Disagreement Is Disappearing on U.S. Cable Debate Shows

Abstract

Prime-time "debate" shows claim to air competing viewpoints, yet no large-scale evidence quantifies how often hosts and guests actually disagree. Measuring such exchanges is difficult: live broadcasts contain overlapping speakers, sarcasm, and billions of captioned words. We construct the first speaker-resolved map of agreement and disagreement across U.S. cable opinion programming. The study assembles 17,000 episodes spanning from 2010–2024 from 24 Fox News, MSNBC, and CNN shows, segments them into host–guest turns, and labels 2.13 million turn-pairs using a large-language-model classifier. We present three findings: (1) the proportion of disagreement/debate on prime time shows a consistent downward trend, dropping by roughly one-third between 2017 and 2024; (2) on-air challenge is partisan and asymmetric—conservatives seldom face push-back on Fox, liberals seldom on MSNBC, with CNN declining toward the midpoint; (3) polarizing issues such as abortion, gun rights, and immigration attract the least disagreement. The work contributes a public corpus, an open-source stance pipeline, and the first longitudinal evidence that televised "debate" is retreating from genuine discussion, which is essential for a thriving democracy.

1 Introduction

Televised debate shows—programmes built around a host who interviews or spars with one or more guests about the political story of the day—are advertised as democracy's front-row seats to an argument. These broadcasts occupy the most valuable real estate in cable schedules: the prime-time block (8p.m.–11p.m. Eastern), when total viewership peaks and advertising rates spike. In 2024, Fox News, MSNBC, and CNN together delivered roughly forty hours of such programming each week, and seven of the ten highest-rated cable slots were Fox opinion shows, each drawing well over two million nightly viewers (Littleton 2024). Because the audience is both large and politically attentive, elites treat these discussions as signals—Presidents quote them, lawmakers schedule hearings around them, and social-media influencers amplify their clips (Gertz 2020). If the "debate" has quietly become one-sided affirmation, the entire information chain downstream is fed from an ideological echo.

Copyright © 2025, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

The puzzle, then, is whether prime-time debate still fulfills its deliberative promise or whether it has degenerated into choreographed agreement that deepens what political scientists call *affective polarization*—the widening gulf of distrust and dislike between partisans that now rivals disagreement over concrete policy (Iyengar and Hahn 2009). Scholars have shown that exposure to like-minded news encourages attitude reinforcement (Stroud 2010), while diversified media diets can moderate views (Mutz and Reeves 2005). Yet almost all evidence rests on who watches which channel or how often a partisan guest is booked, not on whether a conservative and a liberal, once seated opposite the same host, actually clash over substance. Without that micro-level record, we cannot know whether televised debate attenuates or accelerates polarization.

Capturing that record is hard for three reasons. First, live television is unscripted and multi-speaker: turns overlap, tone shifts mid-sentence, and topics pivot quickly. Second, genuine disagreement is a semantic property, not a surface cue; a sarcastic "right, exactly" is opposition, not assent. Naïve proxies—assuming every Democrat on Fox is challenged or every negative adjective signals dissent—fail on counter-examples. Third, the scale is daunting. A single year of prime-time opinion programming on the big three cable networks generates more than a billion words of caption text (Hong et al. 2021). Manual annotation at the speaker-turn level is financially untenable, and existing automated pipelines either collapse all speech into one stream (losing who said what) or ignore rhetorical stance altogether.

Previous studies therefore stop short of the core question. Some sample a handful of episodes and code them by hand, offering depth without generality (Farnsworth and Lichter 2005). Others scrape guest lists or sentiment polarity and infer that a segment was adversarial when the guest's party differs from the host's party (**?**). Such approaches miss intra-party dissent, rhetorical framing, and shifts within a single interview. What we lack is a comprehensive, speaker-resolved measurement of agreement and disagreement across the contemporary cable landscape.

This paper supplies that measurement. We assemble a corpus of 17,000 episodes (approximately 8 000 hours of video) aired between 2010 and 2024 on 24 flagship opinion shows across Fox News, MSNBC, and CNN. An automatic-speech-recognition and speaker-diarisation

pipeline segments each broadcast into host and guest turns. Adjacent host–guest pairs are then classified as *agreement*, *neutral*, or *disagreement* by a high fidelity stance model which is then applied to classify over 2 million host-guest interaction turns. Finally, we identify the names and political party affiliations of the guests along with the detailed topic being discussed in each episode.

Using this rich dataset, we present five main findings. First, explicit host-guest dissent is rare: across 2017–2024 barely 15% of exchanges register as disagreement, with MSNBC the lowest at roughly 13% and Fox the highest at 17%. Second, dissent is in retreat. On Fox and CNN the share of disagreement turns falls sharply each year, and in the few MSNBC shows with long-run data the baseline is flat but never high, signaling a network-wide drift toward concord. Third, the decline is historically contingent rather than inevitable—archival captions from the early 2010s show *Hannity* operating at nearly double today's disagreement rate, evidence that the format once accommodated sharper clash. Fourth, gatekeeping amplifies the trend: the channels not only book ideologically friendly guests but also moderate them differently once on air—Republican invitees face the stiffest push-back on MSNBC, while Democrats encounter the gentlest treatment there, with the mirror pattern on Fox and a narrowing middle on CNN. Finally, the topics that dominate contemporary culture-war politics—abortion, gun rights, immigration—attract the *least* disagreement, turning segments most likely to mobilize audiences online into echo chambers on air.

Together, these results suggest that prime-time "debate" is increasingly a performance of consensus within partisan boundaries rather than a forum for genuine argumentative contestation. By documenting this shift at scale and at the level of individual utterances, the study clarifies how televised political talk contributes to the broader architecture of affective polarization, and why restoring actual disagreement to mass media remains a pressing democratic task.

2 Related work

A vast literature links partisan cable news consumption to attitudinal polarization, uncivil discourse, and diminished civic trust. We organize the most relevant findings into four strands: (1) selective exposure and the rise of echo chambers, (2) the evolution of televised debate, (3) democratic consequences of incivility and one-sided programming, and (4) recent computational audits that map ideological drift at scale. Together they illustrate both what we know and the empirical gaps this study fills.

Selective Exposure and the Cable-News Echo Chamber. Once Americans could choose among dozens of channels, many gravitated toward ideologically congenial outlets, producing self-reinforcing "echo chambers." Laboratory evidence shows Republicans strongly prefer Fox-branded stories while Democrats shun them (Iyengar and Hahn 2009). Longitudinal survey work demonstrates that this sorting hardens attitudes over time (Stroud 2010; Levendusky 2013). Although only a minority of viewers are "active-audience" partisans, their loyalty shapes network programming and, by extension, public debate (Arceneaux and Johnson 2013). Recent large-scale audience data confirm that cable news consumers today almost never stray outside their partisan bubble, with Fox, CNN, and MSNBC audiences sharply segregated (Kim, Lelkes, and McCrain 2022).

From "Shout-Fests" to Parallel Monologues. Early-2000s panel shows such as *Crossfire* thrived on raised voices, interruption, and personal digs. Experiments revealed that televised incivility elevates physiological arousal yet erodes trust in government, even when issue content is held constant (Mutz and Reeves 2005). By the mid-2010s, however, genuine left–right sparring had almost vanished; *Crossfire* was canceled again in 2014, symbolizing CNN's retreat from adversarial debate (Forgette and Morris 2006). A decade-long content audit shows that after the 2016 election Fox News booked steadily more conservative guests while CNN and MSNBC booked more liberals, leaving scant room for on-air disagreement (Kim, Lelkes, and McCrain 2022). When dissent does appear, it is often tokenistic or performative rather than deliberative (Sobieraj and Berry 2011).

Consequences for Democratic Deliberation. Incivility and ideological uniformity carry measurable civic costs. Uncivil exchanges not only entertain but also prime hostility toward political opponents and lower institutional trust (Mutz and Reeves 2005). Because partisan networks rarely present serious cross-cutting dialogue, viewers encounter few competing frames, exacerbating motivated reasoning and affective polarization. The effects cascade beyond the viewing public: elite rhetoric (Gertz 2020) and social-media conversations (Ding, Horning, and Rho 2023) frequently echo televised talking points, amplifying polarization across multiple arenas.

Computational Mapping of Ideological Slant. Recent work leverages natural-language processing (NLP) and large transcript corpora to quantify network bias with temporal precision. By linking every on-air guest to their campaign-donation ideology score, Kim et al. chart hourly shifts in slant from 2010–2020, showing Fox's pronounced rightward drift and CNN/MSNBC's reciprocal move left after 2016 (Kim, Lelkes, and McCrain 2022). Earlier econometric analyses reached similar conclusions using hand-coded tone and carriage data (Martin and Yurukoglu 2017). Such computational audits reveal that partisan skew is dynamic—responsive to electoral cycles, competitive positioning, and programming genre—and they supply the methodological blueprint for the present study.

Taken together, these strands demonstrate that *who* watches (selective exposure), *what* they watch (ideological uniformity), and *how* it is presented (incivility) jointly deepen partisan division. Yet we still lack fine-grained evidence on how moment-to-moment argumentative structure within cable segments shapes downstream discourse. By integrating NLP with interaction-level annotations, this paper addresses that gap and specifies the mechanisms linking televised debate to audience polarization.

Show	Channel	Type	Time span	Episodes
The Tucker Carlson Show	Fox News	Weekday	Jan 2017 – Jun 2023	1491
Hannity	Fox News	Weekday	Jul 2009 – Jan 2014 and Jan 2017 – Dec 2024	2441
Special Report with Bret Baier	Fox News	Weekday	Jul 2009 – Oct 2013 and Jan 2017 – Dec 2024	2672
Outnumbered	Fox News	Weekday	Oct 2015 – Dec 2024	1644
The Ingraham Angle	Fox News	Weekday	Oct 2017 – Dec 2024	1584
Gutfeld!	Fox News	Weekday	Jan 2017 – Dec 2024	1001
Jesse Watters Primetime	Fox News	Weekday	Jan 2022 – Dec 2024	634
The Story with Martha MacCallum	Fox News	Weekday	May 2017 – Dec 2024	1639
Life, Liberty & Levin	Fox News	Weekend	Feb 2018 – Dec 2024	481
Anderson Cooper 360°	CNN	Weekday	May 2019 – Dec 2024	1203
Erin Burnett OutFront	CNN	Weekday	Jun 2019 – Dec 2024	1206
Inside Politics	CNN	Weekday	Jul 2019 – Dec 2024	1210
The Lead with Jake Tapper	CNN	Weekday	Jun 2019 – Dec 2024	1056
The Source with Kaitlan Collins	CNN	Weekday	Jul 2023 – Dec 2024	334
Laura Coates Live	CNN	Weekday	Oct 2023 – Dec 2024	283
State of the Union	CNN	Weekend	Apr 2018 – Dec 2024	316
Fareed Zakaria GPS	CNN	Weekend	Apr 2017 – Dec 2024	353
The Beat with Ari Melber	MSNBC	Weekday	Jan 2023 – Dec 2024	288
Inside with Jen Psaki	MSNBC	2 Weekday	Jun 2023 – Dec 2024	130
Alex Wagner Tonight	MSNBC	Weekday	Jan 2024 – Dec 2024	179
Velshi	MSNBC	Weekend	Feb 2021 – Dec 2024	340
The Sunday Show with Jonathan Capehart	MSNBC	Weekend	Jan 2023 – Dec 2024	117
The Savage Nation	MSNBC	Weekday	Apr 2019 – Dec 2024	648
The Rachel Maddow Show	MSNBC	Weekly	May 2022 – Dec 2024	126

Table 1: Broadcast shows, overall time span, and total episode counts.

3 Dataset

Our corpus originates with the Internet Archive Television News Archive[1], an open repository that mirrors the closed-caption feeds of major U.S. broadcasters. We build on the cleaned release by Laohaprapanon and Sood (2017), but extend it in two important ways. First, debate research requires unambiguous speaker segmentation. Many shows in the Archive omit the $>>$ *speaker*:" token that marks a change of voice; others collapse hosts and guests under the same tag. To maximize precision, we retained only shows in which closed captions explicitly annotate both the host (e.g. $>>$ TUCKER:") and the guest ("$>>$" with no name). This criterion preserved most Fox News content but excluded the bulk of CNN and MSNBC's schedule, whose captioning conventions lack named turns.

Second, to restore balance across networks we supplemented the Archive with episode transcripts from the shows' official Apple Podcasts feeds, which supply RSS-accessible `.mp3` files which we crawled using a lightweight crawler, to obtain every available episode for shows whose captioning was inadequate in the Archive. We then transcribed and obtained speaker turns these episodes using a large Whisper model (Lyu, Lyu, and Chang 2024). Next, we then harmonized the formatting to match the Internet-Archive schema. Consistent with the Archive license, all podcast transcripts were already in the public domain; no paywalled or private audio was accessed.

Our final corpus contained 24 flagship debate or opinion show—nine on Fox News, eight on CNN, and seven on MSNBC—broadcast between January 2010 and December 2024. The sample contains over 17,000 episodes, accumulating just over 8,000 hours of video and roughly 194 million captioned words. Across those episodes we identify 2.13 million speaker turns, and on average, about 190 words per show. The earliest transcript dates to 7 July 2009 and the latest to 14 January 2025, giving a span of approximately 5,670 days (15.5 years, though we have gaps in 2014-2017). Table 1 lists every show alongside its coverage window and episode count. For each utterance we have a timestamp, the interlocutor's role (host or guest), the guest's inferred party affiliation (Section 4.2), and a stance label—*agreement*, *disagreement*, or *neutral*—generated by our classifier (Section 4.1). We also identify the topics discussed in each episode of the show (Section 4.3).

Dataset completeness. Despite its breadth, the corpus remains incomplete. We focus on the three dominant U.S. cable-news channels—Fox News, CNN, and MSNBC—because they account for the overwhelming share of prime-time debate programming. For each network we obtained every episode for which either the Internet Archive or an active Apple Podcasts RSS feed supplied transcripts. Even so, several shows exhibit intermittent coverage: entire weeks are missing when a programme was pre-empted, went on hiatus, or when its podcast feed was later purged from Apple's catalogue. Table 6 details the principal causes of these gaps. Fox, whose caption files consistently include speaker tags and whose podcast feeds remain intact, therefore contributes more episodes and longer time-series

[1]https://archive.org/details/tvarchive

than CNN or MSNBC. This asymmetry constrains some cross-show comparisons, especially for the latter networks, but all analysis we report below explicitly considers episode availability and our results are robust to the missing-data patterns.

Even though television data is widely available through the Internet Archive (Ding, Horning, and Rho 2023) and the Stanford TV data (Hong et al. 2021), data at the level of granularity (including guest names, timestamps, topics, etc) and quality we obtain in this paper is not available previously. We believe the resulting corpus, together with the scraping and alignment procedures documented here, provides a replicable template for future work that seeks speaker-resolved television data at scale. We will be releasing the full dataset and code for obtaining such data upon paper acceptance.

4 Methods

4.1 Detecting disagreement

Data Pre-processing For programmes whose closed captions already encode speaker names—predominantly Fox News—we treat every contiguous block attributed to the host as an anchor turn and concatenate the immediately following guest utterances until the host next speaks. Each such *host–guest bundle* constitutes one analytic pair. We further excise material outside the argumentative core of the show by trimming the transcript at the first host utterance and at the final host utterance, thereby discarding lead-in promos, commercial breaks, and closing credits.

Episodes obtained from Apple Podcasts are `.mp3` files without caption metadata. We convert audio to text and speaker diarisation using the large-V3 WHISPER model (Lyu, Lyu, and Chang 2024), producing a token-time-aligned transcript in which turns are labelled generically (SPEAKER 1, SPEAKER 2, ...). To identify the host we apply a two-step heuristic. First, the speaker with the greatest number of turns is presumed the host; every other speaker is provisionally a guest. If we observe multiple guests, occasionally pushing the host below first place, we override the rule and assign the first utterance in the episode to the host, propagating that label to all matching diarization IDs. Residual advertising copy at the head of the file is removed with show-specific regex filters keyed to recurring phrases such as "welcome to *Inside Politics*." We manually verified every episode where our heuristic did not apply to ensure the dataset was of high quality.

Data Annotation The main analytical backbone of the project is a stance label assigned to every *host–guest bundle*. Because disagreement can flare and subside many times within a single episode, we annotate each contiguous exchange rather than the episode as a whole. Each record therefore contains two text fields—`host_turn` and `guest_turn`—and one categorical target with three possible values: *Agreement*, *Disagreement*, or *Neutral*.

We began by hand-coding five hundred randomly selected pairs from various shows. Two annotators (the authors) worked independently, then reconciled differences in adjudication meetings, achieving an initial Cohen's κ of 0.81 and a post-reconciliation perfect agreement set. The following decision rules governed the exercise:

1. **Agreement** is assigned only when the guest endorses, amplifies, or otherwise affirms the host's preceding claim. *Example*:

   > **Host**: "Just to be clear, you agree the Russian invasion was an unprovoked violation of international law?"
   > **Guest**: "Absolutely, Moscow crossed a bright red line, and the world has to treat it that way."

2. **Disagreement** requires an explicit refutation, contradiction, or correction of the host. *Example*:

   > **Host**: "Premiums are up twenty–five percent—people hate this law because they're paying more."
   > **Guest**: "That's not accurate. Premiums were flat for the two previous years, and the average growth *slowed* after the law passed."

3. **Neutrality** covers all residual cases: clarifying questions, greetings, news reading, off-topic banter, or exchanges in which stance cannot be inferred.

Manual coding of the 2.13 million host–guest bundles was infeasible, so we adopted a staged pipeline that moves from small, human-verified seeds to large-scale annotation with language models.

We built the stance classifier through a four-stage, bootstrapped workflow. We began with a small hand-coded seed and found that task-specific encoders (e.g., fine-tuned DeBERTa-v3-base) topped out at barely 55% accuracy, so we pivoted to large language models, using GPT-4 in an interactive loop to expand the seed to 500 high-quality pairs. Two successive rounds of crowd annotation then supplied an additional 1,712 carefully triaged pairs, vetted with gold checks, consensus filtering, and manual audits. At each expansion we re-trained both commercial (GPT-4o) and open-source models (Llama-3 70B 4-bit, DeepSeek-R1-Distill-Llama-8B, Qwen-2.5 32B 4-bit), iteratively pruning noisy items until performance stabilised. The final DeepSeek model attains 89% accuracy on an unseen validation set, providing the labels used throughout the paper. Full details of the annotation rounds, model variants, and evaluation metrics appear in the Appendix.

This fine-tuned model was then applied to the rest of the 17,000 episodes, generating stance labels for 2.13 million host–guest bundles. For each show, we manually sampled a random sample of 100 pairs to check the performance of our model and the accuracy for all shows was over 85%, thus giving us the confidence that our model and predictions are high fidelity and can be relied upon. Overall, we identify 53.6% neutral, 30.4% agreement and 16.0% disagreement labels for all the 17,000 episodes.

4.2 Guest identification and party coding

Our procedure for determining who the guests were in each episode—and how those guests leaned politically—unfolded in two stages.

Stage 1: extracting guest names with GPT 3.5. For every episode we isolated the host's opening monologue, which typically lists the evening's interviewees. We then

passed that text to the `gpt-3.5-turbo` model through the OpenAI API in a zero-shot setting. A representative prompt for *Life, Liberty & Levin* was:

> This is a transcript of the show ``Life, Liberty and Levin,'' hosted by Mark Levin. He is never a guest. Extract all invited guests who actually join the show from the full opening monologue, and for each give their political affiliation as one of [Democrat,Republican,Other,Unknown]. If the guest is a known office-holder, use their real party. If they are a journalist, member of a non profit, an advocate or any other, infer their closest alignment from your training. If you absolutely cannot determine affiliation, use Unknown -- but only as a last resort. Your reply must be *only* a JSON array (no narrative, no Markdown, no code fences).

This step yields, for each episode, a list of guest names paired with the model's first-pass guess at affiliation. The host's name is systematically excluded.

Stage 2: refining affiliation labels with external data. Journalists, editors, and many subject-matter experts do not have an official party registration, which often led the language model to return `Unknown`. To improve coverage, we matched the extracted names against the DIME database (Bonica 2015), a large-scale compilation of itemized U.S. campaign-finance records. This is a high fildelity resource used in previous research and has been shown to be of high quality, containing a large database of public personas (Kim, Lelkes, and McCrain 2022).

Applying this two-step pipeline to the full corpus provided an affiliation label for nearly every recurring guest, reducing reliance on the `Unknown` category. Figure 1 shows the final distribution, showing, as expected, that Fox News features predominantly Republican guests, while MSNBC features predominantly Democratic ones.

4.3 Topic modeling

To recover the topics of each episode we applied BERTopic (Grootendorst 2022) to every episode transcript in the corpus.

Pre-processing. Captions were lower-cased, stripped of non-alphabetic characters, filtered through the NLTK English stop-word list, and purged of "intro-junk" tokens such as host names and sponsor slogans. The cleaned texts were embedded with the `all-MiniLM-L6-v2` sentence-transformer.

Model configuration and tuning. We initially set the vectorizer to `min_df = 2`, `max_df = 0.9`, and `max_features = 30,000`. Dimensionality reduction used UMAP with cosine distance (`n_components = 3`, `min_dist = 0.1`, `n_neighbors = 10`); soft clustering employed HDBSCAN, which yields a membership probability for every document. Hyper-parameters were tuned by grid search on a 5% development slice, optimizing for *codeless* coherence—our best setting achieved a coherence score of 0.689.

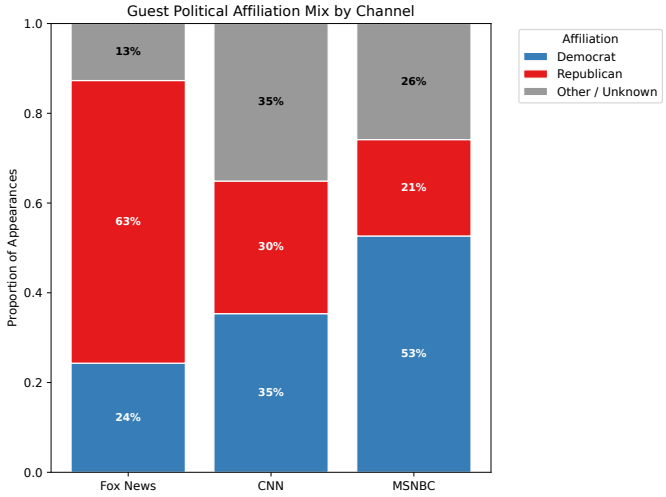


Figure 1: Fraction of guests by political leaning and channel. As expected, we see that Fox News invites mostly Republicans and MSNBC mostly Democrats.

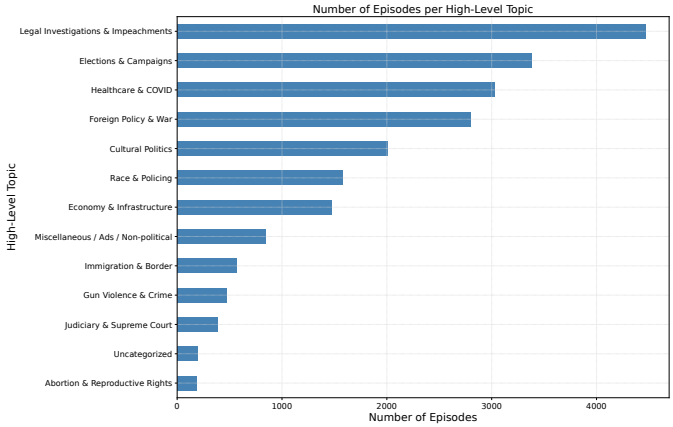


Figure 2: Episodes per high level topic.

Topic consolidation. The tuned model returned 146 raw clusters. Manual inspection revealed substantial lexical overlap, so we merged semantically redundant clusters—using keyword similarity and exemplar phrase matching—down to 80 distinct topics. Finally, two authors grouped those 80 topics into 15 higher-level categories (e.g. ELECTIONS & CAMPAIGNS, CULTURAL POLITICS, COVID-19, etc); any disagreements were resolved by discussion. These 15 top-level labels constitute the topic taxonomy visualized in Figure 2. The list of all 80 topics is added in table 6 in the Appendix.

5 Findings

We organize the findings along three dimensions. First, we present longitudinal disagreement and agreement from 2010

through 2024, documenting how the share of disagreement contracts across most shows and channels (Section 5.1). Second, we pair every utterance with the guest's party identification to show how partisan booking decisions shape conversational tone—Democratic invitees receive more agreement on MSNBC, Republicans on Fox, while CNN remains comparatively balanced (Section 5.2). Third, we disaggregate the data by topic, demonstrating that the lowest levels of disagreement cluster around culture-war flashpoints such as abortion, gun rights, and immigration, a pattern that underscores the echo-chamber character of contemporary debate television (Section 5.3).

5.1 Agreement and Disagreement over time

In this analysis, we compute the average (dis)agreement per show, aggregated by month. For each month, we calculate the mean level of (dis)agreement observed across all shows within that period. Alongside the mean, we also present the 95% confidence intervals, which allows us to track how (dis)agreement varies over time.

Figure 3 shows the temporal trajectories of host–guest disagreement for a (non random) sample of the shows in our dataset. Across the full 2017–2024 window, disagreement on average constitutes only a surprisingly small share of airtime—typically between 10% and 15%—and, for most shows, that share is either flat or trending downward. The pattern is clearest on Fox News. Tucker Carlson Tonight begins its run in 2017 with roughly one third of all host–guest exchanges coded as explicit disagreement, yet this fraction drops to about 15% by the time the show goes off-air in April 2023 (Figure 3a). The ordinary-least-squares slope on the monthly series is –3.9 percentage points per year ($p = 0.000$), confirming a systematic retreat from disagreement. Our finding dovetails with the New York Times content audit of 1,100 episodes, which likewise documents a contraction in dissenting perspectives over the same period (Times 2022). Laura Ingraham's primetime hour shows a similar trend (Figure 3d), albeit from a slightly lower baseline (-1.3 percentage points per year, $p=0.000$), showing that the phenomenon is not idiosyncratic to a single host but characterizes Fox's broader editorial posture.

CNN exhibits a more heterogeneous picture. Anderson Cooper 360° mirrors Tucker Carlson in direction (Figure 3b), though the glide path is gentler: disagreement falls from roughly 20% in 2019 to 14% by late 2024, a decline of 1.2 percentage points per year ($p = 0.000$). Similarly, the Sunday show State of the Union—which is structurally designed to pit partisan surrogates against one another—has a higher base rate of disagreement (near 30 %) in 2018 but goes down to around 20% by 2024 (decline of -2.1 percentage points per year ($p = 0.000$), Figure 3e).

MSNBC's inventory is the most internally consistent: disagreement begins low and stays low. The Savage Nation and Velshi, for instance, rarely strays above 10-15% disagreement in any month (Figure 3f), and the slope of the time-trend is statistically indistinguishable from zero. Other MSNBC offerings in the appendix (Figure 10) replicate this flat-line pattern. Taken together, the cross-network comparison reveals that while Fox is actively converging toward a narrow opinion corridor, CNN is bifurcated between a shrinking adversarial prime-time slot and a resilient Sunday shows, and MSNBC has long since embraced homophily as a default.

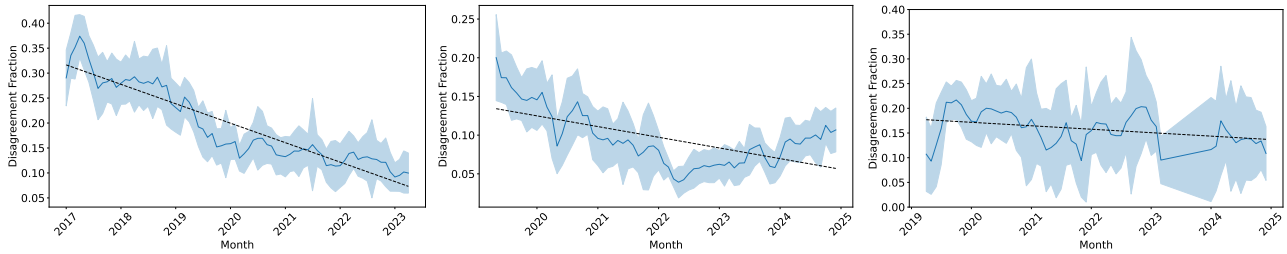
The complementary analysis of agreement (see Figure 4 for a small sample, Figure 11 in the Appendix for all shows) reinforces these conclusions. Where disagreement contracts, agreement expands. On Tucker Carlson Tonight the proportion of explicit agreement rises by almost 5 percentage points over the study period, a mirror image of the disagreement slide. Neither Anderson Cooper 360° nor The Savage Nation exhibits a statistically significant upward drift in agreement, suggesting that their modest losses in disagreement are being absorbed by neutral or ambiguous exchanges rather than by outright concurrence. Crucially, across all shows the combined share of agreement and neutral statements already exceeds 85% of airtime by 2024, leaving scant room for the deliberative contestation that debate television purports to deliver.

Two patterns emerge with genuine substantive weight. First, the data point to a deliberate editorial pivot away from inviting ideologically dissonant guests: as fewer counter-voices appear, on-air push-back correspondingly evaporates. This retreat is clearest on Fox and CNN, where disagreement falls significantly in 7 out of 9 Fox shows and 5 of 8 CNN shows. MSNBC shows no systematic trend—largely because the available time-series for its line-up is shorter—but its baseline levels of disagreement are already minimal. Second, the erosion of dissent is mirrored by a marked rise in affirmation, especially on Fox, where explicit agreement climbs in 5 of 9 shows.

Viewed alongside the topic-level findings in Section 5.3, the pattern is disquieting. The shrinking space for overt disagreement coincides with sustained—or even heightened—coverage of polarizing policy domains such as abortion, gun rights, and immigration. What remains, therefore, is "echo-chamber prime time": divisive themes are aired, yet almost invariably among like-minded interlocutors who amplify rather than interrogate partisan frames. The evidence is unmistakable: contemporary debate television has not mellowed into polite conversation; it has shed the adversarial backbone that once separated it from straightforward opinion broadcasting.

Longitudinal case studies To gauge how network-level patterns translate into editorial decisions on individual shows, we isolate three Fox News franchises for which we possess either uninterrupted long-run data or a well-documented format change: *Hannity*, *Special Report with Bret Baier*, and *Gutfeld!*. Each offers a distinct window onto the evolving relationship between host persona, booking strategy, and discursive diversity.

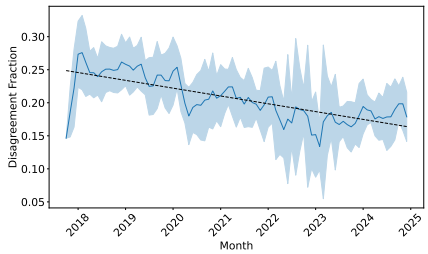
Sean Hannity Sean Hannity's primetime hour has been on the schedule since 2010, but its tone shifted markedly once Donald Trump emerged as the Republican standard-bearer. Hannity not only endorsed Trump on air; he also appeared on stage at a 2018 campaign rally in Cape Girardeau, Missouri, blurring the line between commentator and surrogate (Schwartz and Oprysko 2018). Our data mirror that



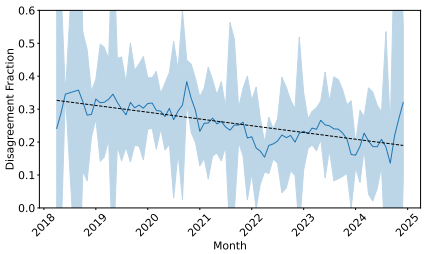
(a) Tucker Carlson (Fox)

(b) Anderson Cooper (CNN)

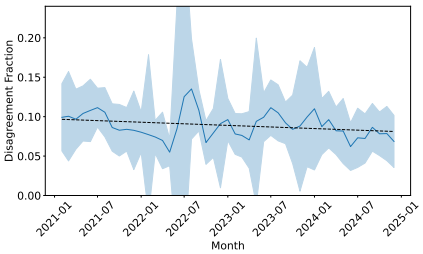
(c) The Savage Nation (MSNBC)



(d) Laura Ingraham (Fox)

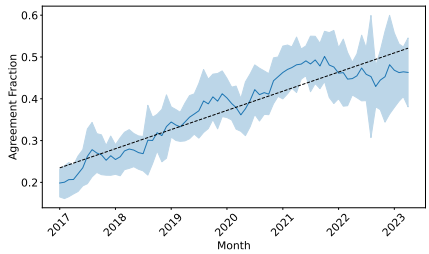


(e) State of the Union (CNN)

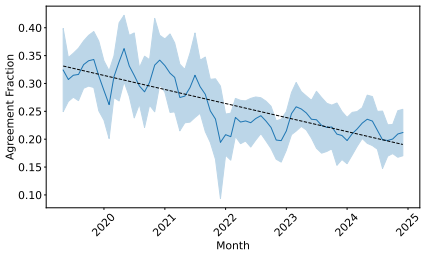


(f) Velshi (MSNBC)

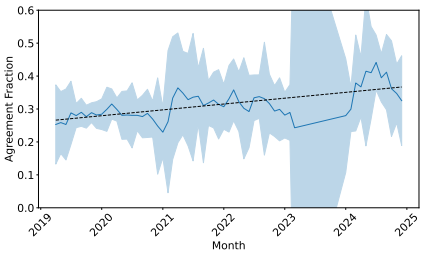
Figure 3: Trends in mean disagreement fraction over the years for a sample of the shows in our dataset.



(a) Tucker Carlson (Fox)



(b) Anderson Cooper (CNN)



(c) The Savage Nation (MSNBC)

Figure 4: Trends in mean agreement fraction over the years for a sample of the shows in our dataset.

realignment. The average disagreement fraction drops from 0.115 in the 2010–2013 window to 0.075 between 2017 and 2024—a 35% contraction that is statistically significant at $p < 0.001$ (Figure 5a). Over the same span, explicit agreement rises by nearly six percentage points (Appendix, Figure 11f), confirming that what little debate and discussion remained has largely been replaced by consonant, in-group affirmations. The show thus exemplifies how personal political engagement by a host can accelerate the migration from debate to echo chamber.

Bret Baier By contrast, Bret Baier—long marketed as Fox's "fair and balanced" chief political anchor (Darcy 2023) —shows no comparable drift. Across the full series (2010–2024) the monthly disagreement share hovers near 8%, with early-period (2010–2013) and late-period (2017–2024) means of 0.092 and 0.075 respectively, a difference that is not statistically distinguishable from zero (Figure 5b) . Agreement frequencies are likewise flat (Appendix, Figure 11o). In other words, even as Fox's opinion block narrows its viewpoint corridor, the network's flagship news show preserves a modest but steady level of dissent which is slightly comforting for the state of journalism and news reporting.

Greg Gutfeld The third case illustrates how structural changes in scheduling can rewire host–guest interaction. Fox announced in April 2021 that *The Greg Gutfeld Show*, formerly a weekend panel show, would move to weeknights at 11 p.m. under the new title *Gutfeld!* (imd 2021).

The shift coincided with a pronounced jump in disagreement: mean dissent climbs from 0.031 in 2016–2021 to 0.122 in 2021–2024 (Figure 5c) . Agreement, by contrast, contracts slightly (Appendix, Figure 11e). One plausible interpretation is that the late-night talk-show format rewards performative sparring, yielding more frequent but still low-stakes disagreement; yet even after the quadrupling, adversarial exchanges fill barely 12% of airtime, leaving the show well within the echo-chamber regime typified by its prime-time siblings.

Taken together, these longitudinal profiles reveal the heterogeneous pathways through which debate show devolve

(or resist devolving) into an ideological monologue. Hannity's alignment with partisan activism precipitates a steady erosion of dissent; Bret Baier's adherence to straight-news conventions stabilizes it; and *Gutfeld!*'s format reboot jolts the metric upward without restoring balance. Each trajectory, however, converges on the same broader outcome documented in Sections 5.2 and 5.3: a shrinking space for genuine contestation even as coverage gravitates toward the most polarizing policy arenas of American politics.

Channel-level trajectories: yearly aggregates, 2017–2024 Figure 6 collapses the show-specific results into annual means of disagreement fractions for each channel. Three facts stand out. First, the direction of movement in disagreement is clear: between 2017 and 2024 the average share of explicit host–guest disagreement falls across the board—by roughly one—third on Fox News, a quarter on CNN, and a fifth on MSNBC. The decline is gradual rather than abrupt, but its consistency across eight successive years suggests a structural shift in booking or production practices rather than ephemeral news-cycle shocks. Second, the ordering of the bars is counter-intuitive (especially for the Overall averages (right most set of bars in Figure 6)). Fox, long portrayed in the literature as an archetypal echo chamber (Jamieson and Cappella 2008), nonetheless retains the highest level of dissent throughout the series, finishing 2024 at just over 0.10—still above CNN's 0.08 and MSNBC's 0.07. One plausible explanation is our sample's composition: since we have more shows from Fox which also include panel, opinion and news shows, whereas for CNN and MSNBC we primarily have only opinion shows. Third, although the networks start from different baselines, the slopes are convergent; the gap between Fox and its competitors shrinks by almost two percentage points over the eight-year window, signaling a sector-wide contraction in adversarial dialogue.

The mirror image appears in Figure 7. Average agreement climbs steadily on all three channels, with Fox again leading the pack: its mean agreement fraction rises from just under 0.40 in 2017 to about 0.55 in 2024, a gain of fifteen percentage points. CNN and MSNBC follow the same upward path, though from lower starting points and at a slightly slower pace. By the end of the period, the combined share of agreement plus neutrality exceeds 90% of airtime for every network–a level at which meaningful counter-argument becomes the exception rather than the rule.

Taken together with the show-level trajectories documented above, the yearly aggregates show clearly a troubling pattern where networks are shedding disagreement more quickly than it is gaining it. The data thus corroborate a core claim of selective-exposure scholarship that modern cable news tends to narrow, rather than widen, the interpretive frame of its audiences (Stroud 2011) but they also sharpen it: the narrowing is not merely a function of audience choice across outlets; it is being actively reinforced within the outlets themselves.

5.2 Disagreement by speakers affiliation

The most straightforward pathway to ideological consonance is simply to curate the guest list. Figure 1 confirms that strategy in practice: Republicans dominate Fox, Democrats crowd MSNBC, while CNN remains roughly balanced, mirroring earlier content-analysis work on cable line-ups (Kim, Lelkes, and McCrain 2022). Yet composition alone cannot explain on-air dynamics; what matters is how those guests are treated once the cameras roll. Figure 8 disaggregates the data by the guest's party label and the network to which they have been invited, allowing us to test whether "friendly" visitors are rewarded with deference and "outsiders" confronted with push-back.

On Fox, Republican guests do enjoy marginally warmer receptions: the average share of explicit agreement for Republicans significantly higher than for Democrats, while disagreement is identical for both groups. Statistical significance tested using a Welch's t-test assuming unequal variance. p <0.01 (Figure 8a). The substantive effect, however, is well under one percentage point suggesting that Fox's echo-chamber reputation is driven more by *who* appears on screen than by large differences in tone once the invitation is extended. Put differently, the ideological filter operates at the booking stage; once a Democrat is allowed through that filter, the host's deference scarcely deviates from the baseline accorded to Republican regulars.

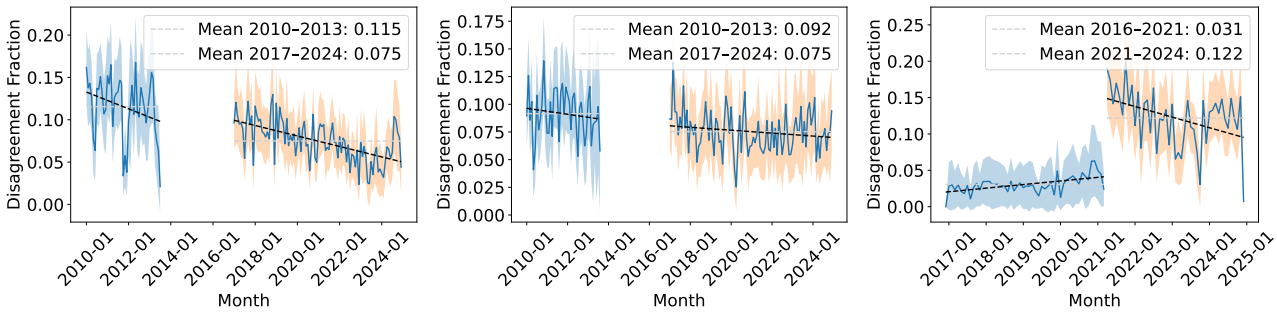
MSNBC displays a clearer in-group bias. Democratic guests receive a 0.40 agreement share, four points higher than Republicans (0.36), and experience slightly less disagreement (0.12 vs. 0.13). Both differences are statistically significant at p <0.01 (Figure 8b). Although the raw differences may look small, recall that disagreement is already exceedingly rare; a one-point reduction in dissent represents nearly 8% of all hostile exchanges on the network. The result tallies with scholarship arguing that liberal-leaning outlets enforce ideological solidarity more tightly than their conservative counterparts once the guest roster is set (Sobieraj and Berry 2011).

CNN, by contrast, shows virtually no partisan differential. The absence of a treatment effect is consistent with CNN's self-branding as a "down-the-middle" outlet and resonates with earlier findings that its conflict cues flow more from issue framing than from guest identity (Arceneaux and Johnson 2013).

5.3 Disagreement by topic

Table 2 ranks the five topics that elicit the least and most host–guest disagreement on each channel. Topics with least disagreement indicate issues where there was little counter voice or debate. Three patterns emerge:

First, the same topics surface in the "least disagreement" column across otherwise opposed outlets. Polarizing topics such as Abortion, Immigration, Gun violence and Policing appear among the five topics on most of channels. These are precisely the issues that animate contemporary culture-war politics, yet they rarely provoke open discussion on the news channels. Second, the topics that generate what most disagreement are largely procedural, event-driven or news related: the economy, elections & campaigns, COVID-19, etc. Third, there are interesting topics with little disagreement specific to channels such as discussions on Culture wars on both Fox News and MSNBC. Similarly, shows related to



(a) Hannity (Fox) (b) Bret Baier (Fox) (c) Gutfeld (Fox)

Figure 5: Disagreement over time for Hannity, Baier and Gutfeld.

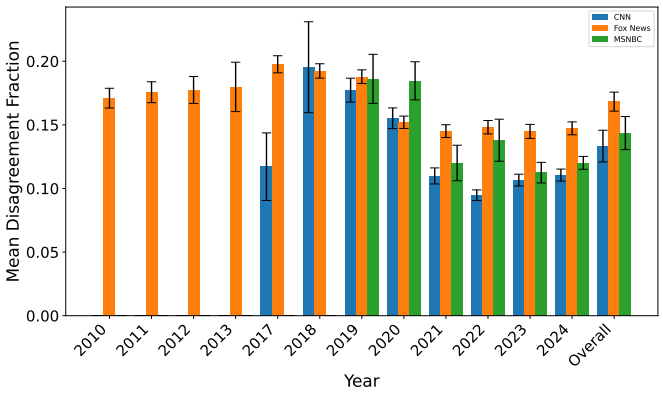


Figure 6: Yearly trends in disagreement for the three channels.

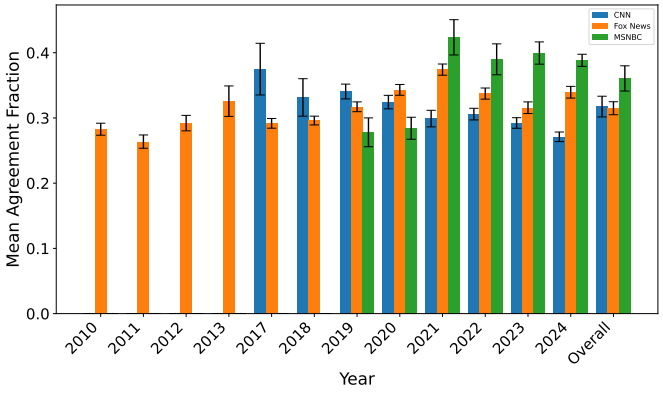


Figure 7: Yearly trends in agreement for the three channels.

the investigations and impeachment of Donald Trump were among the least disagreed on MSNBC.

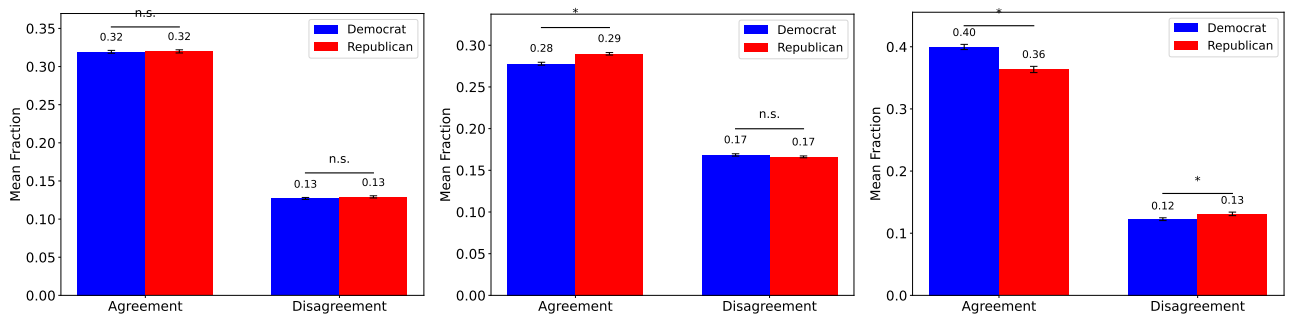
The implication from the findings in the previous two sections is that ideological filtering operates first at the guest-selection stage and then again at the topic-selection stage, leaving only narrow windows in which cross-cutting viewpoints make it to air. Taken together with the longitudinal evidence above, the topic analysis shows a central irony of partisan television. The shows devote ever more airtime to the controversies that polarize American politics (see Figure 9), yet they purge those very segments of meaningful disagreement.

Channel	Least disagreement	Most disagreement
Fox News	Abortion	COVID
	Cultural Politics	Economy
	Immigration	Elections & Campaigns
	Gun Violence	Judiciary
	Race & Policing	Trump Investigations
MSNBC	Abortion	Judiciary
	Trump Investigations	Elections & Campaigns
	Cultural Politics	Economy
	Immigration	Foreign Policy
	Race & Policing	COVID
CNN	Immigration	Economy
	Judiciary	Elections & Campaigns
	Gun Violence	COVID
	Foreign Policy	Cultural Politics
	Abortion	Trump Investigations

Table 2: Topics with the most and least disagreement per Channel. The topic names have been simplified to fit in the table.

6 Discussion

Summary of principal findings. Our speaker–turn analysis reveals a debate ecosystem that is strikingly placid by historical and cross-network comparison. Across the 2017–2024 window fewer than one-in-six host–guest exchanges contains explicit dissent; MSNBC is lowest at 13% and Fox highest at 17% (Figure 8). More troubling is direction of travel: disagreement declines year-on-year for



(a) CNN (b) Fox News (c) MSNBC

Figure 8: Average agreement/disagreement by political party of guests invited on a channel. * indicates statistical significance at $p < 0.01$. $n.s.$ indicates no significance.

almost every Fox and CNN show and is flat—but never high—on the MSNBC shows for which long-run data exist (Figure 10). The trajectory is not inevitable: archival captions from the early 2010s show *Hannity* hovering near 25% dissent, a level now unmatched by any prime-time hour. Guest composition compounds the effect. Networks already skew their invitation lists toward ideologically sympathetic voices, yet our turn-level labeling shows that even once a cross-partisan guest appears the dialogue tilts toward conformity: Republican invitees receive gentler treatment on Fox, Democrats on MSNBC, with CNN occupying a narrowing center. Finally, the very topics that polarize the electorate—abortion, gun rights, immigration—attract the least disagreement, converting ostensibly "hot" segments into performances of in-group consensus.

What do these patterns mean? Two implications stand out. First, the findings operationalize the "spectacle without contestation" hypothesis advanced by selective-exposure theory: cable outlets dramatize ideological battle lines through topic choice and affective framing yet withhold the clash of arguments that would expose audiences to competing rationales. That configuration can intensify *affective* polarization—warmth toward one's own side and hostility toward the other—while leaving factual beliefs untested, a dynamic known to entrench misperceptions and policy gridlock. Second, the disappearance of dissent is not merely a product of audience self-sorting across channels; it is reinforced *within* channels by editorial practices that book, frame, and moderate conversations so as to minimize overt conflict. The data therefore complicate optimistic accounts that place the onus solely on viewer choice: even citizens who seek out disagreement may be offered little once they tune in.

Why the decline? Several mechanisms are plausible. Audience analytics reward affective reinforcement more reliably than cognitive challenge; hosts cultivate personal brands tied to partisan authenticity; and the economics of booking favor a stable roster of predictable allies. The COVID-19 pandemic further normalized remote interviews, making it easier to screen out wild-card guests. Our data cannot pinpoint causal weightings, but the cross-network symmetry—in which left-leaning and right-leaning channels mirror each other—suggests structural, not idiosyncratic, forces.

Is disagreement the right metric? Raw frequency of dissent is a blunt instrument. Some disagreements are performative, others substantive; some are civil, others demagogic. Still, frequency matters for at least three reasons. First, democratic theory assigns deliberation a central role: citizens form better-justified preferences when exposed to counter-arguments. Second, low baseline dissent limits the potential for "constructive conflict," because intensity cannot compensate for absence. Third, our classifier is conservative: sarcasm, hedged rebuttals, and rapid-fire interruptions often register as *neutral*, implying that true argumentative engagement may be even rarer than we report. Future work should layer additional dimensions—tone, evidence quality, emotional valence—onto the stance signal to distinguish shallow sparring from reason-giving debate.

Limitations. The study has four chief constraints. *Coverage*: speaker-label quality forced reliance on podcast feeds for many MSNBC and CNN shows, shortening their temporal span relative to Fox. *Measurement error*: ASR inaccuracies, diarization drift, and stance-classifier mistakes (accuracy=0.89) introduce noise, though audits indicate no directional bias large enough to change headline trends. *Affiliation coding*: DIME donations are an imperfect proxy for ideological leaning, especially for non-U.S. guests or low-profile experts. *Missing modalities*: our analysis ignores visuals, tone of voice, and real-time fact-checking, all of which shape viewer perception of conflict.

Future directions. Integrating sentiment arcs and evidence quality could distinguish "heat" from "light" in televised argument. Linking caption-level disagreement to real-time audience metrics (e.g., minute-by-minute ratings, social-media engagement) would clarify whether viewers reward or punish dissent. Finally, extending the method to streaming talk shows and YouTube political channels would test whether the retreat from contestation is a cable-specific artifact or a broader feature of contemporary political media.

Overall, the erosion of on-air disagreement documented here challenges the longstanding notion of cable news as a public forum for competing ideas. If deliberation is an essential democratic input, then reviving formats that encourage genuine clash—not just spectacle—remains an urgent task for journalists, regulators, and scholars alike.

References

2021. Fox News Shifts 'Greg Gutfeld Show' to Weeknights. *IMDb*. Accessed: 2025-05-16.

Arceneaux, K.; and Johnson, M. 2013. *Changing minds or changing channels?: Partisan news in an age of choice*. University of Chicago Press.

Bonica, A. 2015. Database on ideology, money in politics, and elections (DIME). *Harvard Dataverse*, 2.

Darcy, O. 2023. Fox News anchor Bret Baier's reputation takes hit after text messages reveal what he said in wake of 2020 election. *CNN Business*. Accessed: 2025-05-16.

Ding, X.; Horning, M.; and Rho, E. H. 2023. Same words, different meanings: Semantic polarization in broadcast media language forecasts polarity in online public discourse. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 17, 161–172.

Farnsworth, S. J.; and Lichter, R. S. 2005. *The mediated presidency: Television news and presidential governance*. Rowman & Littlefield Publishers.

Forgette, R.; and Morris, J. S. 2006. High-conflict television news and public opinion. *Political Research Quarterly*, 59(3): 447–456.

Gertz, M. 2020. Study: Trump sent 657 live tweets of fox programming in 2019. *Media Matters for America*.

Goh, H. W.; Tkachenko, U.; Mueller, J.; and Cleanlab, C. C. 2022. Utilizing supervised models to infer consensus labels and their quality from data with multiple annotators. *CoRR*.

Grootendorst, M. 2022. BERTopic: Neural topic modeling with a class-based TF-IDF procedure. *arXiv preprint arXiv:2203.05794*.

Hartman, R.; Moss, A. J.; Jaffe, S. N.; Rosenzweig, C.; Litman, L.; and Robinson, J. 2023. Introducing Connect by CloudResearch: Advancing online participant recruitment in the digital age.

Hong, J.; Crichton, W.; Zhang, H.; Fu, D. Y.; Ritchie, J.; Barenholtz, J.; Hannel, B.; Yao, X.; Murray, M.; Moriba, G.; et al. 2021. Analysis of faces in a decade of us cable tv news. In *KDD'21: Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*.

Iyengar, S.; and Hahn, K. S. 2009. Red media, blue media: Evidence of ideological selectivity in media use. *Journal of communication*, 59(1): 19–39.

Jamieson, K. H.; and Cappella, J. N. 2008. *Echo chamber: Rush Limbaugh and the conservative media establishment*. Oxford University Press.

Kim, E.; Lelkes, Y.; and McCrain, J. 2022. Measuring dynamic media bias. *Proceedings of the National Academy of Sciences*, 119(32): e2202197119.

Laohaprapanon, S.; and Sood, G. 2017. Archive.org TV News Closed Caption Corpus. https://github.com/notnews/archive_news_cc/. Accessed May 5, 2025.

Levendusky, M. S. 2013. Why do partisan media polarize viewers? *American journal of political science*, 57(3): 611–623.

Littleton, C. 2024. Cable News Ratings 2024: Fox Gains, MSNBC and CNN In Post-Election Dropoff. *Deadline*. Accessed 2025-05-16.

Lyu, K.-M.; Lyu, R.-y.; and Chang, H.-T. 2024. Real-time multilingual speech recognition and speaker diarization system based on Whisper segmentation. *PeerJ Computer Science*, 10: e1973.

Martin, G. J.; and Yurukoglu, A. 2017. Bias in cable news: Persuasion and polarization. *American Economic Review*, 107(9): 2565–2599.

Mutz, D. C.; and Reeves, B. 2005. The new videomalaise: Effects of televised incivility on political trust. *American Political Science Review*, 99(1): 1–15.

Schwartz, J.; and Oprysko, C. 2018. Fox News raps Hannity for appearing onstage at Trump rally. *POLITICO*. Accessed: 2025-05-16.

Sobieraj, S.; and Berry, J. M. 2011. From incivility to outrage: Political discourse in blogs, talk radio, and cable news. *Political Communication*, 28(1): 19–41.

Stroud, N. J. 2010. Polarization and partisan selective exposure. *Journal of communication*, 60(3): 556–576.

Stroud, N. J. 2011. *Niche news: The politics of news choice*. Oxford University Press.

Times, N. Y. 2022. Inside the Apocalyptic Worldview of 'Tucker Carlson Tonight' (Published 2022) — nytimes.com. https://www.nytimes.com/interactive/2022/04/30/us/tucker-carlson-tonight.html. [Accessed 29-04-2025].

Wu, Y.; Zhou, Y.; Xu, B.; Wang, W.; and Song, Y. 2024. KnowComp at DialAM-2024: Fine-tuning Pre-trained Language Models for Dialogical Argument Mining with Inference Anchoring Theory. In Ajjour, Y.; Bar-Haim, R.; El Baff, R.; Liu, Z.; and Skitalinskaya, G., eds., *Proceedings of the 11th Workshop on Argument Mining (ArgMining 2024)*, 103–109. Bangkok, Thailand: Association for Computational Linguistics.

Ethics checklist

1. For most authors...
   (a) Would answering this research question advance science without violating social contracts, such as violating privacy norms, perpetuating unfair profiling, exacerbating the socio-economic divide, or implying disrespect to societies or cultures? Yes
   (b) Do your main claims in the abstract and introduction accurately reflect the paper's contributions and scope? Yes
   (c) Do you clarify how the proposed methodological approach is appropriate for the claims made? Yes
   (d) Do you clarify what are possible artifacts in the data used, given population-specific distributions? Yes.
   (e) Did you describe the limitations of your work? Yes.
   (f) Did you discuss any potential negative societal impacts of your work? Yes.
   (g) Did you discuss any potential misuse of your work? No.
   (h) Did you describe steps taken to prevent or mitigate potential negative outcomes of the research, such as data and model documentation, data anonymization, responsible release, access control, and the reproducibility of findings? No.
   (i) Have you read the ethics review guidelines and ensured that your paper conforms to them? Yes
2. Additionally, if your study involves hypotheses testing...
   (a) Did you clearly state the assumptions underlying all theoretical results? NA
   (b) Have you provided justifications for all theoretical results? NA
   (c) Did you discuss competing hypotheses or theories that might challenge or complement your theoretical results? NA
   (d) Have you considered alternative mechanisms or explanations that might account for the same outcomes observed in your study? NA
   (e) Did you address potential biases or limitations in your theoretical framework? NA
   (f) Have you related your theoretical results to the existing literature in social science? NA
   (g) Did you discuss the implications of your theoretical results for policy, practice, or further research in the social science domain? NA
3. Additionally, if you are including theoretical proofs...
   (a) Did you state the full set of assumptions of all theoretical results? NA
   (b) Did you include complete proofs of all theoretical results? NA
4. Additionally, if you ran machine learning experiments...
   (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? No
   (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? Yes
   (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? Yes
   (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? NA
   (e) Do you justify how the proposed evaluation is sufficient and appropriate to the claims made? Yes
   (f) Do you discuss what is "the cost" of misclassification and fault (in)tolerance? Yes
5. Additionally, if you are using existing assets (e.g., code, data, models) or curating/releasing new assets, **without compromising anonymity**...
   (a) If your work uses existing assets, did you cite the creators? Yes
   (b) Did you mention the license of the assets? NA
   (c) Did you include any new assets in the supplemental material or as a URL? NA
   (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? Yes
   (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? Yes
   (f) If you are curating or releasing new datasets, did you discuss how you intend to make your datasets FAIR (see **?**)? Yes
   (g) If you are curating or releasing new datasets, did you create a Datasheet for the Dataset (see **?**)? NA
6. Additionally, if you used crowdsourcing or conducted research with human subjects, **without compromising anonymity**...
   (a) Did you include the full text of instructions given to participants and screenshots? Yes
   (b) Did you describe any potential participant risks, with mentions of Institutional Review Board (IRB) approvals? Yes
   (c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? No
   (d) Did you discuss how data is stored, shared, and deidentified? Yes

Appendix

Detecting agreement/disagreement

Here we detail the entire pipeline we followed to obtain the agreement/disagreement/neutrality labels for the host/guest pairs.

Our stance–labeling workflow unfolded in four stages, moving from small, manually curated seeds to a production-scale large-language-model (LLM) annotator.

Stage 1: baseline with task-specific encoders. We first adopted the DialAM-2024 framework of Wu et al. (2024), fine-tuning DEBERTA-V3-BASE on 160 of 200 hand-labeled bundles from *Tucker Carlson Tonight*. Test accuracy reached only 55%, and augmenting the training set with the original DialAM corpus drove accuracy down to 48.5%. These results ruled out conventional encoder–classifier models for our domain.

Stage 2: iterative GPT-seed expansion. Turning to instruction-tuned LLMs, we grew the seed set to 500 bundles via an interactive loop with ChatGPT (GPT-4). The model labelled batches of 20 pairs in zero-shot mode; authors corrected its mistakes and fed the revised items back as demonstrations before issuing the next batch. A subsequent 80/20 train–test split yielded 73% accuracy for zero-shot GPT-4 and 79% after fine-tuning GPT-3.5-TURBO. These figures established a workable floor but left room for improvement.

Stage 3: crowd round 1 and consensus filtering. To enlarge the training pool we recruited 150 U.S. participants on CloudResearch (Hartman et al. 2023). We sampled 1,000 bundles drawn equally from *Tucker Carlson Tonight*, *Hannity*, and *The Rachel Maddow Show*. Each 20-pair block was triple-annotated on a purpose-built website featuring guideline examples and embedded gold checks. Quality screening (gold failures, implausibly short completion times) removed 12% of submissions. Inter-annotator reliability was modest (Fleiss κ = 0.302; Krippendorff's α = –0.005), so we applied CLEANLAB (Goh et al. 2022) to infer consensus labels, retaining 459 high-confidence pairs. Combining these with the 500 GPT-seed pairs (959 total) and re-training multiple models produced the accuracies reported in Table 3.

Stage 4: crowd round 2, manual triage, and final model. Because model performance improved with training-set size, we commissioned a second CloudResearch round for another 1,000 bundles. The raw labels, however, degraded accuracy (Table 4). Manual audit revealed liberal use of "agreement" for mere acknowledgments, violating our stricter definition. We therefore hand-verified every ambiguous triple, retained 712 balanced, high-quality pairs, and merged them with the earlier 959, yielding a 1,671-pair gold corpus. Fine-tuning the best open-source model (MISTRAL-7B-INSTRUCT) on this set restored accuracy to 84%; GPT-4o, fine-tuned on the same data, achieved 89% (Table 5).

6.1 Topic analysis

Figure 9 shows the high level topics per channel over time. We see clear trends like the increase in coverage in COVID/Race & policing in 2020, consistent coverage for foreign policy content, etc.

The full list of 80 topics are shown below.

1. 2012 GOP debate on Pakistan and national security post-bin Laden
2. 2012 GOP primaries and healthcare policy debates
3. 2020 Democratic Primary: Sanders vs. Bloomberg, Super Tuesday, and Media Scandals
4. 2020 Democratic primary candidates
5. 2020 Trump Foreign Policy
6. 2024 Republican Primary Campaigns in Iowa and New Hampshire
7. Abortion Rights and the Overturning of Roe v. Wade
8. Ads
9. Afghanistan Withdrawal and Evacuation Crisis at Kabul Airport
10. Andrew Cuomo's Resignation Amid Nursing Home and Sexual Harassment Scandals
11. Black Lives Matter, socialism, and cultural politics
12. Black Voter Representation, Kamala Harris Campaign, and Tim Walz Commentary
13. Brett Kavanaugh Confirmation Hearings and Sexual Assault Allegations
14. Brexit
15. British Royal Family Coverage: Queen Elizabeth's Legacy and King Charles's Ascension
16. COVID origins, Christianity, immigration, and media figures
17. Campus Encampments, Free Speech, and Debates on Antisemitism in Higher Education
18. Chinese Spy Balloon Incidents and U.S. Airspace Surveillance Controversies
19. Collective bargaining and emergency authority
20. Conservative Commentary on Marijuana, Free Speech, and Cultural Politics
21. Covid-19
22. Critiques of the Chinese Communist Party and US-China Strategic Rivalry
23. DOJ and FBI Leadership During Mueller Investigation and Trump-Era Controversies
24. Debates Over Tax Reform, Immigration, and Progressive Policy Priorities
25. Debt Ceiling Debates and Elon Musk's Influence in Political Discourse
26. Democrats
27. Diplomatic Relations and Summit Talks with North Korea
28. Early Trump administration power struggles and turnover
29. Egyptian Revolution and Mubarak's Fall Amid Muslim Brotherhood Rise
30. Electoral Strategy and Media Framing in Battleground States Like Wisconsin
31. Gabby Petito case coverage
32. George Floyd Protests and Police Response in Minneapolis and Beyond
33. George Floyd's Murder and the Prosecution of Derek Chauvin
34. Georgia Senate Elections and Runoff Campaigns
35. Gutfeld! Comedy
36. Hurricane Coverage: Storm Surge, Flooding, and Landfall Impacts
37. Hurricanes and FEMA Disaster Response in Hawaii, Texas, and the Southeast
38. IRS Fraud

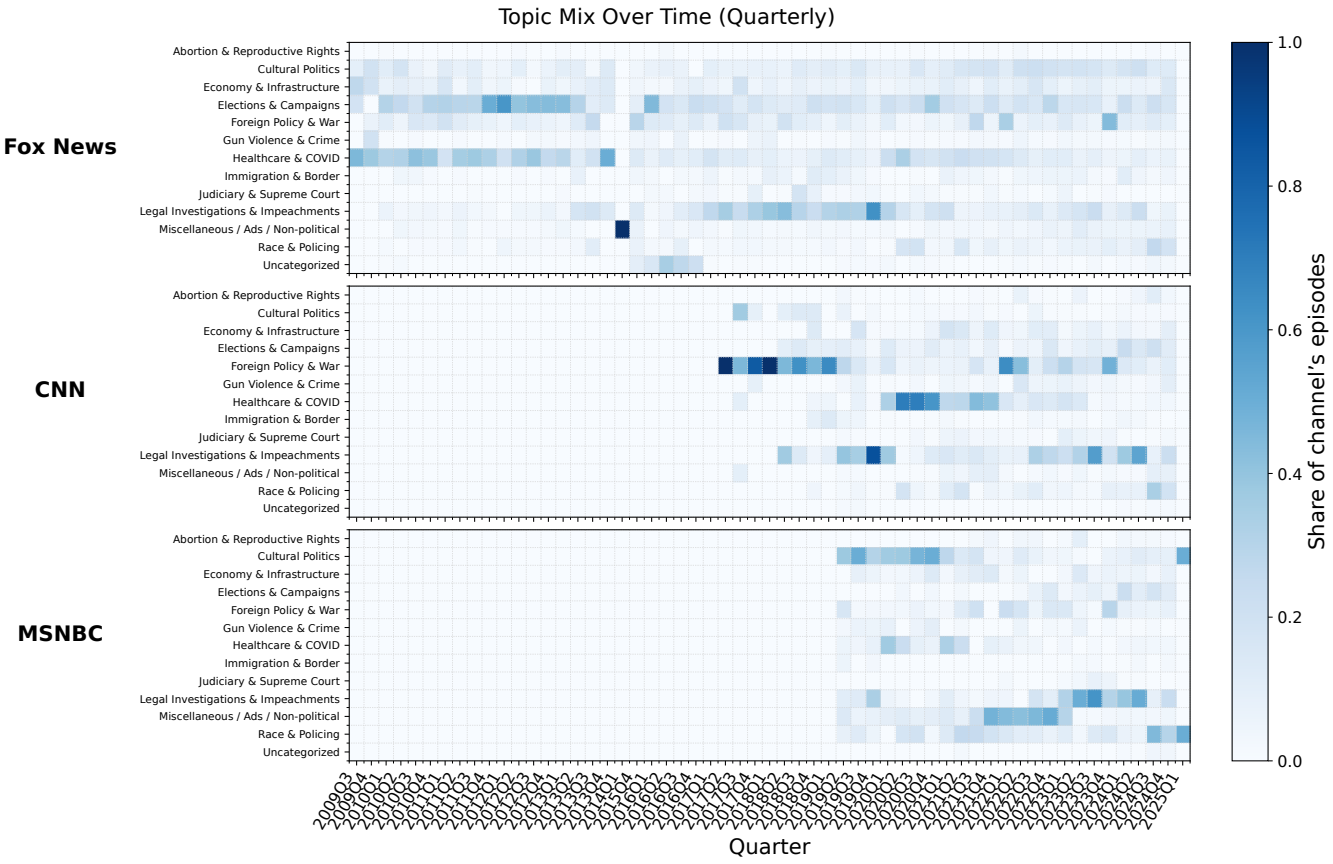


Figure 9: Topics per channel over time.

Model	Dataset	Train, Test	Accuracy
DeepSeek-R1-Distill-Llama-8B	High quality subsets from the crowdsourced dataset	(339, 120)	66.97%
DeepSeek-R1-Distill-Llama-8B	High quality subsets from the crowdsourced dataset + our own annotated 500 pairs	(498, 184)	70.64%
Llama-3.3 70b 4bit	High quality subsets from the crowdsourced dataset	(339, 120)	69.64%
Llama-3.3 70b 4bit	High quality subsets from the crowdsourced dataset + our own annotated 500 pairs	(498, 184)	74.98%
Qwen-2.5 32b 4bit	High quality subsets from the crowdsourced dataset	(339, 120)	64.08%
Qwen-2.5 32b 4bit	High quality subsets from the crowdsourced dataset + our own annotated 500 pairs	(498, 184)	65.22%
GPT-4o	High quality subsets from the crowdsourced dataset	(339, 120)	79.43%
GPT-4o	High quality subsets from the crowdsourced dataset + our own annotated 500 pairs	(498, 184)	**80.43%**

Table 3: Comparison of fine-tuned model accuracy using different combination of datasets (phase 1)

Model	Dataset	Train, Test	Accuracy
DeepSeek-R1-Distill-Llama-8B	High quality subsets from the crowdsourced datasets	(528, 245)	62.45%
DeepSeek-R1-Distill-Llama-8B	High quality subsets from the crowdsourced datasets + our own annotated 500 pairs	(861, 344)	68.90%
Llama-3.3 70b 4bit	High quality subsets from the crowdsourced datasets	(528, 245)	62.04%
Llama-3.3 70b 4bit	High quality subsets from the crowdsourced datasets + our own annotated 500 pairs	(861, 344)	69.00%
Qwen-2.5 32b 4bit	High quality subsets from the crowdsourced datasets	(528, 245)	64.08%
Qwen-2.5 32b 4bit	High quality subsets from the crowdsourced datasets + our own annotated 500 pairs	(861, 344)	62.50%
GPT-4o	High quality subsets from the crowdsourced datasets	(528, 245)	**72.24%**
GPT-4o	High quality subsets from the crowdsourced datasets + our own annotated 500 pairs	(861, 344)	66.87%

Table 4: Comparison of fine-tuned model accuracy using different combination of datasets (phase 2)

39. ISIS-Linked Terror Attacks in Europe: Paris, Manchester, and Brussels
40. Inflation, Energy, and Economic Slowdown
41. Israel-Gaza Conflict and Ceasefire Discussions
42. January 6th
43. Jeffrey Epstein Case and Legal Accountability of Labor Secretary Acosta
44. Kevin McCarthy
45. Kyle Rittenhouse Trial, Verdict, and Broader Jury Trials in Media Focus
46. Las Vegas Mass Shooting and Debates Over Gun Violence and Background Checks
47. Leaks, executive oversight, and antisemitism in political media discussions
48. Manchin's role in Democratic infrastructure and vaccine policy negotiations
49. Mar-a-Lago Classified Documents Case
50. Mass School Shootings and Mental Health Crisis in the US
51. Matt Gaetz Ethics Scrutiny and Cabinet Appointments During Presidential Transition
52. Mueller, Obamacare, and political controversies
53. National identity, foreign influence, and Second Amendment debates in conservative media
54. Nationalist cultural commentary and anti-communist rhetoric
55. Nikki Haley's 2024 Campaign, Southern Primaries, and IVF Policy Debates
56. Obamacare
57. Paul Pelosi Attack Incident and Related Legal Proceedings
58. Police Brutality and Public Outcry Following the Deaths of Floyd, Taylor, and Brooks
59. Postal service, shutdowns, and COVID-era political coverage
60. Religion, patriotism, and cultural conflict in U.S. media (Christianity, NFL, anthem protests)
61. Republican National Convention Coverage Amid Assassination Attempt and Political Tensions
62. Right-Wing Critiques of Race, Immigration, and Progressive Politicians
63. Right-Wing Framing of the Democratic Party and Radical Leftist Ideologies
64. Rudy Giuliani's Legal Troubles: Election Worker Defamation and Mueller-Era Ties
65. Russia-ukraine war and nato's strategic response
66. San Francisco Homelessness
67. Saudi Arabia's Crown Prince and the Killing of Jamal Khashoggi
68. Scandals and Public Figures (e.g., Avenatti, Bernie)
69. Syrian Civil War and U.S. Policy on Assad, Chemical Weapons, and ISIS

Model	Dataset	Train, Test	Accuracy
DeepSeek-R1-Distill-Llama-8B	High quality subsets from the crowdsourced datasets (manually checked)	(459, 253)	89.61%

Table 5: Final model used

70. Tea Party-Era Health Reform Debates
71. Titan Submersible Disaster and Ocean Exploration Catastrophe Investigation
72. Trump Administration Infighting, Mueller-Era Investigations, and DACA Policy Debates
73. Trump Hush Money Trial Featuring Cohen and Daniels
74. Trump Impeachment Hearings and Whistleblower Testimony
75. U.S. Drone Strike on Qasem Soleimani and Escalating Tensions with Iran
76. US-Mexico border security and cartel-driven immigration
77. Urban Crime and Policing in Central Park and Manhattan with NYPD Coverage
78. VP Kamala Harris and Democratic Campaign Strategies on Abortion
79. Virginia and New Jersey Gubernatorial Races with National Culture War Overtones

Gaps in the Data

Due to a hiatus or simply not being able to scrape data from the Internet Archive platform or even Podcast, there are some gaps in the data. They are noted in table 6.

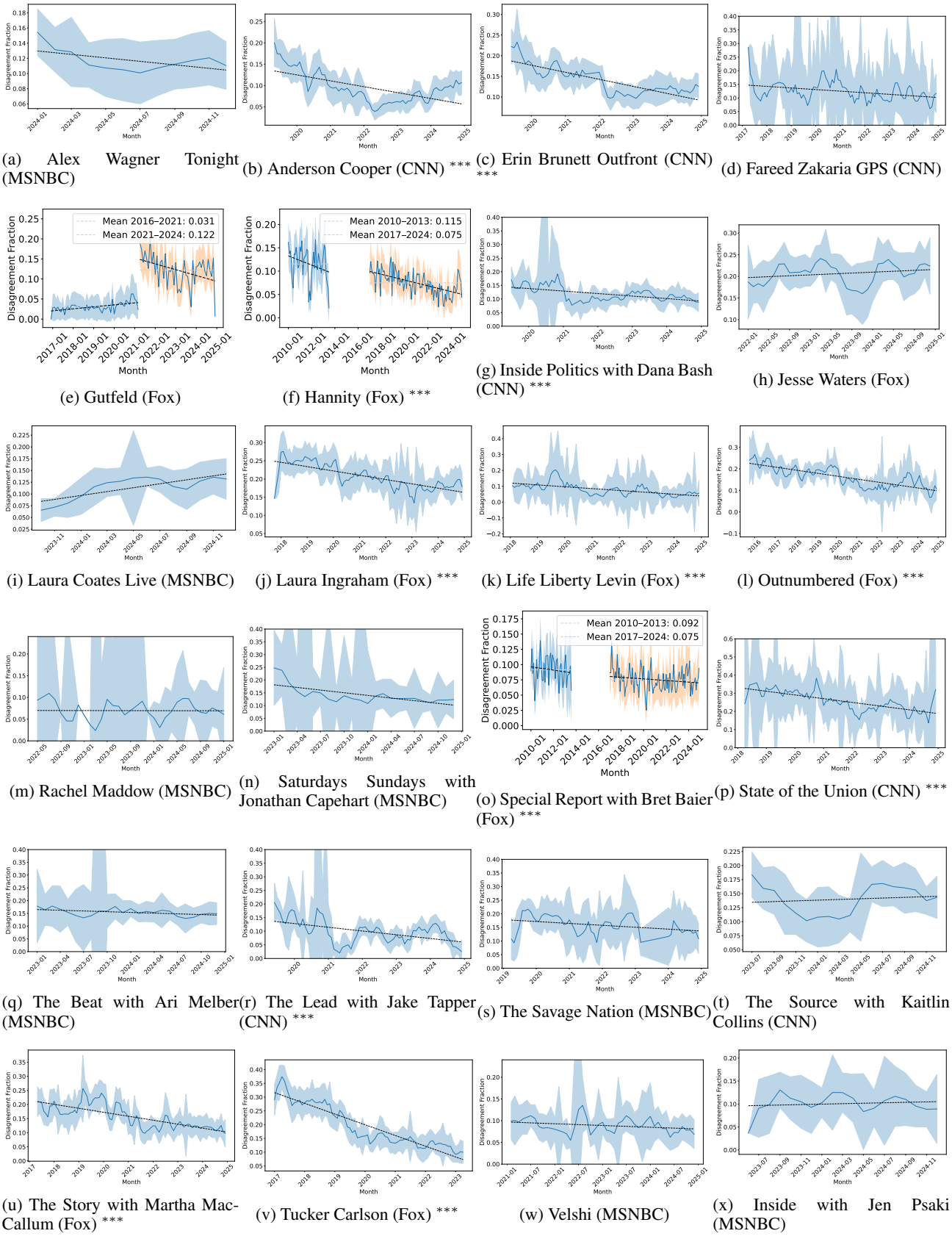
Show	Coverage Gaps or Anomalies
Hannity	No data from 2014 to 2016.
The Rachel Maddow Show	Only one episode in September 2022, all labeled as agreement or neutrality because it was just Rachel and a reporter talking through the day's news. Weekly airing format may reduce disagreement frequency.
Outnumbered	Very few episodes before September 2015; analysis starts from October 2015.
Special Report with Bret Baier	No episode available from November 2013 to December 2016.
Life, Liberty & Levin	Some months contain only 2–3 episodes with limited speaker turns; often results in 0 recorded disagreements.
The Savage Nation	Did not air from March 2023 through December 2023.
Erin Burnett OutFront	Did not air from September 2021 to December 2021 and again from October 2023 to December 2023.

Table 6: Data coverage gaps and broadcast anomalies for selected shows.

Agreement/Disagreement plots for all shows

Figure 10 shows the disagreement plots for all shows.

Figure 11 shows the agreement plots for all shows.



(a) Alex Wagner Tonight (MSNBC)

(b) Anderson Cooper (CNN) ***

(c) Erin Brunett Outfront (CNN) ***

(d) Fareed Zakaria GPS (CNN)

(e) Gutfeld (Fox)

(f) Hannity (Fox) ***

(g) Inside Politics with Dana Bash (CNN) ***

(h) Jesse Waters (Fox)

(i) Laura Coates Live (MSNBC)

(j) Laura Ingraham (Fox) ***

(k) Life Liberty Levin (Fox) ***

(l) Outnumbered (Fox) ***

(m) Rachel Maddow (MSNBC)

(n) Saturdays Sundays with Jonathan Capehart (MSNBC)

(o) Special Report with Bret Baier (Fox) ***

(p) State of the Union (CNN) ***

(q) The Beat with Ari Melber (MSNBC)

(r) The Lead with Jake Tapper (CNN) ***

(s) The Savage Nation (MSNBC)

(t) The Source with Kaitlin Collins (CNN)

(u) The Story with Martha MacCallum (Fox) ***

(v) Tucker Carlson (Fox) ***

(w) Velshi (MSNBC)

(x) Inside with Jen Psaki (MSNBC)

Figure 10: Disagreement patterns over time for all shows. *** indicates a significant decline ($p < 0.001$) in slope.

(a) Alex Wagner Tonight (MSNBC)

(b) Anderson Cooper (CNN)

(c) Erin Brunett Outfront (CNN)

(d) Fareed Zakaria GPS (CNN)

(e) Gutfeld (Fox)

(f) Hannity (Fox) ***

(g) Inside Politics with Dana Bash (CNN)

(h) Jesse Waters (Fox)

(i) Laura Coates Live (MSNBC)

(j) Laura Ingraham (Fox) ***

(k) Life Liberty Levin (Fox) ***

(l) Outnumbered (Fox) ***

(m) Rachel Maddow (MSNBC)

(n) Saturdays Sundays with Jonathan Capehart (MSNBC)

(o) Special Report with Bret Baier (Fox)

(p) State of the Union (CNN)

(q) The Beat with Ari Melber (MSNBC)

(r) The Lead with Jake Tapper (CNN)

(s) The Savage Nation (MSNBC) ***

(t) The Source with Kaitlin Collins (CNN)

(u) The Story with Martha MacCallum (Fox)

(v) Tucker Carlson (Fox) ***

(w) Velshi (MSNBC)

(x) Inside with Jen Psaki (MSNBC) ***

Figure 11: Agreement patterns over time for all shows. *** indicates a significant increase ($p < 0.001$) in slope.