) alongside "regular" viewers (those who watch four or five times per week).

[64] We make a number of additional parameter assumptions to make the model more tractable. In particular, we assume $\alpha = 2$ (quadratic matching in transmission, which most closely matches the dynamics of a standard SIR model); and we abstract away from healthcare capacity constraints by assuming that $\iota = 1$.

[65] The matrix is based on data provided by Replica, which uses anonymized cellphone GPS data to simulate a "synthetic population" that "closely approximates both age and industry distributions from the Census ACS, as well as granular ground-truth data on mobility patterns from a variety of different sources" (Akbarpour et al., 2020).

[66] Thus, in contrast to Acemoglu et al. (2020), there is no single transmission rate β governing the probability by which a susceptible agent will be infected when they come into contact with an infected agent; this rate is an increasing function c in the β_j parameters of the infected agent and the susceptible agent. To our knowledge, there are no estimates of $c(\cdot, \cdot)$ for COVID-19. For tractability, we assume that when agents from groups a and b with $\beta_a \neq \beta_b$ come into contact, the "effective transmission rate" is given by $c(\beta_a, \beta_b) = \max\{\beta_a, \beta_b\}^2$, intuitively capturing the intuition that it is the less cautious agent that drives the transmission probability. For example, the primary benefit of face masks is that they help prevent infected people from spreading COVID-19 to others; they are less effective in protecting the wearer against contracting COVID-19 from others (Bai, 2020). However, our results are qualitatively similar if we instead assume $c(\beta_a, \beta_b) = \beta_a \beta_b$.

relative viewership: $\bar{N}_{yt} = \bar{N}_{ot} = 0$, $\bar{N}_{yu} = N^+_{yu} + N^+_{yt}$, $\bar{N}_{ou} = N^+_{ou} + N^+_{ot}$. Also by construction, transmissibility in the county with mean relative viewership is always equal to transmissibility among untreated in the county with a one standard deviation higher relative viewership: $\bar{\beta}_{yu}(t) = \bar{\beta}_{ou}(t) = \beta^+_{yu}(t) = \beta^+_{ou}(t)$, for all t. To ease notation, we write $\bar{\beta} := \bar{\beta}_{yu} = \bar{\beta}_{ou}$, $\beta^+_u := \beta^+_{yu} = \beta^+_{ou}$, $\beta^+_t := \beta^+_{yt} = \beta^+_{ot}$. We report all parameter values in Table E1.

We take the timing of behavioral changes in response to COVID-19 from our survey, which are presented in Panel B of Figure 3, as primitives in our model. The treatment effect of *Hannity* viewership relative to *Tucker Carlson Tonight* viewership on the total number of people who report having changed their behavior to act more cautiously in response to COVID-19 is approximately 0 on February 1, increases to peak on March 1, and then decreases. The difference had not yet returned to zero by the date of the survey, but assuming the observed trend continued, we would expect it to return to zero by mid-April. We thus fix $\bar{\beta}(t) = \beta^+_n(t) = \beta^+_c(t)$ for $t =$ Feb 1 and $t \geq$ Apr 15. Since, in our survey, both the increase in estimated treatment effects between February 1 and March 1 and the decrease between March 1 and April 3 are approximately linear, we linearly interpolate values of β between February 1 and March 1 and between March 1 and April 15. Informed by recent epidemiological estimates (e.g., Unwin et al. 2020), we allow the transmission rate to decline linearly from April 15 to May 1. This leaves us with five parameters to estimate: $\bar{\beta}$(Feb 1) $= \beta^+_u$(Feb 1) $= \beta^+_t$(Feb 1), $\bar{\beta}$(Mar 1) $= \beta^+_u$(Mar 1), β^+_t(Mar 1), $\bar{\beta}$(Apr 15) $= \beta^+_u$(Apr 15) $= \beta^+_t$(Apr 15), and $\bar{\beta}$(May 1) $= \beta^+_u$(May 1) $= \beta^+_t$(May 1).

COVID-19 cases were vastly underreported, with some preliminary estimates suggesting that as many as 93% of cases may be undetected (Stock et al., 2020). This was particularly true in the United States, which suffered from testing shortages during the early stages of the pandemic.[67] As a result, we focus on fitting the trajectories of *deaths* implied by our coefficient estimates. We proceed by simulating death trajectories under different values of parameters, selecting the combination that minimizes a loss function based on the sum of squared residuals between the 2SLS estimates and the simulated trajectories.[68]

Panel A of Figure E2 plots the fitted trajectories of β for the untreated (which comprise the entire county with a mean viewership difference and the vast majority of the county with a one standard deviation higher viewership difference) and for the treated (the remaining fraction of the county with a one standard deviation higher viewership difference).[69] The peak difference in $\bar{\beta}$ and β^+_t on March 1 is approximately 27%.[70] The estimated paths imply that the treated

[67] See, for example, "Why America's coronavirus testing barely improved in April", *The New York Times*, May 1, 2020.

[68] We begin our simulations on February 1, five days before the day of the first confirmed COVID-19-related death in the US (see "First Known U.S. COVID-19 Death Was Weeks Earlier Than Previously Thought", *NPR*, April 22, 2020.)

[69] We repeat this exercise for our OLS estimates; the results are reported in Appendix Figure E1.

[70] This difference is approximately equal to the March 1 persuasion rate we identify from the survey data (24.1%), though the two estimates are of course not directly comparable. Weighting by the size of each group, the maximum

population did not adjust their behavior at all throughout most of February and only began doing so in March, while the non-treated population gradually adjusted behavior throughout the period before the April 15 convergence. For ease of comparison with other studies, we can also calculate the trajectories of the effective reproduction number R_t: the expected number of susceptible individuals an individual infected at time t will him or herself infect. At $t = 0$, this is approximated by $R_0 \approx \frac{\beta^2}{\gamma} = 3.18$; R_t falls to approximately 1.81 by April 15 among the untreated and approximately 1.15 among both groups by May 1. These values are broadly similar to recent estimates of the effective reproduction rate, e.g. Atkeson et al. (2020).

Panel B of Figure E2 plots the simulated treatment effect and the estimated 2SLS treatment effects. Our model fits the estimated treatment effects fairly well. Adding additional degrees of freedom by modeling agent heterogeneity, "super-spreader" events, and network structure would allow us to better fit the shape of estimated treatment effects (McGee, 2020), but these are beyond the scope of our exercise.

Our model also allows us to examine what fraction of people who died were members of the treated group, i.e. the group whose transmissibility was affected by a one standard deviation increase in relative viewership. We estimate that approximately 5% of the additional deaths occur in the treated group, with the additional deaths occurring in the untreated group. Since there is substantial uncertainty about the true values of the exogenously taken input parameters of the model, and since our model fails to capture important features such as county-to-county spillovers, we should be cautious when interpreting this estimate. Nonetheless, the model highlights the relevance of externalities in generating our estimated treatment effects. This is broadly consistent with evidence that although symptomatic COVID-19 cases and COVID-19 deaths were concentrated among the elderly, incidents of transmission "were the highest among school-aged children, between children and their parents, and between middle-aged adults and the elderly" (Monod et al., 2021).[71]

Taken together, our results suggest that behavioral responses among viewers early on in a pandemic – due to differential media coverage of the virus – can give rise to modest but meaningful differences in transmissibility among the broader population, which ultimately translate into effect sizes of roughly the same magnitude as those we estimate.

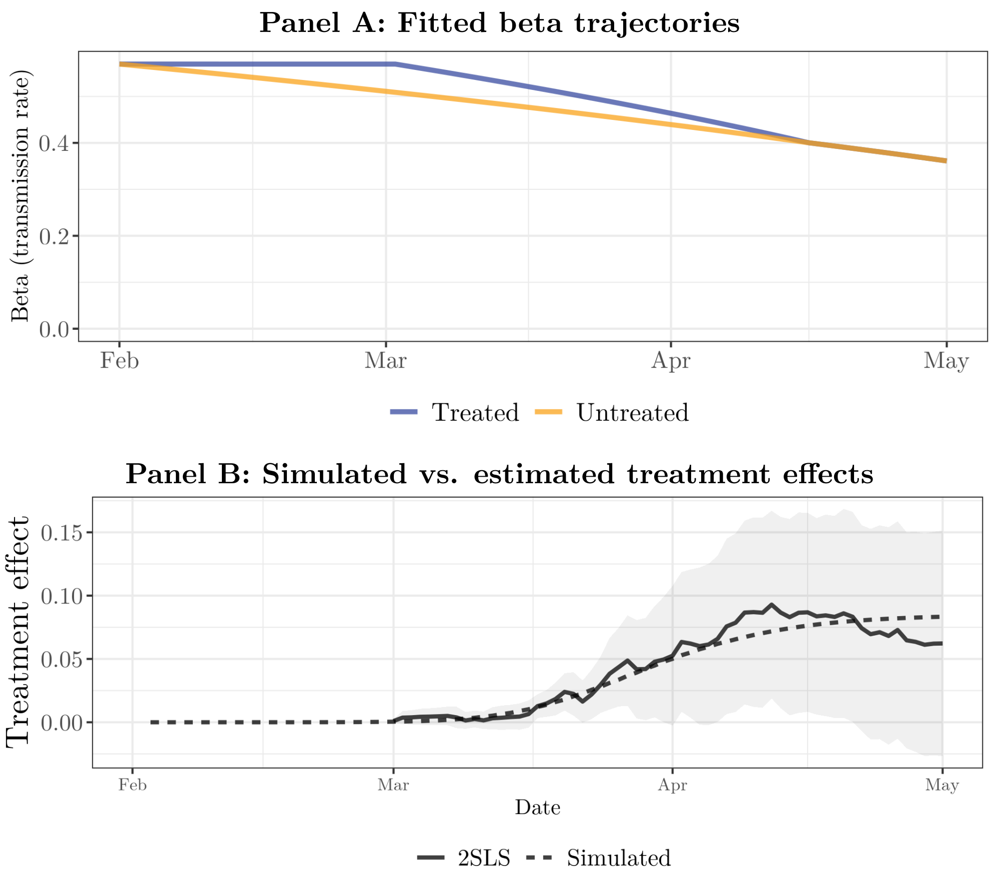
---

difference in the *average* beta in the county with a mean viewership difference vs. the county with a 1 SD higher viewership difference is around 2%.

[71] Our results are in line with those of Banerjee et al. (2020), which also finds large spillovers in the context of health behaviors during the COVID-19 pandemic.

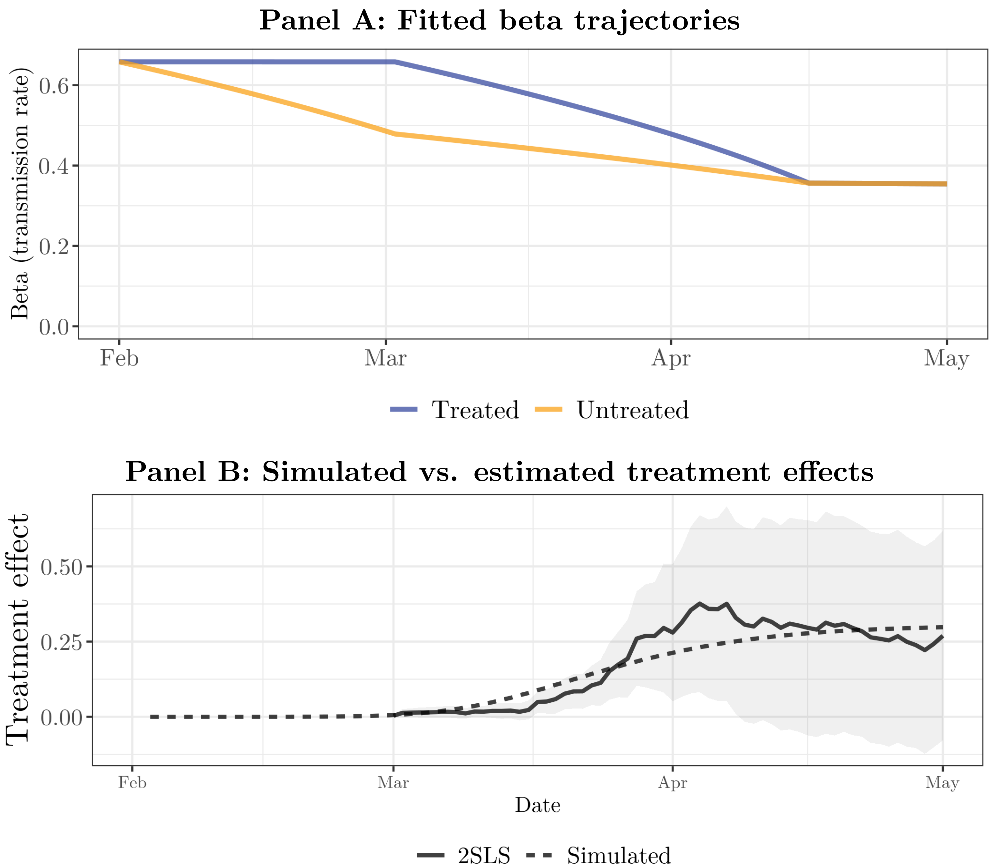
E.1 Figures and Tables

Figure E1: MG-SIR simulations (OLS)



Notes: Panel A of Figure E1 plots, in orange, the β trajectory implied by our simulation for non-compliers (which comprise the entire county with a mean viewership difference and the vast majority of the county with a one standard deviation higher viewership difference) and, in blue, the corresponding trajectory for compliers (which comprise the remaining fraction of the county with a one standard deviation higher viewership difference). Panel B plots the simulated treatment effect and the estimated treatment effects. lines.

Figure E2: MG-SIR simulations (2SLS)



Notes: Panel A of Figure E2 plots, in orange, the β trajectory implied by our simulation for non-compliers (which comprise the entire county with a mean viewership difference and almost the entire county with a one standard deviation higher viewership difference) and, in blue, the corresponding trajectory for compliers (which comprise the remaining population of the county with a one standard deviation higher viewership difference). Panel B plots the simulated treatment effect and the estimated treatment effects.

Table E1: Exogenous model parameters

Parameter	**Description**	**Value**	**Source**
P_o	Share of simulated population above the age of 65	0.3216	American Comunity Survey (ACS)
$\bar{N}_{yt}$	Share of treated among young in representative county with mean viewership	0	
$\bar{N}_{ot}$	Share of treated among old in representative county with mean viewership	0	
N_{yt}^{+}	Share of treated among young in representative county with 1 SD higher viewership	0.0097	Nielsen
N_{ot}^{+}	Share of treated among old in representative county with 1 SD higher viewership	0.0112	Nielsen
$i(0)$	Initial fraction of infected individuals	3.030×10^{-8}	Estimated 10 infections in US on Feb 1
$I_j(0)$	Initial share of infected individuals in group j	$i(0) \times N_j$	
$S_j(0)$	Initial share of susceptible individuals in group j	$N_j - I_j$	
$R_j(0)$	Initial share of recovered individuals in group j	0	
$D_j(0)$	Initial share of dead individuals in group j	0	
γ	Estimated recovery arrival rate	0.125	Allcott et al. (2020) (derived)
δ_y	Estimated fatality arrival rate among young individuals	6.354×10^{-4}	Ferguson et al. (2020) (derived)
δ_o	Estimated fatality arrival rate among older individuals	0.0101	Ferguson et al. (2020) (derived)
α	"Returns to scale" in matching of individuals	2.000	Acemoglu et al. (2020)
ρ	Matrix of group interaction rates (first row/column for young, second for old)	$\begin{bmatrix} 1.51 & 0.57 \\ 0.53 & 0.47 \end{bmatrix}$	Akbarpour et al. (2020)

F Survey Instrument

F.1 Survey Experiment: Trust in Opinion

F.1.1 Fox News

At the end of this survey, we will ask you a question about **median weekly earnings in the US from recent years**. You will have only a few seconds to answer this question, so you will not have the chance to look up the answers online: please just make your best guess.

If your answer lies within 5 percent of the official value, you will win a $10 Amazon gift card. We will ask you to provide your email at the end of the survey so that we can send you your card in case you make an accurate guess. If you prefer not to provide your email address, you can leave the box blank, but we will be unable to contact you to send you your gift card if you win. We will delete all email addresses immediately after the completion of the study.

To help you make your guess, you will be able to choose one of a few video clips of TV shows. These shows were all broadcast on the same date, which we selected at random. We will ask you to make a guess about **median weekly earnings in the US prior to the airing of the show**. We will not tell you about the date of the show, so you will have to make your best guess solely based on the show content.

Please pick which clip you would like to watch:

- ○ Special Report (with Brett Baier)
- ○ The Story (with Martha MacCallum)
- ○ Tucker Carlson Tonight (with Tucker Carlson)
- ○ Hannity (with Sean Hannity)

We will show you a clip from the TV show you indicated. Before watching the clip, please answer a few questions:

Why did you choose the show that you did? Please answer in at least 2 sentences.

→

If you had to guess, what do you think is the probability that the clip from each show contains the information you need to make your guess?

0: Definitely does not contain the information 100: Definitely contains the information

0 10 20 30 40 50 60 70 80 90 100

Tucker Carlson Tonight (with Tucker Carlson)

Special Report (with Brett Baier)

Hannity (with Sean Hannity)

The Story (with Martha MacCallum)

The below link leads you to the show Tucker Carlson Tonight. Please return to this page when you are done watching.

https://archive.org/details/FOXNEWSW_20200808_000000_Tucker_Carlson_Tonight/start/2400/end/2580

→

Did you watch the entire segment linked on the previous page?

- Yes, I watched the entire segment
- No, I watched only parts of the segment
- No, I did not watch any parts of the segment

→

On the next page, we will ask you to make a guess about the economy at the time of the show's airing. You will only have 10 seconds to make the guess.

→

What do you think?

According to the Bureau of Labor Statistics, what were median weekly earnings in the US **prior to the airing of the show?**

06

→

F.1.2 MSNBC

At the end of this survey, we will ask you a question about **median weekly earnings in the US from recent years**. You will have only a few seconds to answer this question, so you will not have the chance to look up the answers online: please just make your best guess.

If your answer lies within 5 percent of the official value, you will win a $100 Amazon gift card. We will ask you to provide your email at the end of the survey so that we can send you your card in case you make an accurate guess. If you prefer not to provide your email address, you can leave the box blank, but we will be unable to contact you to send you your gift card if you win. We will delete all email addresses immediately after the completion of the study.

To help you make your guess, you will be able to choose one of a few video clips of TV shows. These shows were all broadcast on the same date, which we selected at random. We will ask you to make a guess about **median weekly earnings in the US prior to the airing of the show**. We will not tell you about the date of the show, so you will have to make your best guess solely based on the show content.

Please pick which clip you would like to watch:

- ◯ The Last Word (with Lawrence O'Donnell)
- ◯ The Beat (with Ari Melber)
- ◯ MSNBC Live
- ◯ The Rachel Maddow Show

We will show you a clip from the TV show you indicated. Before watching the clip, please answer a few questions:

Why did you choose the show that you did? Please answer in at least 2 sentences.

→

If you had to guess, what do you think is the probability that the clip from each show contains the information you need to make your guess?

0: Definitely does not contain the information | 100: Definitely contains the information

0 10 20 30 40 50 60 70 80 90 100

The Rachel Maddow Show

MSNBC Live

The Beat (with Ari Melber)

The Last Word (with Lawrence O'Donnell)

The below link leads you to the show MSNBC Live. Please return to this page when you are done watching.

https://archive.org/details/MSNBCW_20200807_190000_MSNBC_Live/start/120/end/300

→

Did you watch the entire segment linked on the previous page?

- ◯ Yes, I watched the entire segment
- ◯ No, I watched only parts of the segment
- ◯ No, I did not watch any parts of the segment

→

On the next page, we will ask you to make a guess about the economy at the time of the show's airing. You will only have 10 seconds to make the guess.

→

What do you think?

According to the Bureau of Labor Statistics, what were median weekly earnings in the US **prior to the airing of the show?**

01

→

F.2 Behavioral Change Survey

Demographics questions

What is your exact age?

What is your gender?

- Male
- Female

With which political party do you identify?

- Democratic Party
- Republican Party
- Independent

Do you have a job outside of taking surveys?

- ◯ Yes: full-time (35+ hours a week)
- ◯ Yes: part-time (less than 35 hours a week)
- ◯ No: homemaker
- ◯ No: currently seeking employment
- ◯ No: student
- ◯ No: retired
- ◯ No: other

What was your family's gross household income in 2019 in US dollars?

- ◯ Less than $15,000
- ◯ $15,000 to $24,999
- ◯ $25,000 to $49,999
- ◯ $50,000 to $74,999
- ◯ $75,000 to $99,999
- ◯ $100,000 to $149,999
- ◯ $150,000 to $200,000
- ◯ More than $200,000

Which of the following best describes your race or ethnicity?

- African American/Black
- Asian/Asian American
- Caucasian/White
- Native American, Inuit or Aleut
- Native Hawaiian/Pacific Islander
- Other

Are you of Hispanic, Latino, or Spanish origin?

- Yes
- No

What is the highest level of education you have completed or the highest degree you have received?

- Less than high school degree
- High school graduate (high school diploma or equivalent including GED)
- Some college but no degree
- Associate degree in college (2-year)
- Bachelor's degree in college (4-year)
- Master's degree
- Doctoral degree
- Professional degree (JD, MD)

→

F.3 Media Consumption Questions

Which, if any, of the following major TV news stations do you watch at least once a week?

- [ ] CNN
- [ ] MSNBC
- [ ] Fox News
- [ ] Other

→

F.3.1 Fox News

You indicated that you watch Fox News at least once a week. How often do you watch each of the following shows on Fox News?

	Never	Occasionally	Every day or most days
Sean Hannity	○	○	○
The Ingraham Angle	○	○	○
Other Fox show	○	○	○
The Five	○	○	○
The Story with Martha MacCallum	○	○	○
Tucker Carlson	○	○	○

→

F.3.2 CNN News

You indicated that you watch CNN at least once a week. How often do you watch each of the following shows on CNN?

	Never	Occasionally	Every day or most days
Anderson Cooper 360	○	○	○
Erin Burnett OutFront	○	○	○
CNN Tonight	○	○	○
Cuomo Prime Time	○	○	○
Other CNN show	○	○	○
The Situation Room	○	○	○

→

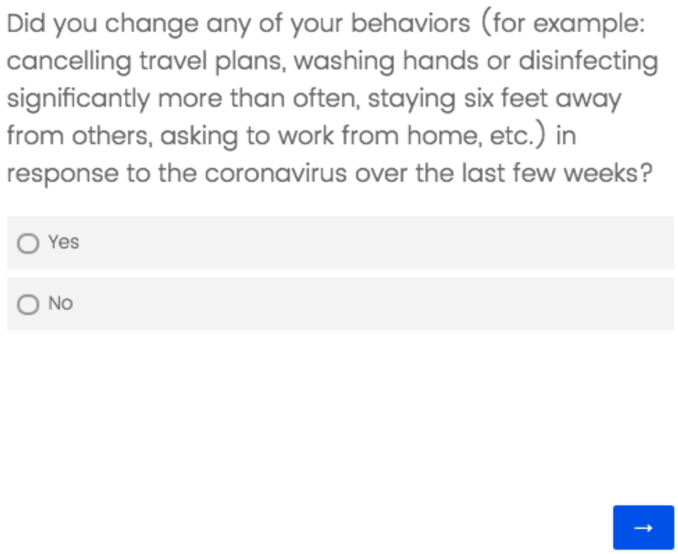
F.3.3 MSNBC News

You indicated that you watch MSNBC at least once a week. How often do you watch each of the following shows on MSNBC?

	Never	Occasionally	Every day or most days
The Beat with Ari Melber	○	○	○
Other MSNB show	○	○	○
All In with Chris Hayes	○	○	○
The Last Word with Lawrence O'Donnell	○	○	○
The 11th Hour with Brian Williams	○	○	○
The Rachel Maddow Show	○	○	○

→

F.4 Behavior Change Questions



Did you change any of your behaviors (for example: cancelling travel plans, washing hands or disinfecting significantly more than often, staying six feet away from others, asking to work from home, etc.) in response to the coronavirus over the last few weeks?

- Yes
- No

→

When did you first significantly change any of your behaviors (For example, cancelling travel plans, washing hands or disinfecting significantly more than often, staying six feet away from others, asking to work from home, etc.) in response to the coronavirus? How did you change your behavior? Why did you change your behavior?

On which date, did you first significantly change any of your behaviors in response to the coronavirus? (For example, cancelling travel plans, washing hands or disinfecting significantly more than often, staying six feet away from others, asking to work from home, etc.).

	Month	Day
Date of change in behavior		

→

F.5 Post-Outcome Questions

What is your zipcode of residence?

→

Thank you very much participating in this survey. If you have any comments, please let us know below.

→