

The Effect of Streaming Chat on Perceptions of Debates

Victoria Asbury*, Keng-Chi Chang[†], Katherine McCabe[‡],
Kevin Munger[§], Tiago Ventura[¶]

May 1, 2020

Abstract

Broadcast political media is becoming more social. Technology has advanced to the point where now many online video “livestreams” come with embedded live-streaming chatboxes, uniting the on-screen and social components into one real-time, integrated experience. We investigate how these chatboxes may shape perceptions of political events. We conducted a field experiment during the September 2019 Democratic Primary Debate where subjects were assigned to view the debate with or without streaming chatboxes. Subjects were encouraged to view the debate on the ABC homepage (with no chatbox), on FiveThirtyEight.com (expert chat) or on Facebook (social chat). We use text analysis to characterize the types of comments in the two chat streams. Our experimental findings indicate that Democratic subjects assigned to the Facebook chat condition reported lower affect towards Democrats and a worse viewing experience. The tone of candidate-directed comments also matter: We find that the number of negative comments about a candidate on the social chat predicts a decreased feeling thermometer rating of that candidate, while the number of positive comments predicts increased belief that that candidate will improve in the polls.

*Harvard University

[†]University of California San Diego

[‡]Rutgers University

[§]Penn State University

[¶]University of Maryland

Media consumption is becoming a more social and interactive endeavor. Media producers encourage audiences to engage with their content via social media, and media consumers take to social media as a “second screen” to see what others think as they watch various types of content, from pre-recorded season finales to live events. More recently, the advent of live video with *integrated* streaming chat is gaining in popularity among younger generations. “Streaming chat” offers a viewing experience where the live video and real-time commentary are embedded on a screen together, encouraging viewers to immerse themselves in both sources at the same time.

Young politicians like Congresswoman Alexandria Ocasio-Cortez have famously adopted livestreams with streaming chat as a powerful political tool, and increasingly, major political events are also being broadcast through platforms that offer streaming chat options. This technological change has the potential to modify the effect of live broadcasts on the viewing public by changing a number of parameters of the experience that had been constant for decades.

Specifically, previous scholars have shown that exposure to commentary about political events (specifically, political debates) can alter perceptions of these events. The rise of “dual” or “second screening”—where viewers follow along with social commentary *during* the live broadcast—offers an enhanced source of influence in real time (Gil de Zúñiga, Garcia-Perdomo and McGregor, 2015; Vaccari, Chadwick and O’Loughlin, 2015). However, despite the more recent and rapid rise in integrated streaming chat during live political events, we have little evidence to show the potential effects of this technological change on political attitudes and behaviors.

To examine these potential effects, during the September 2019 Democratic Primary Debate, we pre-registered¹ and conducted a digital field experiment, modeled after the Gross, Porter and Wood (2019) panel experimental design, to study the influence of these real-time comment sections on the public’s perceptions of the debate, the participating candidates, and overall trust in the democratic process. In a two-wave survey, using Amazon’s Mechanical Turk to recruit subjects, we randomly assigned and encouraged 1095 participants to watch the debate on one of three online platforms: the ABC News website, which provided a livestream of the debate without a streaming chat; the ABC News Facebook page, which provided social commentary from public Facebook users alongside the video; and the FiveThirtyEight website, which provided live expert commentary from political analysts alongside the video. We simultaneously

¹Pre-registration available at [REDACTED]

collected the comment feeds from the social and expert conditions to characterize how a key aspect of the treatment— the content of the chatboxes— differed between conditions. Directly following the debate, participants completed a follow-up survey.

Because this is a field experiment, our design is high in ecological validity. Subjects consume a prominent political media event in real time, from the comfort of their own homes; we merely prompt them to vary the platform they use to view the debate. Our sample consists entirely of people who were already planning to watch the debate. This is not, of course, representative of the general population, but we believe that this non-representativeness is epistemically desirable. Forcing media upon people who would never choose to consume it produces potentially misleading counterfactuals (Arceneaux and Johnson, 2013).

Even though we lack explicit control over the content of the two treatments, we are able to trace the important dimensions of the expert and social chat contexts using text analysis to infer the nature of one of the most critical aspects of the treatments. Descriptively characterizing the “social” chat feed on the official Facebook livestream of the debate is not trivial, and understanding the contours of this social chat can help illuminate the potential effects of exposing subjects to different platforms. To this end, we hand-coded 6,500 “social” comments for both topic and sentiment, to supplement machine-coding methods.

Our results reveal stark differences in the real-time commentary on the “social” Facebook and “expert” FiveThirtyEight pages. The “social” commentary was significantly more negative in tone than the “expert” commentary and more likely to include comments that score high on scales designed to measure online toxicity. This staid description undersells the distasteful and sometimes extremely offensive nature of the Facebook chat. The volume and intensity of the comments to which viewers of the debate with the streaming chat modality were exposed has no analogue.

The results of our randomized experiment tightly match these descriptive results. Democratic respondents who were encouraged to watch the debate on the ABC News Facebook page came away from the debate with more negative feelings toward Democrats, aligning with the negative depictions of the debate participants in the real-time comments.

The text analysis also allows us to differentiate comments about different candidates, enabling us to capture changes in a number of candidate-level opinions. The number of negative comments about a given candidate is strongly predictive of decreased feeling thermometer evaluations of that candidate in the “social” condition relative to the

control. Notably, Vice President Joe Biden was the target of the highest number of negative comments (the modal such comment was about his advanced age) and saw the largest decline on the feeling thermometer scale—about 10 points, on the 100-point scale. Senator Kamala Harris received the second most negative comments (mostly vile misogyny) and also saw a decline on the feeling thermometer of about 10 points.

On the other hand, the number of *positive* comments about a candidate was highly correlated with an increased perception that that candidate would do better in the polls after the debate. The number one candidate on both positive comments and predicted poll bump was Senator Bernie Sanders; number two on both dimensions was Andrew Yang.

It is worth noting that some of these findings went against some of our initial expectations. The development of new technologies for people to interact with others while consuming media is undoubtedly exciting in its potential to stimulate viewers, and we anticipated the potential for positive effects of the social chat on engagement. However, our findings point out potential pitfalls of real-time comments. Just as the broadcast era of political news proved to have priming and agenda setting effects on political attitudes (Iyengar and Kinder 1987), the streaming chat era may have its own impact in disrupting learning and information processing as citizens are exposed to the carnivalesque thoughts of their peers online.

Bennett and Iyengar (2008) famously describe a “new era of minimal effects,” primarily due to self-selection among media consumers. Streaming chat feeds are a novel technology in the density, diversity and intensity of the media stimuli they deliver, and they present a case where non-minimal effects can be expected.

Introduction of Streaming Chat

The technological capacity for audiences to engage in mass communication on one screen while consuming broadcast media on another (most often a “media event” that coordinates the interest of a large group of people (Dayan and Katz, 1992)) has existed for over a decade, and while earlier studies documented the existence and prevalence of the phenomenon Jungherr (2014); Larsson and Moe (2012), theoretical interest in dual or double screening began in earnest with Gil de Zúñiga, Garcia-Perdomo and McGregor (2015) and Vaccari, Chadwick and O’Loughlin (2015).

“Dual screening” or “second screening,” where viewers of political television media

participate in real-time conversations in online platforms like Twitter, has been shown to change the effect of media consumption on attitudes (Barnidge, Gil de Zúñiga and Diehl, 2017; Gil de Zúñiga, Garcia-Perdomo and McGregor, 2015; McGregor and Mourão, 2017; Vaccari, Chadwick and O’Loughlin, 2015). Traditionally in political science, the news media are thought to have the power to set the agenda and change the criteria by which we, as consumers, evaluate political figures and form attitudes on political issues (Iyengar and Kinder, 1987). The addition of a second screen could, at minimum, muddle the effects of media on public opinion. However, the extent to which this occurs depends on the amount of attention consumers give to the “second screen” and whether they treat online commentary similar to actual news content in deciding whether to “accept” the information in the commentary as considerations that can influence their beliefs (Zaller, 1992).

Our study focuses on an innovation to dual-screening. In recent years, streaming chatboxes alongside a related video feed, all on one screen, have become increasingly popular. Already the norm in livestreams of sports and video game competitions, chatboxes have recently been added to online streams of political media. In many ways “integrated real-time streaming chat” (hereafter referred to as just “streaming chat”) amplifies the potential effects of the second screen. Viewers are now immersed in both the video livestream and the real-time comments during the viewing experience.

The question, therefore, remains: Do real-time comments have the power to shape what is salient and accessible as consumers–potential voters–form their attitudes on political figures and events? And does it have the power, in turn, to change people’s attitudes?

We propose three theoretical pathways by which the addition of streaming chat windows to broadcast media might affect perceptions of the media event: frequency, content, and context of the comments that stream. Our discussion is premised on a technosocial context similar to the Facebook chat in our study, but we hope that identifying theoretically distinct pathways will extend the temporal validity of our theory and results in what is likely to be a rapidly evolving communication technology (Munger, 2019).

- **Frequency:** A high volume of streaming comments increases the potential for distraction and information overload.
- **Content:** Discussion of issues serves as a prime, increasing their salience. The

nature of the messages as a discussion network increases the likelihood of persuasion.

- **Context:** The composition of the commenters is not obvious to the viewer, which will lead them to over-estimate the quality of the information source for making inferences about public opinion.

We will now trace the theoretical antecedents of each of these pathways and hypothesize their potential consequences within the context of the setting of our study: the September 2019 Democratic Debate. As we explain in more detail below, our two treatment conditions involve assigning subjects to view two different types of streaming chat—on Facebook (social condition) and on FiveThirtyEight (expert condition)—which vary each of these parameters, producing different predictions. “Expert chat” is a much less prominent phenomenon than streaming social chat, and less novel, so while we consider social chat our primary treatment condition of interest, the expert chat prediction serves as a useful check on the mechanisms driving the effects we observe for streaming chat.

The Setting: The September Democratic Debate

Debates are an exemplar of the kind of political “media event” for which streaming chat is relevant, but they are also political media events in their own right. Primary presidential debates are crucial vehicles for disseminating information and setting campaign agendas. Indeed, studies have found primary debates can influence perceptions of candidate viability and electability (Yawn et. al 1998), affect towards candidates and issue salience (Best and Hubbard 1999), knowledge of candidates’ policy positions (Benoit et. al 2002), and voter preference (Benoit et. al 2010). Thus, primary debates, particularly those early in the primary season, are crucial political events, as they are mechanisms for introducing candidates to the voters, many of whom are not decided on a given candidate.

In the past, newspaper journalists were the chief political actors that signaled to readers what information in a presidential debate mattered *after the debate ended* (Benoit 2017). As cable news became an increasingly important domain for political analyses, cable news hosts and commentators also became important agents who communicate to audience members how debate performances should be interpreted

(Fridkin et. al 2008). The rise of dual-screening, however, has the potential to break down journalistic gate-keeping and influence perceptions in real-time.

The September 12, 2019 Democratic primary debate hosted by ABC News and Univision was a major media event. The September debate was the third debate of the 2020 presidential election season, however, it was the first time all of the top-polling Democratic candidates would be on one stage and able to confront each other directly. An ABC News press release after the night of the debate noted: “ABC News Democratic Debate coverage drew 2.9 million unique visitors and 11 million video views (ABC News Live and VOD) across ABC News digital properties and distributed partners including Facebook Live, Twitter, Apple News, YouTube, Hulu and Hotstar.” Given the amount of interest in the third primary debate and the high level of live-streaming and digital engagement the event attracted, the presidential debate featured in this study serves as an appropriate context for exploring the effect of streaming chat on political attitudes.

Frequency: Distraction and Affective Processing

Some early research on second screening focused on the capacity of the second screen to distract the audience’s attention (Gottfried et al., 2017; Van Cauwenberge, Schaap and Van Roy, 2014; Van Cauwenberge, dHaenens and Beentjes, 2015). However, research that centers the purposiveness of the act of second screening emphasizes the agency of the viewers in selecting when and where to consult the second screen (McGregor and Mourão, 2017), and whether to actually participate by posting commentary (Vaccari, Chadwick and O’Loughlin, 2015). For example, in order to follow a hashtag on Twitter, someone has to actively open Twitter and search for relevant comments – an act that is inherently purposive. Further, Gil de Zúñiga, Garcia-Perdomo and McGregor (2015) determines the type of person who is actually likely to engage in second screening, thus establishing the extent of these effects. This is particularly important because dual-screening requires viewers to actively engage with a second source of information. If they were merely watching the broadcast, incidental exposure to the second screen would be impossible.²

Consuming media with streaming chat is less likely to be purposive as more media sources embed chats with their video streams, forcing scholars to reconsider findings

²Another possibility is that users of a given chat platform (here, Twitter) become incidentally exposed to the discussion of the media event and either gain some information about it or even decide to tune in. Integrated streaming chat renders this process irrelevant.

about second screening which center the purposiveness of the act. This shift also increases the range of people exposed to streaming chat; as [Gil de Zúñiga, Ardèvol-Abreu and Casero-Ripollés \(2019\)](#) demonstrate in the case of another novel political communication technology (WhatsApp chat among extended family groups).

Even more significantly, the unity of the broadcast and the stream in the visual field makes streaming chat theoretically distinct from second screening. The richness of the visual information provided on a single screen showing a broadcast combined with streaming chat is unparalleled. Streaming chat is often located at the same eye level as the video broadcast, so the viewer consumes both automatically. This undercuts the purposiveness that even moving the eye between two screens provides, increasing the relevance of concerns about distraction. Just by tuning into a livestream with streaming chat, even if a viewer intends to focus on the video, it is likely they may be incidentally exposed to the real-time comments. In contrast, dual-screener watching a broadcast and also tracking a given hashtag on Twitter that refers to the media event being broadcast must click into the Twitter interface, scroll down, and then click at the top to load the most recent tweets. The experience of second screening is thus intrinsically more purposive than is observing the streaming chat flow across the screen.

We hypothesize that this element of the streaming chat platform design will change how viewers experience the media event relative to contexts where viewers can focus exclusively on the video:

- **Hypothesis 1:** Increased distraction from the streaming chat will cause viewers to consider the debate less enjoyable and informative, but the increased number of stimuli will cause them to be more engaged.

The expert chat has a dramatically lower frequency of comments, decreasing the problem of distraction. The addition of a small number of considered and informed comments, we predicted, would only enhance the process.

- **Hypothesis 1(e):** Increased commentary from the expert chat will cause viewers to consider the debate more enjoyable, informative and engaging.

Recent research on the experience of consuming political media in the contemporary context of total saturation and extreme affect can cause people to feel overwhelmed, anxious and angry ([Wagner and Boczkowski, 2019](#)). There is a wealth of evidence for the

important role these emotions play in political communication and participation (see [Wagner and Morisi \(2019\)](#) for a recent summary), so we also test whether streaming chat causes viewers to experience these negative emotions.

- **Hypothesis 2:** Streaming chat will cause viewers to be more angry and anxious.

The expert chat does not threaten to overwhelm the viewer, instead promoting a measured and reasoned response to the broadcast.

- **Hypothesis 2(e):** The expert chat will cause viewers to be less angry and anxious.

Content: Priming

Two of the processes most central to contemporary theories of political communication are *framing* and *priming*. [Cacciatore, Scheufele and Iyengar \(2016\)](#) argue that the two are often confused, and we follow the conceptual distinction they draw to evaluate priming effects in our study. Priming changes the relative salience of different characteristics or issues in determining how viewers evaluate politicians.³ Theoretically grounded in the cognitive process of “spreading activation,” priming affects which thoughts are immediately subsequent to considering a given politician ([Iyengar, 1987](#)).

The concept was developed for the era of broadcast media, and within the debate setting, originally concerned commentary immediately before or after a debate broadcast, given the temporally linear nature of the medium. Second screening enables near-instantaneous priming, as the viewer can consume commentary on a given debate response while it is still ongoing. Past work has shown that second screening increases the likelihood of persuasion among consumers of broadcast media ([Barnidge, Gil de Zúñiga and Diehl, 2017](#)), likely due to the immediacy of audience feedback.⁴

Furthermore, the primes produced in unmoderated streaming chat potentially include primes that would not be produced by traditional media sources. This includes primes that are openly racist, sexist or otherwise discriminatory, and primes that are simply untrue or, bluntly, completely asinine.⁵ In a related context, [Anspach and](#)

³It is possible that these effects may represent learning new, salient information about candidates in addition to increasing the salience of topics that voters may already be familiar with about the candidates.

⁴Previous research has also found that engagement in contentious political discussions on social media increases political persuasion ([Gil de Zúñiga, Barnidge and Diehl, 2018](#)).

⁵In hand-coding all of the comments in our streaming chat sample, our favorite example was “WHAT IS CORY BOOKER’S REAL NAME????? what is he hiding”.

Carlson (2018) find that social commentary on news posts shared on platforms like Facebook is both misleading yet consequential. Holding the content of a news story’s summary constant, the addition of social commentary that misrepresents that summary can cause people to recall incorrect information rather than that contained in the story. The social information is more salient.

In many cases, the topics or “primes” present in the debate comments are bad for the democratic process, encouraging voters to evaluate candidates on criteria that are unrelated to their performance and which reinforce existing inequalities. Consider the example of Biden’s acuity. During the debate in this study, Julian Castro specifically criticized Biden’s age, and many comments in the streaming chat mentioned Biden’s health or acuity at various points when Biden was talking. This immediacy could supplant whatever policy position Biden was attempting to link himself to, instead priming concern about his fitness for office. Particularly in a primary election, unlike the debate stage, the debate comment section may include detractors from the opposing party who discuss topics that represent negative primes that will be used by the opposing party in the general election, but which may be considered too damaging to employ by copartisans in the primary.

On this pathway, the streaming social and expert chats are substantively similar in terms of theoretical predictions. Our first hypothesis for priming effects is straightforward: that the candidates mentioned most often in the streaming chat will be salient in the minds of viewers, and therefore, experience an increase in name recognition and familiarity. The second hypothesis could not be fully specified in advance, and was not pre-registered. We anticipate that the frequency and valence of particular candidate-specific topics will influence their candidate evaluations. Empirically, we focus on the social chat. We anticipate that the presence of detractors on social chat will uniquely increase the quantity of prominent negative primes about the candidates, and we hypothesize that this will lead to a decrease in evaluations of the candidates who are the focus of these negative comments.

- **Hypothesis 3:** Candidates mentioned most often in the streaming chat will see an increase in name recognition.
- **Hypothesis 4:** Viewers of the streaming chat will reduce their evaluation of candidates in the presence of a prominent negative prime.
- **Hypothesis 3(e):** Candidates mentioned most often in the expert chat will see

an increase in name recognition.

- **Hypothesis 4(e):** Viewers of the expert chat will reduce their evaluation of candidates in the presence of a prominent negative prime.

Context: Inferences About the Public

Our last set of hypotheses stem from two premises. First, streaming social chats are more likely to be composed of low-quality, toxic comments than other forms of media due to the quasi-anonymity of comments. Second, the public may incorrectly project from these comments what the broader public thinks about a livestreamed event. Below we explain both of these claims to derive our hypotheses.

One key feature of the social feeds coordinated using hashtags on Twitter that are standard in the dual screening literature is that the streams are semi-permanent, and each account that comprises the stream conveys significant social information in its profile. In contrast, the stream in the embedded Facebook chat is made quasi-anonymous by the speed at which it flows; many other streaming chats are completely anonymous. In the first block of the debate, for example, we estimate there were 60.3 comments per minute. While viewers might begin to notice frequent commenters, there were too many for long-term reputational costs to obtain.

This quasi-anonymity has significant implications for the types of people who comprise the streaming chat and for the types of messages they send.

For instance, anonymity increases "flaming" or personal attacks in online communities (Muneeam and Crandall, 2011), and removing anonymity can elevate the level of civility in discourse in online newspaper comments sections (Santana, 2014). The presence of anonymity also moderates the capacity for norm enforcement online; for moderate norms like that against political incivility, anonymity decreases the effectiveness of norm enforcement (Munger, 2017).

The presence of high levels of toxicity in a chat room can produce a positive feedback loop, as people with a low tolerance for toxicity opt out of sending messages (Theocharis et al., 2016). Furthermore, joining an online community where toxic language is normalized causes people to comment in a more toxic manner Cheng et al. (2017). As a result (and as we demonstrate empirically below), streaming chats are likely to be highly toxic. This toxic partisan incivility provides an explanation for how cross-partisan communication can actually *increase* affective polarization.

Far from the ideal conditions under which the contact hypothesis predicts will decrease intergroup hostility, watching toxic streaming chats entails spending time with the cruelest, most aggressive and least thoughtful group of your partisan opponents. Although these commenters are likely a radically non-representative sample, the conditions of streaming chat obscure this fact. The presence of incivility in online political conversations has been shown to increase affective polarization (Suhay, Bello-Pardo and Maurer, 2018), and the current context further increases the likelihood that subjects will update their opinion of their fellow citizens.⁶

Barnidge, Gil de Zúñiga and Diehl (2017) demonstrate the importance of second screening by showing how it can enhance persuasion among viewers of political content. Theoretically, this occurs because the second screen provides “social cues...[which] provide direct evidence of social opinion, which people take to be representative of public opinion, even if they are not” (p313). This last claim references Lerman, Yan and Wu (2016), who demonstrate the existence of the “majority illusion” that can be produced by social networks.

Although communication scholars tend to rely on psychological processes to explain the effects of new media technologies on beliefs, logic like that demonstrated in Lerman, Yan and Wu (2016) makes predictions that are observationally equivalent using a rational actor framework. Economists have recently been developing theories of “correlational neglect” that describe the fact that people tend to underestimate the amount of correlation between different sources of information they encounter, particularly on social media.

Social media is much “denser” than either traditional media or socialization; the consumer is able to observe a large number of signals in a very short period of time. The issue is that our intuitions about the informational content of each signal are misleading because the signals are highly correlated with one another. Although this effect should not change the *direction* in which a social media updates her beliefs about the world, it does tend to magnify any belief updating, resulting in what Ortoleva and Snowberg (2015) describe as overconfidence, which in turn causes ideological extremeness.

In the aggregate, however, this may not be normatively undesirable; (Levy and Razin, 2015) describe the conditions under which an electorate with “correlationally neglectful” voters can actually make better decisions. Voters generally tend to un-

⁶This theory comports with survey evidence from Pew, in which “64% say their online encounters with people on the opposite side of the political spectrum leave them feeling as if they have even less in common than they thought” (Duggan and Smith, 2016).

dervalue novel information relative to their longstanding beliefs about the world (see, for example, the persistence of partisanship), and correlation neglect cuts the other direction.

Enke and Zimmermann (2017) provide experimental evidence for the theoretical model developed in (Levy and Razin, 2015), and demonstrate that the primary cause behind correlational neglect is that calculating correlations is cognitively taxing. Subjects with lower scores on cognitive tests exhibit higher correlational neglect, which cannot be reduced by tripling the rewards for accuracy. Experimentally increasing the complexity of the task increases correlational neglect.

This combination of low-quality comments and a context that encourages viewers to make inferences about public opinion as a whole produces three sets of hypotheses:

- **Hypothesis 5:** Streaming chat will cause a decrease in trust in the process of debates and democracy.
- **Hypothesis 6:** Streaming chat will cause an increase in affective polarization.
- **Hypothesis 7:** Respondents will change their estimate of the success of each candidate in future polls based on the sentiment directed at that candidate in the streaming chat.

The context parameters do vary significantly between the two chat conditions. The expert chat is civil and reasoned (bordering on wonk-y), evincing a deep enthusiasm for the process of debates and democracy. Furthermore, the expert chat clearly represents itself as such, and should thus not cause viewers to change their opinion of others' opinions about the candidates.

- **Hypothesis 5(e):** Expert chat will cause an increase in trust in the process of debates and democracy.
- **Hypothesis 6(e):** Expert chat will cause a decrease in affective polarization.
- **Hypothesis 7(e):** Respondents will not change their estimate of the success of each candidate in future polls based on the sentiment directed at that candidate in the expert chat.

Research Design

To examine the effects of media platforms on political attitudes, we conducted a digital field experiment using Amazon Mechanical Turk (MTurk) to recruit participants. Recruitment took place in two parts. During Wave 1, a Human Intelligence Task (HIT) was made available to all MTurk workers residing in the US with an approval rating of 95 percent or higher. Wave 1 included 2352 respondents. Respondents who specified that they were likely to watch the debate, had a Facebook account, and could watch the debate on a computer were deemed eligible and invited to participate in a second survey after the debate ⁷

A total of 1315 respondents were qualified for participating in the second survey, and 1095 of these respondents indicated an interest to participate. Our sample includes only people who are interested and likely to watch political debates and able to watch these debates online, which means our treatment effects are estimated on a sample that reflects a population that might plausibly be “treated” in this manner during similar events in the real world, an advantage for external validity.

Experimental conditions: ABC, Facebook and FiveThirtyEight

Our experiment is an encouragement design with three conditions. The design is modeled after Gross, Porter and Wood (2019) who investigate the effect of viewing political debates on different broadcast networks by using an online field experiment to randomize the channel where subjects viewed the debate. In our study, respondents who indicated that they were interested in participating in the second survey were randomly assigned and asked to watch the debate on one of the following livestream platforms: FiveThirtyEight (Expert condition, N=243 Democrats; N=364 all), ABC News Facebook page (Social Media condition, N= 263 Democrats, N=365 all), or ABC News (Control condition, N= 264 Democrats, N=366 all). Respondents were given links to these websites at the end of the Wave 1 survey and were also sent a reminder email prior to the debate with a direct link to the livestream. For screenshots of three debate platforms, see Figure 10 in the Appendix. At the end of the debate, a link to the follow-up survey was emailed to each respondent who indicated an interest in participating in Wave 2.

⁷Eligible respondents were offered a bonus payment of \$1.50 and an entry into a \$100 raffle for their participation in the second survey

During the Wave 2 survey, respondents were asked questions about their reactions to the debate, their emotional state, predictions regarding polls, familiarity with the candidates, trust in the government, and political polarization. Respondents also rated their feelings towards the two major political parties and the Democratic presidential candidates during a series of questions using feeling thermometers. After answering the survey questions, respondents provided responses to detailed compliance checks, provided open-ended responses noting what stood out about each candidate, and were then re-asked their party identification. (See online appendix for question wording.)

The Wave 2 respondents are generally balanced across assignment conditions: N=305 in the ABC Control condition (84% recontact rate), N= 298 in the FiveThirtyEight Expert condition (82% recontact rate), and N=305 (84% recontact rate) in the Facebook Social Media condition. Our sample includes 648 Democratic respondents who completed Wave 2,⁸ and 576 Democratic respondents who completed Wave 2 and reported watching at least part of the debate.

Our primary analyses will be limited to this subsample of Democrats who watched at least part of the debate, as many of the questions in our follow-up survey specifically ask about reactions to the debate, for which respondents can only answer if they did watch. We focus on Democrats, as they comprise the substantial majority of the sample, and we anticipated potential heterogeneous effects by partisanship.⁹ The table below summarizes the number of respondents in each condition assigned during Wave 1, as well as respondents who will comprise the analysis sample of Democrats who watched at least part of the debate.

As with most encouragement designs, treatment assignment does not equate to the receipt of treatment. Our study has two-sided noncompliance (Gerber and Green, 2012) where some subjects in the Control condition who watched the debate may have been “treated” by viewing social or expert commentary, and some subjects in the treatment conditions who watched the debate may have opted not to view the debate on the assigned platform (going untreated).¹⁰ We conduct supplemental analyses that account

⁸This subsample includes similar recontact rates across conditions: control (N=224, 84.8% recontact rate), social (N=221, 83.5% recontact rate), and expert condition (N=182, 84.0% recontact).

⁹Supplemental results for Republicans are in the appendix.

¹⁰At the end of the Wave 2 survey, we asked how and on what platform respondents watched the debate. About 70% of respondents (73% of Democrats) self-reported watching at least part of the debate and watching the debate on the platform where they were assigned. Self-reported rates of watching the debate on the assigned platform was somewhat higher in the control (74% overall, 75% among Democrats) and social (75% overall, 77% among Democrats) conditions than the expert condition (62% overall, 66% among Democrats). We believe the lower propensity for respondents to

for compliance in the results section. Additional details on compliance and balance across experimental conditions are in the appendix.

Table 1: Sample sizes by condition

	Full Sample	Control	Expert	Social	Total N
	Assigned	366	364	365	1095
	Participated in Wave 2	305	298	305	908
Participated in Wave 2 and Watched Debate		279	254	276	809
Democrats Only (including leaners)		Control	Expert	Social	Total N
	Assigned	264	243	263	770
	Participated in Wave 2	224	203	221	648
Participated in Wave 2 and Watched Debate		204	174	198	576

Analysis of Expert and Social Media Comments

Our results first examine the content of the streaming chats on the ABC News Facebook page or FiveThirtyEight. We performed a series of computational text analyses of the comments found in each stream and a hand-coding of the Facebook stream for subject, tone and topic. Our text analyses are organized to focus on the three theoretical mechanisms through which we believe chat streams may influence perceptions of the debate and the candidates: frequency of the chat, topics discussed that may prime viewers, and the context or tone of the commentary.

To perform the analysis, we scraped the comments from each of the streaming chats. For the FiveThirtyEight chatbox, we were able to extract all the comments made by experts during the debate. For Facebook, we first manually expanded the chatbox, then we extracted the comments from the available HTML file. We were not able to extract all the comments in the chatbox, but a large sample of the comments which Facebook makes available through our manual search. In addition, we did not collect replies to the comments and reactions from the users. The Facebook streaming chat started before and continued after the debate. We decided to analyze comments made during the three-hour debate broadcast windows, since we asked our respondents to watch only the debate.

watch the debate and watch the debate on the FiveThirtyEight website, when assigned, may be due to the quality of the viewing platform for the debate. See appendix for details.

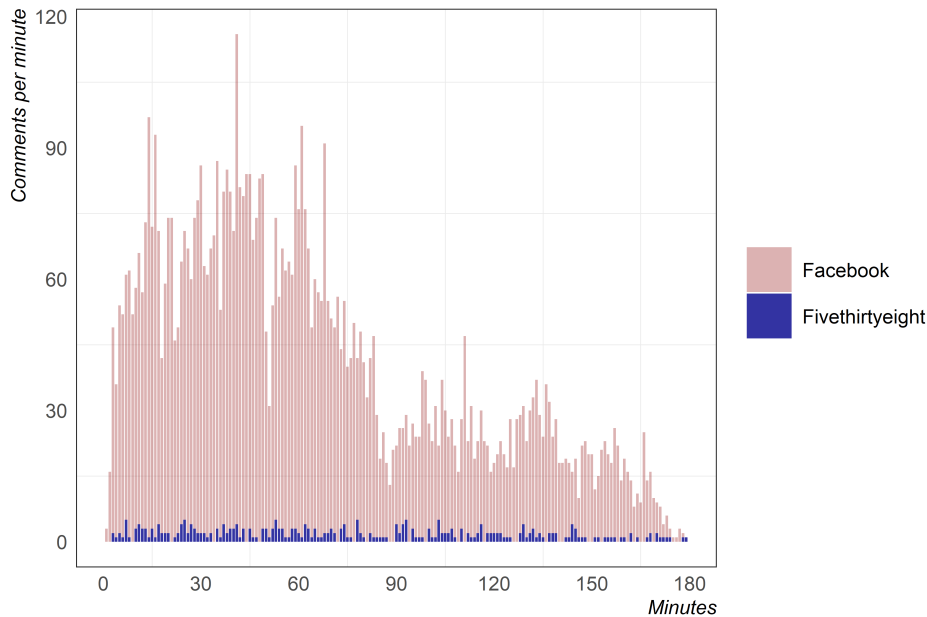
Frequency of Comments

In the expert chat, we collected 314 comments with 23,437 words in total, and 74.6 words on average by comment. On Facebook, we have a sample of 6,915 comments with 53,092 words and an average of 7.75 words by comment.

Overall, comments in the social condition are more frequent and considerably shorter compared to the expert chat. Assuming our sample represents most of the comments a user would be exposed to during the time of the debate, based on our sample, users observed 1.68 comments per minute in the expert chat and 38 comments per minute in the social chat. The difference in frequency makes sense with our theory of a more chaotic environment on the social condition chatbox compared to the expert chat.

Figure 1 presents the evolution of the comments by minute on both conditions. The figures clearly depicts the high frequency of comments on the social condition, and the concentration of them in the first hour of the debate. On the other side, the expert chat is less intense and more evenly distributed across the four hours of debate. In the first block of the debate, for example, there were 60.3 comments per minute on Facebook, and only 2.3 on the expert chatbox.

Figure 1: Distribution of the Comments by minute during the Debate



Context

In this second section, we assess how participants are communicating on the chat streams, looking at various dimensions of toxicity. These aggregate measures may influence overall levels of trust in the democratic process and levels of polarization.

We use Google’s *Perspective* API, a content moderating tool that is the industry standard for automatic detection of toxic content in written comments.¹¹ We present the results considering the levels of toxicity using the score created by the model classifying the comments.

Figure 2 presents the proportion for the toxicity scores in four dimensions - toxicity, severe toxicity, threat and insult. We calculate the proportion classifying as “1” all the comments which surpass the score of 0.5 in each dimension. The differences between the social media and expert chat are stark – levels of toxicity in all the four dimensions are undetectable in the expert chat, while more than 15% of the comments on Facebook are considered toxic. Overall, the Facebook chat exhibits high levels of negative comments, with insults and toxic comments appearing often during the debate. (See Appendix for the full distribution of toxicity scores by condition.)

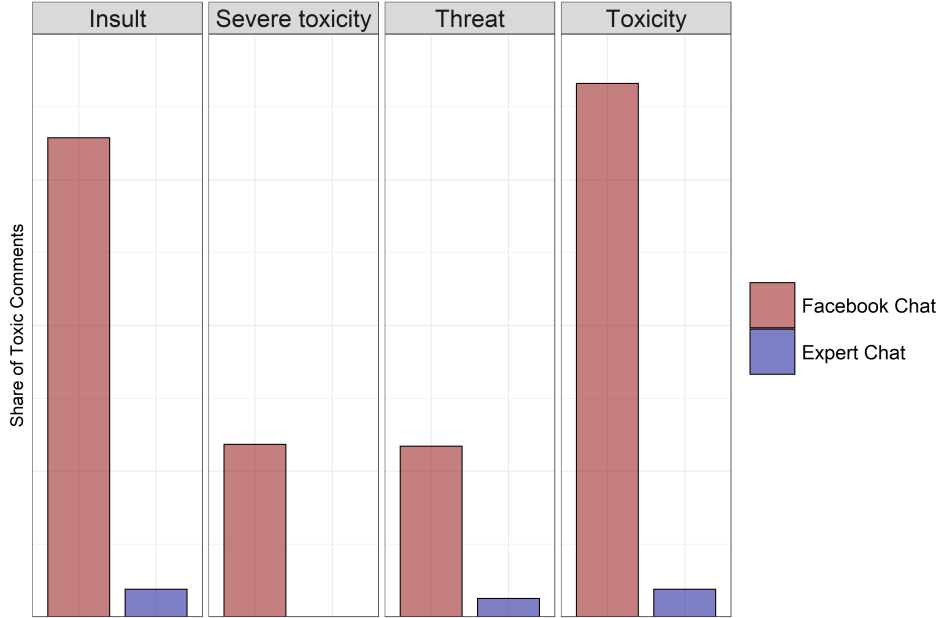
Content: Priming

Our theory expects priming effects from particular characteristics or issues raised on the chatboxes during the debate. Our findings so far indicate that toxicity and negative tone is predominant on the Facebook chatbox compared to the expert condition. However, our automated methods do not reveal which issues or characteristics of the candidates were more salient for the users talking about the debate. Therefore, with the purpose of identifying these issues, we hand-coded 6,500 comments in the social condition.¹² In our hand-coding, we focus on the Facebook comments because our goal is to disentangle the priming effects, in particular negative priming, among the social and the expert

¹¹Perspective uses a convolutional neural net model to score the toxicity of an input text. Toxic is defined as a rude, disrespectful, or unreasonable comment that is likely to make one leave a discussion. The model was built using millions of comments from the internet, using human-coders to rate the comments on a scale from very toxic to very healthy, and using this large data as training information for the machine learning algorithm. We uploaded the content of the comments in each treatment condition.

¹²We hand-code comments during approximately the first 2.5 hours of the debate. We stop coding during a moment when protesters disrupted the debate, and many comments started describing the debate as “over.” This decision focuses our coding on the substantial majority of comments that were made during the active debate.

Figure 2: Proportion of Toxic Comments



condition, and allowing a better understanding of the experimental results.

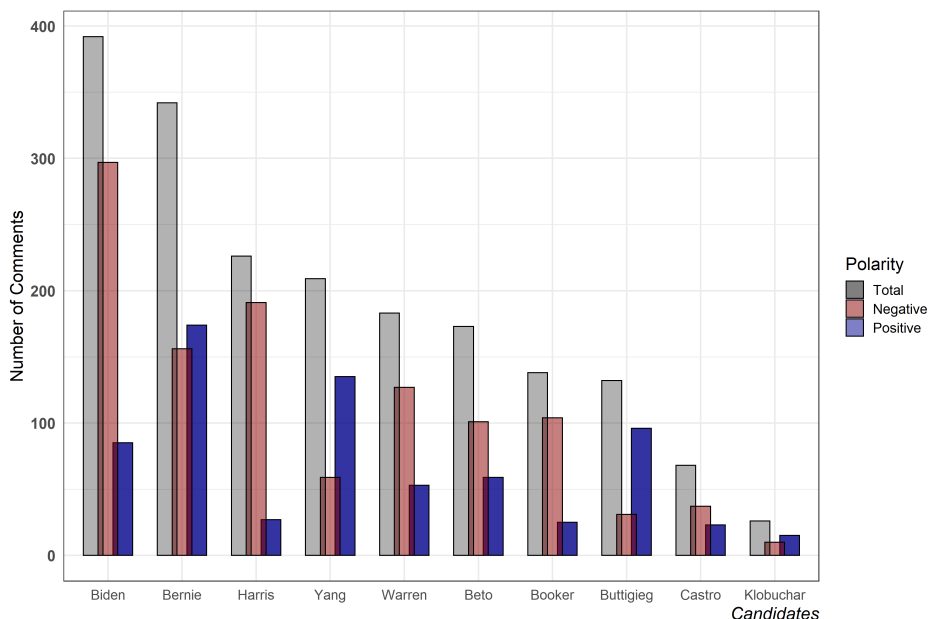
First, we read all the comments on the Facebook streaming chat with the purpose of classifying when a comment was directed at one of the candidates in the debate.¹³ We considered a comment as being directed to the candidates when the text explicitly mentioned the candidates, or implicitly making a reference to some policy, characteristics, or political stand of the candidate. Three coders hand-coded all the comments, and, to avoid false positives, we included only comments classified as being about the candidates by at least two out of the three coders. We end with a sample of 1,889 comments. Two additional coders read through all the selected comments, and classified the content according to the topics discussed and sentiment or polarity (whether a comment was negative, neutral or positive about the candidate).

Figure 3 presents our results. On the left is the the total number of comments about each candidate in the Facebook streaming chat, and the overall polarity associated. Overall, Biden and Sanders received most of the attention followed by Harris and Yang. The strong presence of Yang is surprising considering he was not one of the front-runners of the primaries; he was, however, widely considered the “candidate of the internet,” and his Universal Basic Income proposal generated a lot of attention. For most of

¹³Supplemental analyses in the appendix use automated dictionary-based methods for describing the frequency of candidate mentions and the sentiment of comments.

the candidates, the comments are mostly negative, with special attention to Harris and Biden, with 85% and 75% of negative comments out of their total share. On the other side, Sanders, Yang, Buttigieg and Klobuchar are the only candidates who received proportionally more positive comments, though Klobuchar received very few comments, overall.

Figure 3: Comment Polarity by Candidate, Facebook Streaming Chat



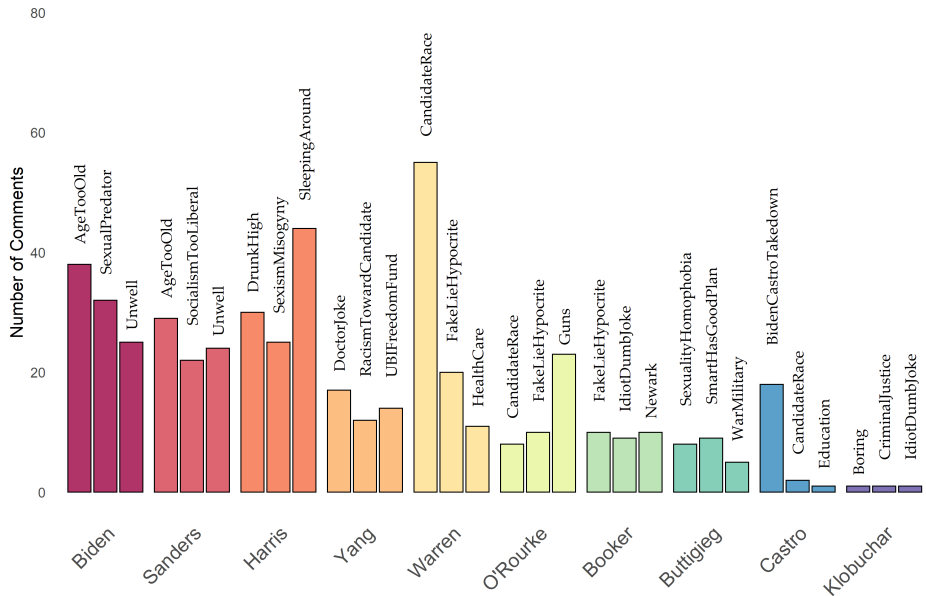
To understand which issues are being primed in the streaming chat, Figure 4 displays the results of our hand-coding of the comments. We read each comment and classified the subject of the comment. Comments were allowed to be about multiple topics, and if a comment was unclear or its meaning ambiguous, it was left uncoded.¹⁴ This analysis gives a much richer understanding of the content of the comment. Because most of the comments are negative, this, in part, means categorizing the types of insults directed at the candidates.

Overall, there was a remarkable degree of consistency in these attacks—so much so that several commenters accused commenters of being Russian trolls. We were

¹⁴Appendix Table 4 presents a more detailed analysis of the words that define these hand-labeled topics. We also omit from this analysis comments that were generically negative or positive, such as “Yang 2020.”

unfortunately unable to collect any information about these commenters, but there were clearly some copy-pasted textual “memes” repeated throughout the debate. We think it is entirely possible that a significant number of the comments were produced by a small number of trolls foreign or domestic. This is not a limitation of the current study: this is a realistic feature of open Internet platforms in 2020.

Figure 4: Top Three Topics for each Candidate



Biden was the most frequently mentioned candidate, and the top three topics were all extremely critical: mocking his age (topic labeled AgeTooOld); accusing him of being creepy or handsy (topic labeled “SexualPredator”); and suggesting that he was physically unwell. Sanders was also mocked for being old, and was often described as a socialist or too far to the left.

These criticisms were harsh, and while they had generally been avoided during the Democratic primary, in that these criticisms were not brought up by their debate opponents¹⁵ they may not be considered outside the bounds of legitimate democratic deliberation. The comments directed at Harris, however, were generally beyond the pale. The most common comment accused her of “sleeping her way to the top,” in

¹⁵The exception happened during this debate: Castro intimated that Biden was too old to be running, which generated the vast majority of comments about Castro during this debate.

sometimes graphic terms. Many others made fun of her affect (e.g., suggesting she was “drunk”), and/or were straightforwardly misogynistic.

Other notable topics include praise for Yang’s UBI proposal and Buttigieg’s polished and well-researched plans. There were several mentions of O’Rourke’s strong anti-gun statement and Booker’s history as mayor of Newark. The most common single-candidate topic, however, was mockery of Warren by reference to her having claimed Native American heritage. Many of the comments in this category were simply echoes of Trump’s “Pocahontas” moniker.

Experimental Results

We now turn to our survey experimental results. We present results for Democratic respondents (including those who lean toward the Democratic Party when forced to choose between the Democratic and Republican Party) who completed the Wave 2 survey indicated watching at least part of the debate (N=576).

In each analysis, we present two types of average effects: First, we report the average marginal effect of assignment to the Expert condition, relative to the Control condition, and assignment to the Social condition, relative to the Control. These are calculated using linear regressions of each outcome on the treatment assignment indicators. The figures to follow present the coefficient point estimate, as well as 90% and 95% confidence intervals.

In addition, we report one type of complier average causal effect (CACE) using instrumental variables (IV) regression. To do so, we consider the two pairwise comparisons of Expert vs. Control and Social vs. Control, separately. We code each respondent in the analysis as being simply “treated” or “untreated” for each comparison based on whether or not they reported watching the debate on the assigned platform.

For each comparison, we conduct a two-stage least squares regression where treatment assignment is used as an instrument for whether the respondent actually received the treatment (Gerber and Green, 2012). We find compliance rates of 75% for the Expert condition and 81% for the Social condition by this method.¹⁶ We report the point estimate of the CACE and 90% and 95% confidence intervals for this analysis in the

¹⁶This is calculated within the sample of Democrats who watched at least part of the debate. Compliance is based on the first stage regression where treatment receipt is regressed on treatment assignment. The CACE is the ratio of the coefficient for the effect of treatment assignment on each outcome over the compliance rate.

figures. We interpret these effects as local— the effect only among those who complied and watched the debate on the assigned platform because they were assigned to do so.¹⁷

Testing Hypotheses Related to Frequency Effects

Our first set of hypotheses discussed how participants would rate their overall debate experience. In particular, we hypothesized that streaming social chat may lead viewers to find the debate less enjoyable and informative, though potentially more engaging due to the increase in stimuli. In addition, we hypothesized that the comments in social chat may also increase the extent to which viewers experienced negative emotions during the debate—specifically, anger and anxiety. We generally anticipated the opposite effects for the experience of the expert chat.

Figure 5 displays the results. Overall respondents assigned to the Social condition had more negative experiences with the debate than respondents assigned to the Control condition. As expected, respondents who were assigned to watch the debate on the Facebook platform found the debate both less informative (n.s.) and less enjoyable ($p < 0.05$) on average. Contrary to our expectations, respondents assigned to the Social condition also found the debate less engaging ($p < 0.05$). After adjusting our results for multiple comparisons based on the five outcomes, the p-values for how enjoyable and engaging the debate is move above conventional levels of statistical significance. Respondents in the Social condition also reported slightly, but not significantly, higher rates of feeling angry and anxious from watching the debate. Thus, at minimum, the results show that the encouragement to watch the debate on Facebook, did not lead debate watchers to have a more satisfying experience. Instead, we have some support that exposure to social media comments can lead to both a less informative and overall negative experience.

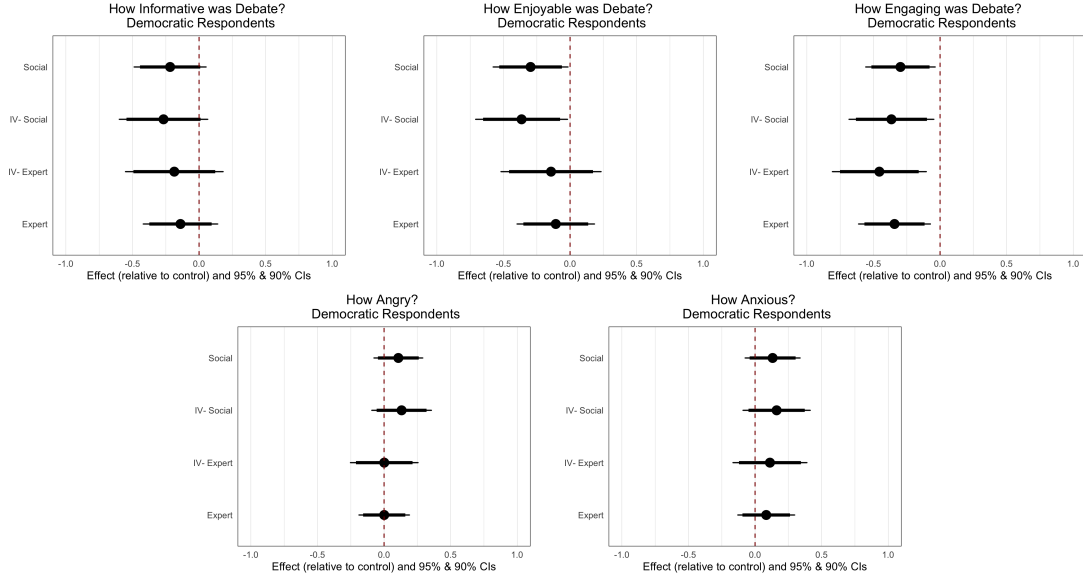
In addition, contrary to our expectations, assignment to the Expert condition did not have opposing effects to the Social condition. Respondents in the Expert condition generally expressed similar, though often slightly more muted, reactions to those in the Social condition. In almost every case, the effects (relative to the Control) were indistinguishable from zero. The one exception is that respondents in the Expert condition found the debate significantly less engaging.¹⁸ In this case, the increase in stimuli

¹⁷Note: Our design may include "always-takers" and "never-takers" who do not adjust their viewing behavior based on their treatment assignment. We assume no defiers.

¹⁸This result remains significant at the $p < 0.10$ level after adjusting for multiple testing of the five outcomes.

served only to turn people off of the debate.

Figure 5: Debate Experience



Testing Hypotheses Related to Content Effects

Our second set of hypotheses anticipates that there will be a connection between the focus of the comments and the extent to which respondents report they are now familiar with the candidates after the debate. The results are displayed in Appendix Figure 11. Overall, we see little movement in either the Social or Expert conditions relative to respondents assigned to watch the debate on the ABC News Website. We suspect that this may be because the focus of the comments in the chat streams often tracked with the amount of speaking time each candidate had and pre-existing poll performance.

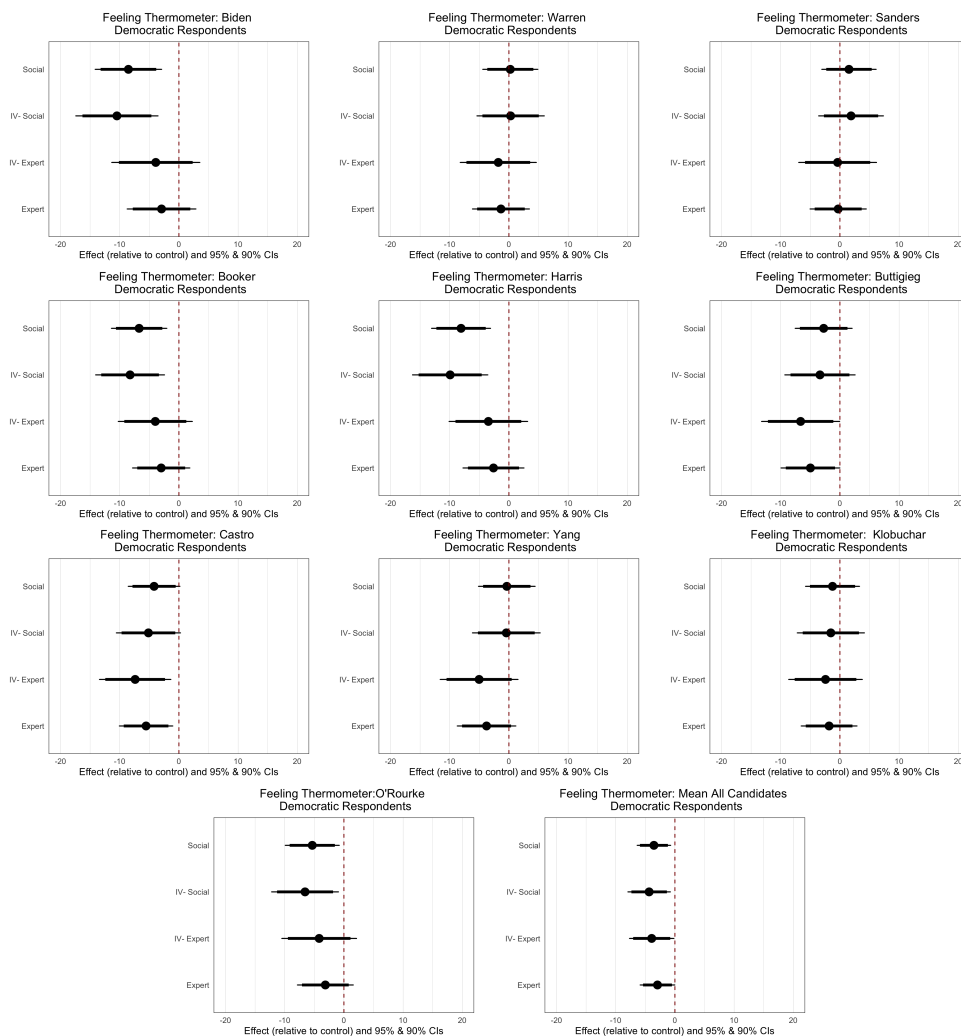
We also hypothesize that candidate-specific favorability may be influenced by the degree to which candidates were discussed in negative ways in the commentary. To assess this, we examine feeling thermometer ratings toward each of the candidates (Figure 6).

As shown in Figure 6, we find that respondents in the Social condition came away with more negative perceptions of all debate participants, on average (bottom right), relative to the Control condition.¹⁹

¹⁹Figure 12 plots the average feeling thermometers for each candidate by condition. Overall, Sanders and Warren had the warmest average ratings, while Klobuchar had the least warm ratings. However,

Comparing across conditions, respondents assigned to watch the debate on the ABC News Facebook page came away with more negative feelings toward Biden, Harris, and Booker. These results remain significant ($p < 0.05$) when adjusting for ten candidate outcomes. Respondents also become more negative in their feelings toward O’Rourke and Castro, though these results move above conventional levels of significance after adjusting for multiple testing. Results for the Expert condition are generally more muted, except in the case of evaluations of Castro, which are significantly more negative in the Expert condition.

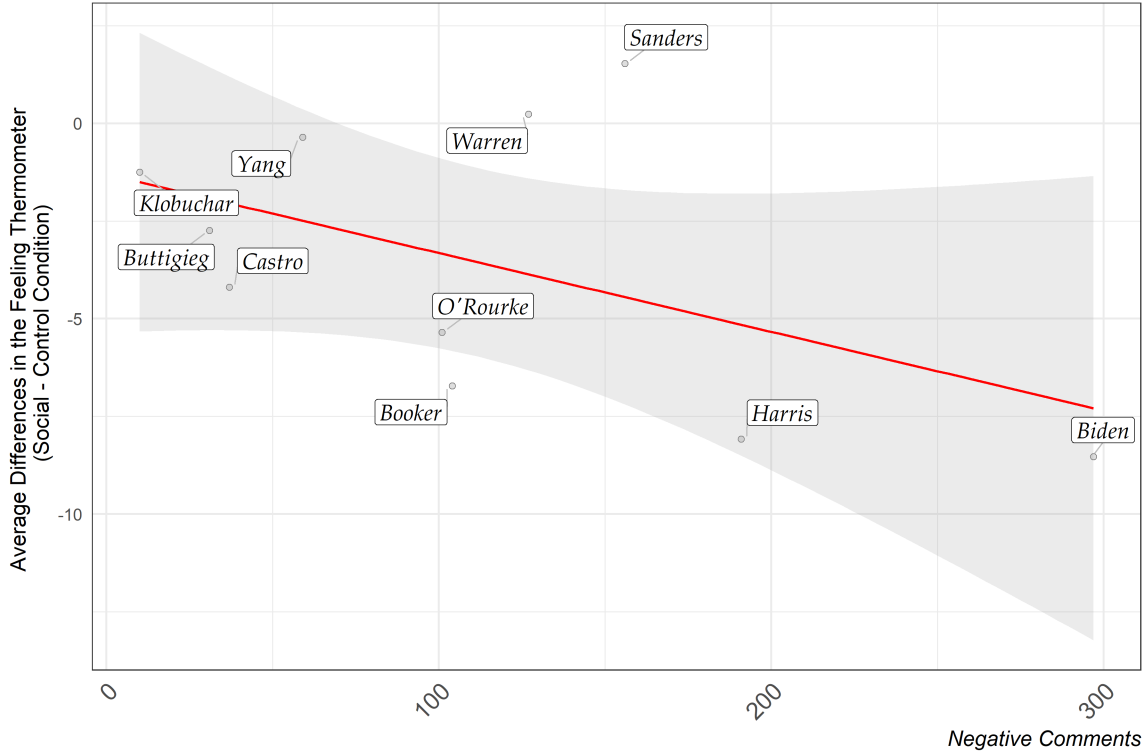
Figure 6: Candidate Feeling Thermometers



candidates generally received neutral to positive ratings. No Democratic debate candidate was viewed extremely coldly, on average.

These candidates, in particular Biden, Harris, Booker and O’Rourke, are among those who received proportionally more negative comments in our text analysis of the social streaming chat. To provide a intuitive visualization of this pattern, Figure 7 plots the average difference in the feeling thermometer between the social and the control condition for each candidate on the sum of comments we labelled as negative on the Facebook streaming chat.

Figure 7: Negative Comments Decrease Candidate Feeling Thermometers



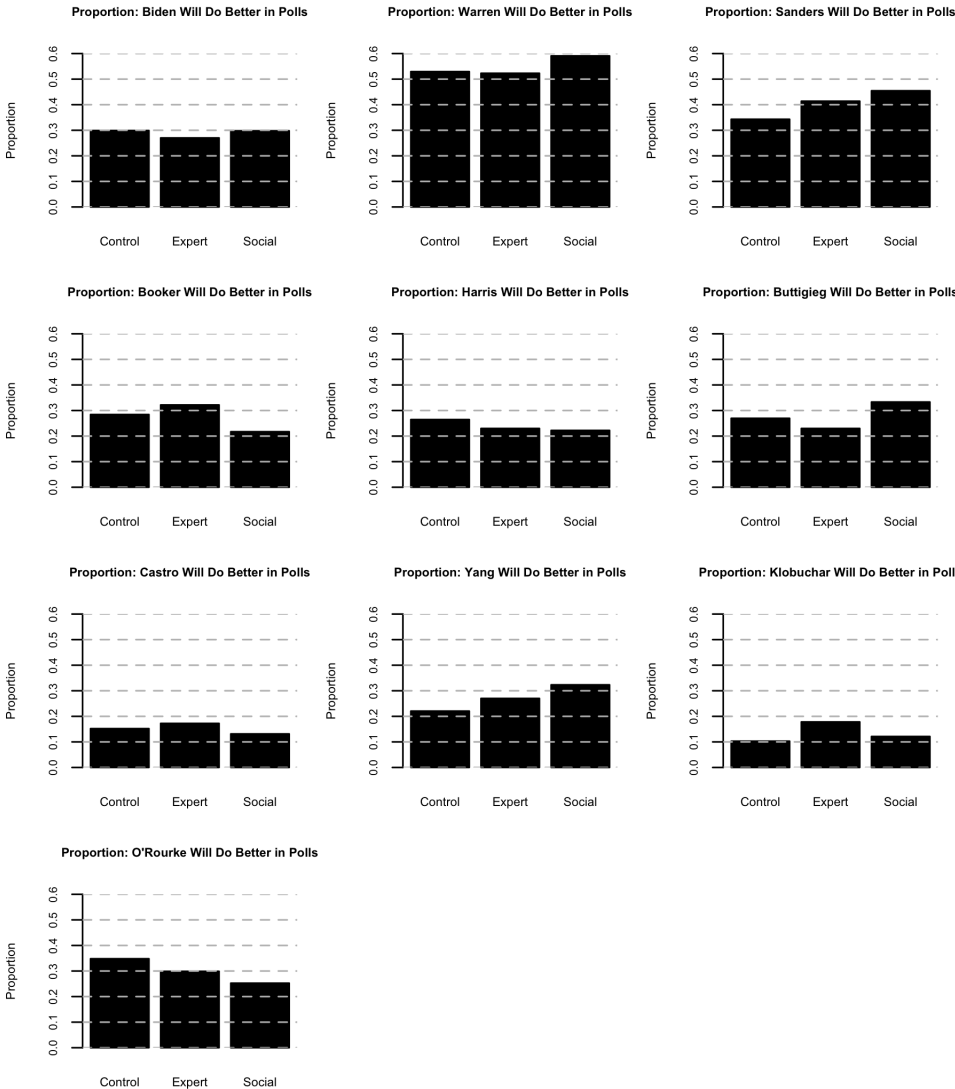
Testing Hypotheses Related to Context Effects

A final set of hypotheses assessed the extent to which respondents would infer from the commentary how the general public views the candidates. To test this, we assess whether respondents think that the candidate will do better in the polls as a result of the debate (Figure 13).

Across all conditions, respondents indicated that some candidates were more likely to perform better in the polls after the debate than other candidates. Figure 8 plots the proportion of respondents who indicated a candidate would likely do better in the

polls the following week by condition. We notice some variation by condition, though the variation between conditions is often smaller than the variation between candidates. Respondents in each condition pointed to Warren and Sanders as the better-performing candidates, while respondents were less likely to indicate that front-runner (at the time) Biden would do better in the polls after his debate performance.

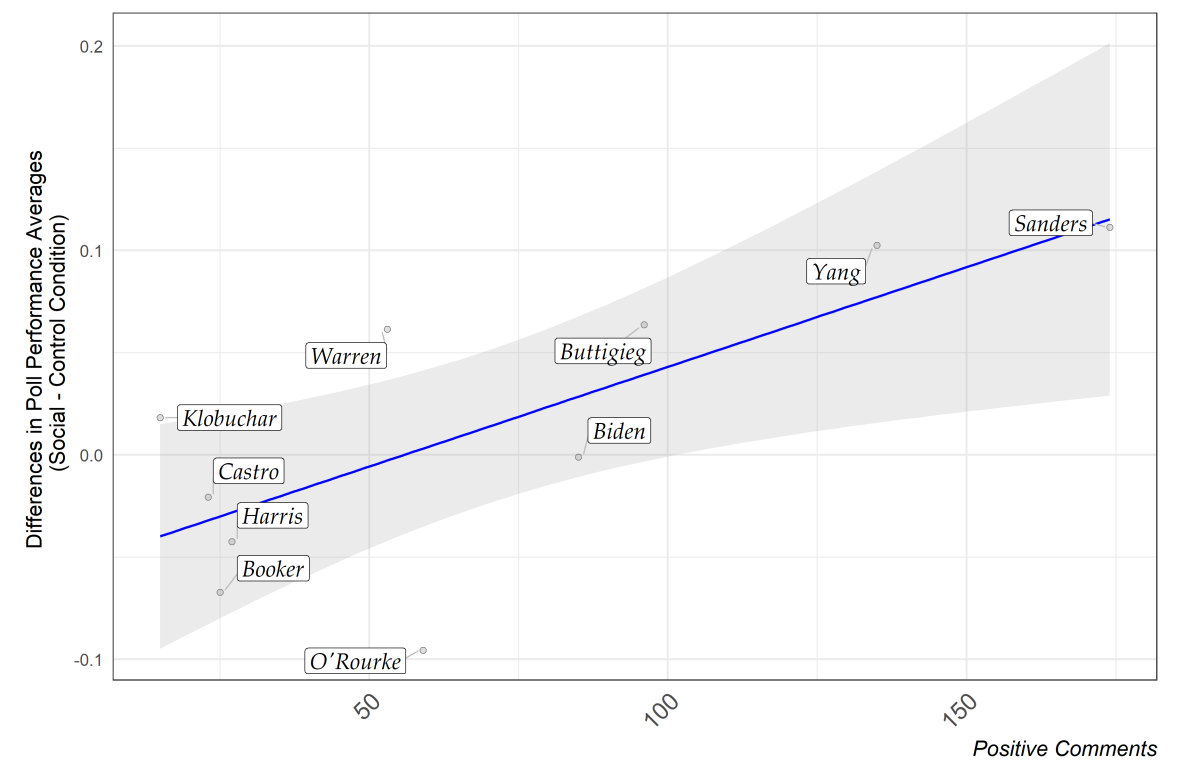
Figure 8: Candidate Poll Performance Averages



Respondents in the social condition were less likely to predict Booker or O’Rourke would perform well in the polls, as we indicate on figure 8, and respondents in the social condition were more likely to predict that Yang and Sanders would do better

in the election polls after the debate. Notably, Yang and Sanders were candidates who received relatively more positive comments on the ABC News Facebook stream compared to other candidates. Figure 9 displays the correlation between the number of positive comments a candidate received in the social chat and the difference in projected poll performance between the social and control conditions.

Figure 9: Positive Comments Increase Expected Poll Performance



In addition, we hypothesized that viewing the debate with social chat may lead respondents to become more polarized, and decrease their trust in the political process. Similar to the first set of hypotheses, we do not have shared expectations for the Expert chat. Instead, we anticipated that viewing expert analysis may help to increase trust and reduce polarization. Figure 14 (Appendix) displays feeling thermometer ratings toward the Democratic Party, Republican Party, and a measure of affective polarization (the difference between these two ratings). Respondents in the treatment conditions tended to come away from the debate with slightly, but not significantly, more negative ratings of the Democratic Party (in contrast to their more negative feelings about many Democratic debate participants). We find null results for treatment effects on ratings

of the Republican Party of affective polarization.

Similarly, despite our expectations, we find no significant effects of assignment to the treatment conditions on perceptions of trust in political parties, the wisdom of the public in making political decisions, the desire to want political officials to work with those with whom they may disagree, or perceptions of the role of the debates as part of the democratic process (Figure 15, in the Appendix).

Conclusion

Does streaming chat alter viewer perceptions of political debates? We first demonstrate that real-time social commentary can be, in practice, a highly toxic, low-quality, overwhelming, and negative experience that differs greatly from watching a political event without comments or with more in-depth and slow-paced expert commentary.

Table 2 summarizes our hypotheses for frequency, content, and context effects on experiment participants. First, we anticipated that the introduction of a social chat stream could be distracting for participants due to the quantity and fast-paced nature of the comments, leading to a less satisfying and emotionally taxing experience. Our text analysis results validated our expectation that the social chat had very frequent comments, and our survey results find that respondents assigned to the social condition came away with slightly, though not always significantly, less enjoyable, less informative experiences and felt more anxious and angry than respondents in the ABC News condition. In addition, while we anticipated the stimulation of the chat could still lead participants to feel engaged in the video more so than watching the debate without real-time commentary, we found that respondents in both the social and expert conditions found the debate significantly less engaging.

We did not find support for our hypotheses that the expert commentary would lead viewers to find the debate more informative and enjoyable and less emotionally taxing. As we discuss in the main text, part of the reason for these more muted effects could be due to the structure and design of the FiveThirtyEight stream, which made it harder to watch the debate video.

Second, we hypothesized that the content of the streaming chats would influence name recognition and candidate evaluations. We did not find effects on name recognition. However, we do find evidence that aligns with potential priming effects. Here, we focus our analysis on our primary treatment— the social chat. Based on content

Table 2: Summary of Hypotheses

	Social	Expert
Frequency		
Hyp 1	Debate less enjoyable and informative	Debate more enjoyable and informative
	Debate more engaging	Debate more engaging
Hyp 2	More angry and anxious	Less angry and anxious
Content		
Hyp 3	Increase in name recognition for candidates mentioned most often	Increase in name recognition for candidates mentioned most often
Hyp 4	Reduce evaluation toward candidates with negative primes	Reduce evaluation toward candidates with negative primes
Context		
Hyp 5	Decrease in trust	Increase in trust
Hyp 6	Increase in affective polarization	Decrease in affective polarization
Hyp 7	Infer future poll performance based on sentiment in chat toward candidate	No change in inferred future poll performance

analysis, candidates who were subject to high frequencies and proportions of negative comments, such as Biden, Booker, and Harris, were rated significantly lower in feeling thermometer ratings by respondents in the social condition relative to those in the ABC News condition.

Lastly, we hypothesized that the negative, toxic nature of the comments may spill over into evaluations of the overall democratic process and levels of polarization. We do not find treatment effects on these outcomes. We do, however, find evidence that people may infer—potentially incorrectly—that the views expressed on the streaming social chat may reflect the sentiment of the public, more generally. For example, we find that candidates who received more frequent positive comments, such as Sanders and Yang, were predicted to do better in the polls by respondents in the social condition.

While this research design achieves high ecological validity as a digital field experiment of a real-world event, it is not without limitations. As previously noted, one limitation is the ability to accurately assess compliance. Research has shown that survey respondents often report watching presidential debates more often than what administrative would suggest (Prior, 2012). Though we attempted to minimize incentives to provide misleading compliance information, it is possible that not all respondents in the analysis sample actually watched the debate. It is possible that overstating compliance could actually understate the effects of streaming commentary on our outcomes. Future research designs that have the ability to more actively monitor compliance could offer alternative methods for identifying complier average causal effects.

In addition, while we are confident that streaming chat will continue to be in use in political broadcasts in the future, our findings may vary depending on the similarity of the context and nature of the streaming chats. For example, it may be possible for social

chats to have fewer deleterious effects if moderators could control the quantity of user posts or moderate the content. Contexts in which social chats have more highly curated streams may lead to different hypotheses. While this may serve as a scope condition on the results of the study, we believe our theoretical pathways of frequency, content, and context effects offer a road map for future researchers to develop hypotheses considering how their settings differ on these dimensions. Even within the current study, we suggest how the unique features of the social streaming chat may lead to different consequences from the expert chat.

Overall, our study points to several consequences of streaming chat for political attitudes by shaping a more negative and less engaging experience that has the potential to spill over into how viewers evaluate candidates and their viability in elections. These findings suggest that traditional theories of media effects must account for the potential role of streaming chat, as it increasingly pervades live political broadcasts.

References

- Anspach, Nicolas M and Taylor N Carlson. 2018. “What to believe? Social media commentary and belief in misinformation.” *Political Behavior* pp. 1–22.
- Arceneaux, Kevin and Martin Johnson. 2013. *Changing minds or changing channels?: Partisan news in an age of choice*. University of Chicago Press.
- Barnidge, Matthew, Homero Gil de Zúñiga and Trevor Diehl. 2017. “Second screening and political persuasion on social media.” *Journal of Broadcasting & Electronic Media* 61(2):309–331.
- Bennett, W Lance and Shanto Iyengar. 2008. “A new era of minimal effects? The changing foundations of political communication.” *Journal of communication* 58(4):707–731.
- Cacciatore, Michael A, Dietram A Scheufele and Shanto Iyengar. 2016. “The end of framing as we know it and the future of media effects.” *Mass Communication and Society* 19(1):7–23.
- Cheng, Justin, Michael Bernstein, Cristian Danescu-Niculescu-Mizil and Jure Leskovec. 2017. “Anyone Can Become a Troll: Causes of Trolling Behavior in Online Discussions.”.
- Dayan, Daniel and Elihu Katz. 1992. *Media events*. harvard university press.
- Duggan, M and A Smith. 2016. “The political environment on social media.” *Pew Research Center* 25.
- Enke, Benjamin and Florian Zimmermann. 2017. “Correlation neglect in belief formation.” *The Review of Economic Studies* 86(1):313–332.
- Gerber, Alan S and Donald P Green. 2012. *Field experiments: Design, analysis, and interpretation*. WW Norton.
- Gil de Zúñiga, Homero, Alberto Ardèvol-Abreu and Andreu Casero-Ripollés. 2019. “WhatsApp political discussion, conventional participation and activism: exploring direct, indirect and generational effects.” *Information, Communication & Society* pp. 1–18.

- Gil de Zúñiga, Homero, Matthew Barnidge and Trevor Diehl. 2018. "Political persuasion on social media: A moderated moderation model of political discussion disagreement and civil reasoning." *The Information Society* 34(5):302–315.
- Gil de Zúñiga, Homero, Victor Garcia-Perdomo and Shannon C McGregor. 2015. "What is second screening? Exploring motivations of second screen use and its effect on online political participation." *Journal of Communication* 65(5):793–815.
- Gottfried, Jeffrey A, Bruce W Hardy, R Lance Holbert, Kenneth M Winneg and Kathleen Hall Jamieson. 2017. "The changing nature of political debate consumption: Social media, multitasking, and knowledge acquisition." *Political Communication* 34(2):172–199.
- Gross, Kimberly, Ethan Porter and Thomas J Wood. 2019. "Identifying media effects through low-cost, multiwave field experiments." *Political Communication* 36(2):272–287.
- Iyengar, Shanto. 1987. "Television news and citizens' explanations of national affairs." *American Political Science Review* 81(03):815–831.
- Iyengar, Shanto and Donald R Kinder. 1987. "News that matters: Agenda-setting and priming in a television age." *News that Matters: Agenda-Setting and Priming in a Television Age* .
- Jungherr, Andreas. 2014. "The logic of political coverage on Twitter: Temporal dynamics and content." *Journal of communication* 64(2):239–259.
- Larsson, Anders Olof and Hallvard Moe. 2012. "Studying political microblogging: Twitter users in the 2010 Swedish election campaign." *New media & society* 14(5):729–747.
- Lerman, Kristina, Xiaoran Yan and Xin-Zeng Wu. 2016. "The" majority illusion" in social networks." *PloS one* 11(2):e0147617.
- Levy, Gilat and Ronny Razin. 2015. "Correlation neglect, voting behavior, and information aggregation." *American Economic Review* 105(4):1634–45.
- McGregor, Shannon C and Rachel R Mourão. 2017. "Second screening Donald Trump: Conditional indirect effects on political participation." *Journal of broadcasting & electronic media* 61(2):264–290.

- Mungeam, Frank and Heather Crandall. 2011. "Commenting on the news: How the degree of anonymity affects flaming online."
- Munger, Kevin. 2017. "Dont@ Me: Experimentally Reducing Partisan Incivility on Twitter."
- Munger, Kevin. 2019. "Temporal Validity in Online Social Science."
- Ortoleva, Pietro and Erik Snowberg. 2015. "Overconfidence in political behavior." *American Economic Review* 105(2):504–35.
- Prior, Markus. 2012. "Who watches presidential debates? Measurement problems in campaign effects research." *Public Opinion Quarterly* 76(2):350–363.
- Santana, Arthur D. 2014. "Virtuous or vitriolic: The effect of anonymity on civility in online newspaper reader comment boards." *Journalism practice* 8(1):18–33.
- Suhay, Elizabeth, Emily Bello-Pardo and Brianna Maurer. 2018. "The polarizing effects of online partisan criticism: Evidence from two experiments." *The International Journal of Press/Politics* 23(1):95–115.
- Theocharis, Yannis, Pablo Barberá, Zoltán Fazekas, Sebastian Adrian Popa and Olivier Parnet. 2016. "A bad workman blames his tweets: the consequences of citizens' uncivil Twitter use when interacting with party candidates." *Journal of communication* 66(6):1007–1031.
- Vaccari, Cristian, Andrew Chadwick and Ben O'Loughlin. 2015. "Dual screening the political: Media events, social media, and citizen engagement." *Journal of Communication* 65(6):1041–1061.
- Van Cauwenberge, Anna, Gabi Schaap and Rob Van Roy. 2014. "TV no longer commands our full attention: Effects of second-screen viewing and task relevance on cognitive load and learning from news." *Computers in Human Behavior* 38:100–109.
- Van Cauwenberge, Anna, Leen dHaenens and Hans Beentjes. 2015. "How to take advantage of tablet computers: Effects of news structure on recall and comprehension."
- Wagner, María Celeste and Pablo J Boczkowski. 2019. "Angry, frustrated, and overwhelmed: The emotional experience of consuming news about President Trump." *Journalism* p. 1464884919878545.

Wagner, Markus and Davide Morisi. 2019. Anxiety, Fear, and Political Decision Making.
In *Oxford Research Encyclopedia of Communication*.

Zaller, John. 1992. *The nature and origins of mass opinion*. Cambridge university press.

The Effect of Streaming Chat on Perceptions of Debates

Supporting Information Files (SIF)

Supplemental Analyses

Experimental Conditions and Compliance

Below we show screenshots from each of the experimental conditions and describe respondent compliance.

We assessed whether respondents watched the debate through their self-reported answers to an initial question in Wave 2 and response options in the post-debate questions. Importantly, respondents were told that their answer to this question would not affect their compensation, reducing potential biases in these answers related to pressures from social desirability or experimenter demand effects. The analysis sample of Democrats who watched at least part of the debate is balanced between the control (204 respondents where 91% of wave 2 Democrats indicated watching) and social conditions (198 respondents where 90% of wave 2 Democrats reported watching). However, there are slightly fewer respondents who followed through in watching the debate in the expert condition (174 respondents where 86% of respondents indicated watching at least part of the debate). We recognize that the effects of the expert condition may be biased due to the slightly lower rates at which respondents followed through in watching the debate. Our primary analyses focus on the difference between the social and control conditions, for which there is balance across conditions.

We believe the lower propensity for respondents to watch the debate and watch the debate on the FiveThirtyEight website, when assigned, may be due to the quality of the viewing platform for the debate. The video box for watching the debate on the FiveThirtyEight site was significantly smaller in size compared to the video boxes on the ABC News and Facebook websites. Instead, the comments and analysis comprised a wider portion of the web browser screen. In addition, the FiveThirtyEight website launched the livestream just minutes prior to the debate, while the other sites had activated their livestreams earlier in the evening. We received multiple emails and open-ended feedback on the survey from respondents in the Expert condition reporting difficulties finding the stream or watching it, including one comment that compared the video to the "size of a sidebar ad" and another that said the video screen was "very small and made viewing the debate incredibly uncomfortable." We consider this to be part of the real-world viewing experience on different online platforms: The quality of the experience and ability to watch the debate and commentary on a platform represent an essential *part* of the effect of making these livestreams available to debate viewers

on perceptions of the debate.

Respondents in the Control condition were not explicitly told *not* to view comments on social media or other websites during the debate. We made this choice because restricting the way someone views the debate in this manner would not reflect a real-world viewing experience. Likewise, in the Social and Expert conditions, respondents were told that there would be comments alongside the video, but they were not explicitly told to focus on those comments exclusively. Similar to the real-world viewing experience, it is likely that participants varied in their attentiveness to the comments, and some subjects supplemented their debate experience by also viewing real-time comments on other sites, in addition to those on their assigned platform.²⁰

²⁰At the end of the Wave 2 survey, we ask respondents follow-up questions about how they viewed the debate. In the Control condition, 24 respondents still reported watching Facebook comments embedded next to the debate livestream, while 17 respondents reported viewing comments on FiveThirtyEight. In the Expert condition, 25 respondents reported viewing comments embedded next to the Facebook livestream. In the Social condition, 12 respondents reported viewing comments on FiveThirtyEight. For our analyses, we have to assume that we do not have any "defiers" in our study (Gerber and Green, 2012)– subjects who become less likely to take up the treatment on account of being assigned to the treatment condition.

Figure 10: Screenshots of three platforms: FiveThirtyEight, ABC News Facebook page, and ABC News

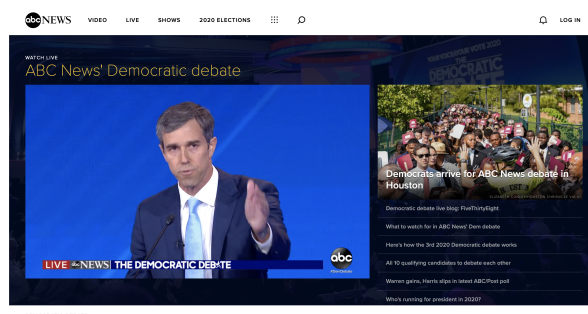


Table 3: Demographics of Wave 1 and Wave 2 Samples

Variable	Full W1 Sample	Recontacted Sample	Appears in W2	W2 Watched	W2 Watched on Assigned
Age	37.947	36.803	37.104	37.335	37.563
Female	0.511	0.488	0.476	0.47	0.472
White	0.732	0.698	0.692	0.705	0.713
College Degree	0.562	0.59	0.594	0.602	0.593
Democrat/Leaner	0.632	0.703	0.714	0.712	0.742
Conservatism	0.434	0.404	0.397	0.397	0.37
News Interest	0.678	0.764	0.768	0.774	0.78
News Frequency	0.787	0.845	0.849	0.854	0.863
Read Comments	0.538	0.603	0.599	0.602	0.597
Intention to Watch Debate	0.572	0.803	0.809	0.812	0.818
11 N	2352	1095	908	844	659

Note: “W2 Watched” refers to respondents that indicated they watched at least part of the debate, while “W2 Watched on Assigned” refers to respondents who watched at least part of the debate and self-reported watching the debate online, on the assigned platform.

Supplemental Results

In these analyses we add Wave 1 candidate familiarity as a control variable in the regressions. Conditional on pre-treatment familiarity with the candidates, Figure 11 displays the results of assignment to the Social or Expert conditions.

Overall, we see little movement in either the Social or Expert conditions relative to respondents assigned to watch the debate on the ABC News Website. We suspect that this may be because the focus of the comments in the chat streams often tracked with the amount of speaking time each candidate had and pre-existing poll performance. In the Expert chat, Biden, Sanders, and Warren had the most comments. In the Social chat, Biden and Sanders similarly were mentioned the most of the debate participants. In this way, at least in terms of candidate mentions, the content of the chat streams did not significantly vary from the debate, itself.

Figure 11: Candidate Familiarity

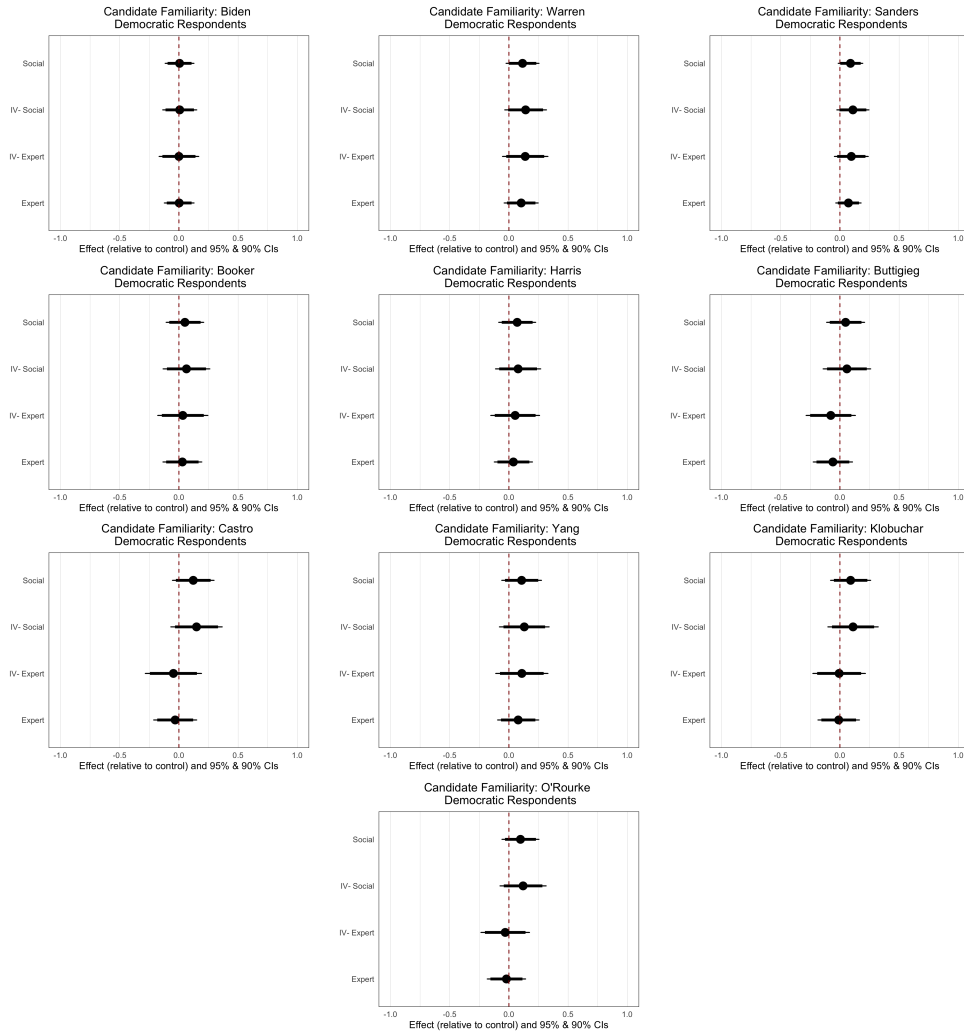


Figure 12: Candidate Feeling Thermometer Averages by Condition

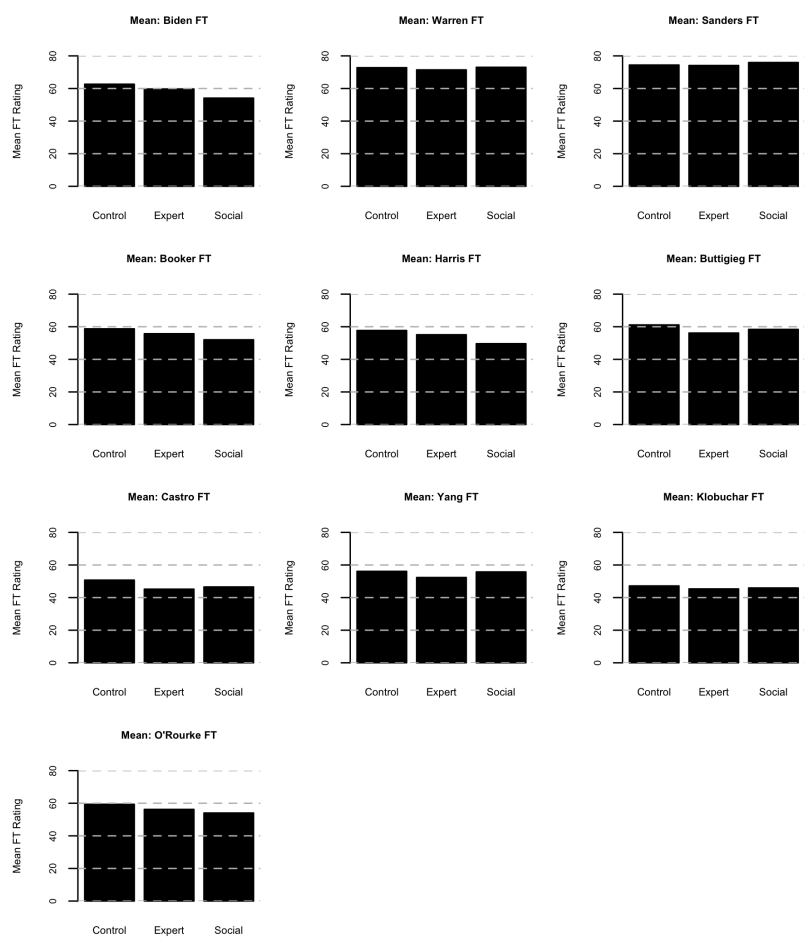


Figure 13: Candidate Poll Performance

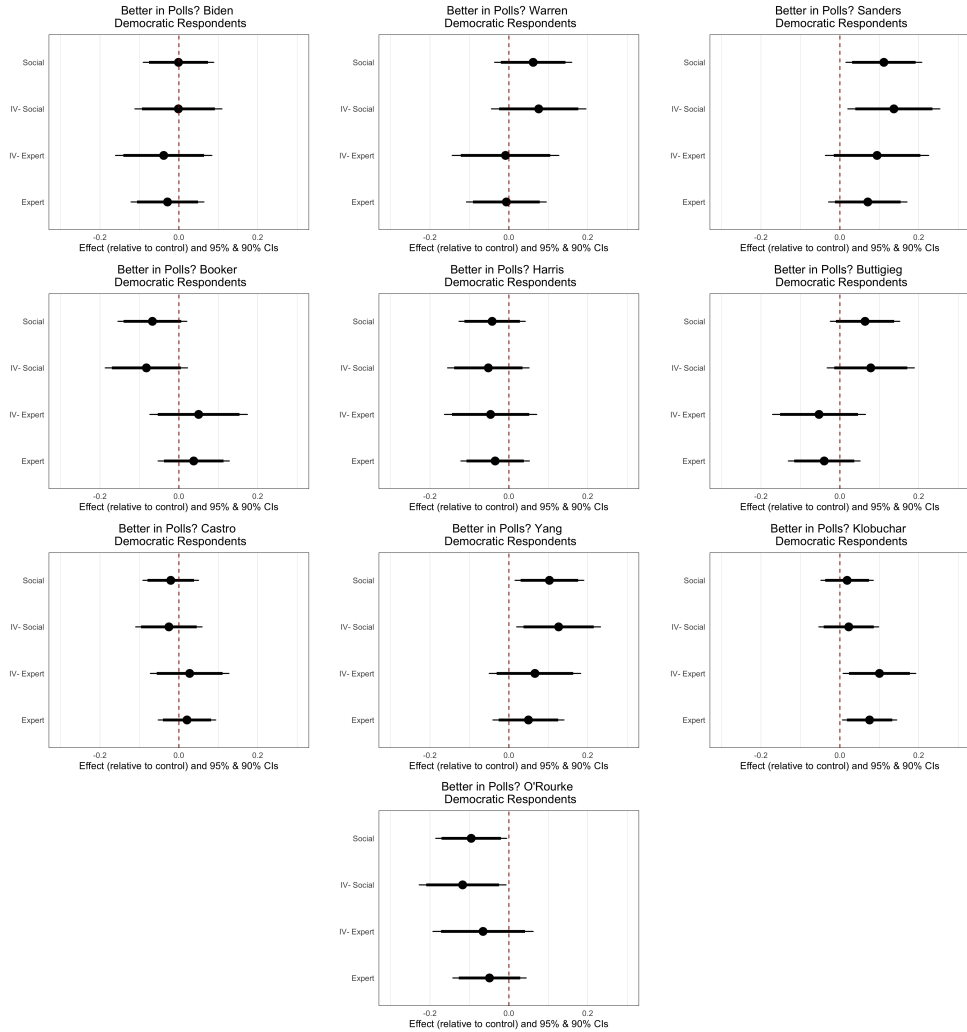


Figure 14: Party Feeling Thermometers

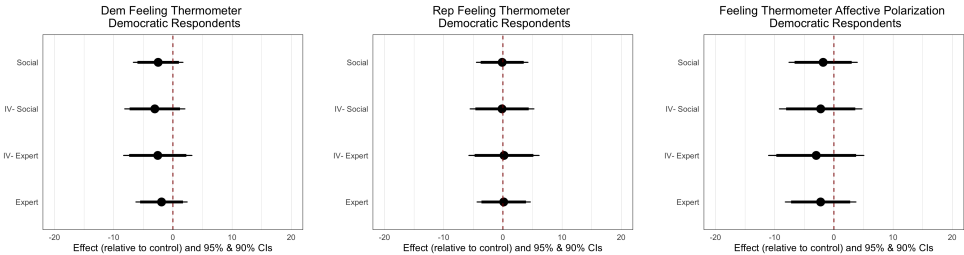
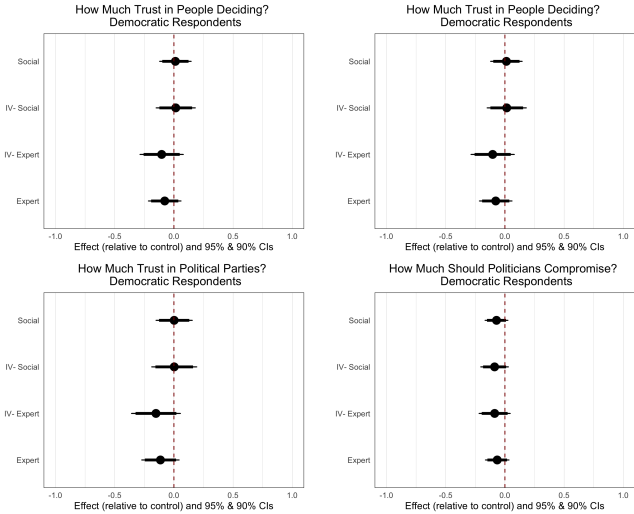
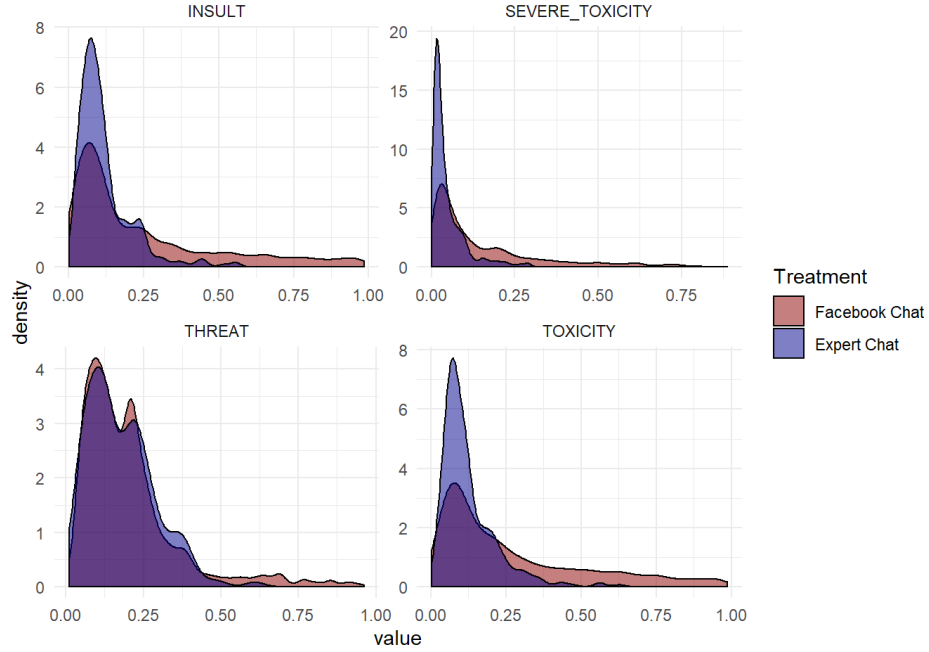


Figure 15: Trust Measures



Text Analysis: Toxicity

Figure 16: Distribution of Toxicity Scores of Social Comments



Text Analysis: Topics

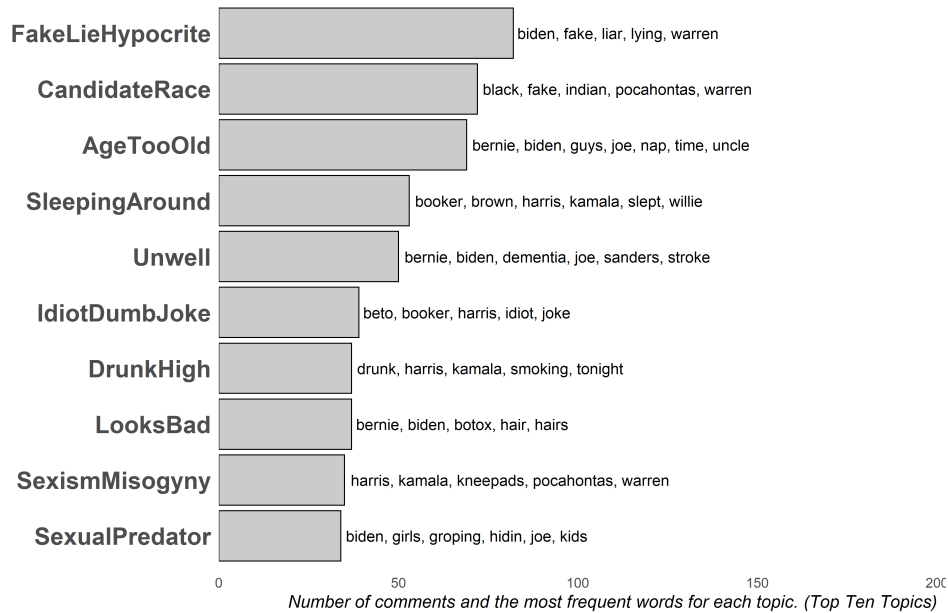
In this section, we present the main topics per candidate, and the most frequent words in each topic calculated only for comments related to the candidate. As we discussed in the main text, the main topics for most candidates reflect negative primes related to personal traits - age, race or gender - or a political features - gun control - of the candidates.

Table 4: Main Topic per Candidate

Candidate	Topic	Most Frequent Words (Within Candidate)
Biden	Age Too Old	alzheimers, dementia, guys, left, nap, past, player, record, senile, short, sleepy, time, uncle
	Sexual Predator	girls, groping, hidin, kids, likes, uncle
	Unwell	alzheimers, blood, dementia, memory, senile
Sanders	Age Too Old	agenda, agree, angry, arguing, asleep, babbling, bag, benie, citizen, coffin, communist, cross, dang, depends, dinosaur
	Unwell	call, die, gonna, hes, stroke, voice
	Socialism	commie, communist, dictator, homes, money, socialism, socialist, socialist
Harris	Sleeping Around	brown, eewwwwww, hoe, job, kneepads, knees, lady, legs, mouth, screwed, sleep, slept, top
	Drunk High	drunk, ha, smoking, sounds, tonight, weed
	Sexism Misogyny	eewwwwww, hoe, job, kneepads, knees, lady, legs, mouth, wheres, willy
Yang	Doctor Joke	america, asians, dentist, doctors, dont, drs, fool
	UBI Freedom Fund	dollars, free, lol, money, month, thousand
	Racism Toward Candidate	americans, asia, asian, driver, drs, eat, hands, healthcare
Warren	Candidate Race	american, black, fake, indian, native, pocahontas, pocohantas, pocohontas, slacks, speaks, woman
	Fake Lie Hypocrite	fake, indian, liar, lying, teacher
	HealthCare	dont, health, healthcare, insurance, obama, pay, private
O'Rourke	Guns	ar, ar15, automatic, children, el, gun, hell, lol, paso,trump
	Fake Lie Hypocrite	beato, dude, exact, fake, joke, latino, liar, mexican, mr.fake, office
	Candidate Race	beato, changed, dude, fake, hes, hispanic, irish, joke, latino, mexican
Booker	Fake Lie Hypocrite	black, community, cory, fake, guy, stop, trump
	Newark	alcohol, america, b9, bite, cant, cory, crisis, decent, district, drinking, drunks, fe0f, fix, ghetto
	Idiot Dumb Joke	asf, badge, biggest, black, bokker, corey, cory, declare, dumb, easy
Buttigieg	Smart Has Ideas	agree, air, answers, breath, brilliant,fluently, fresh, hear, insightful, intelligence
	Sexuality Homophobia	bet, bootygay, bootygig, boy, buddybutt, buttybutt, dumbford, gay, guy, homo
	War Military	afganistan, cuck, disappointment, material, military, president, served, service, talk
Castro	Attacking Biden	biden's, ageism, attacking, beating, blow, blows, called, cant, chooses, continues, cost
	Candidate Race	boy, gabriel, hispanic, juan, sissy
	Education	american, housing, reappartheid, schools, segregation
Klobuchar	Boring	makes, sleep, sleepy
	CriminalJustice	crimminals, free, kobachar, people, prison, rest, send
	Idiot Dumb	blowachar, idiot

We also present here the overall distribution of the topics detected in our hand-coded classification. Figure 17 plots the five most prevalent topics among the comments directed to at least one of the candidates. We also present the most frequent words in each topic. Different from the previous table, we present here the most frequent words polled across the comments for all the candidates, including their names.

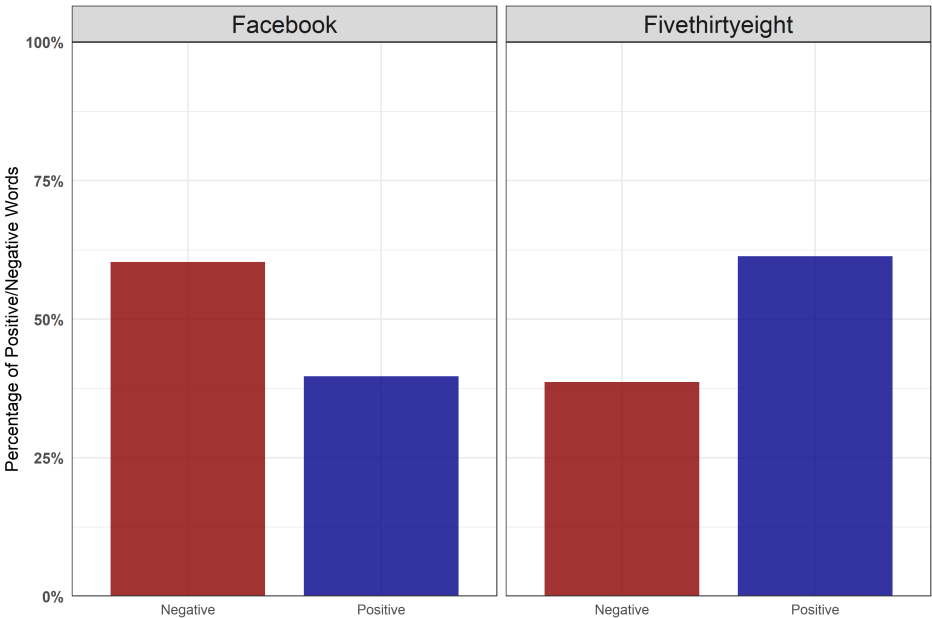
Figure 17: Topic Distributions and Key Terms



Dictionary Methods for Sentiment Analysis

In this section, we provide results using a simple classification, using dictionary methods, to measure polarity on the expert and social chat streams. Figure 18 shows the share of positive/negative words in each treatment condition. Eliminating words with no polarity, we find that more than half of the classified words on the social media chatbox have a negative polarity. In contrast, only about one-third of terms in the expert chat were classified as having negative sentiment. This basic method show us how the dialogue on social media is mostly marked by aggressive and negative language by the users.

Figure 18: Sentiment Analysis (Dictionary Methods)

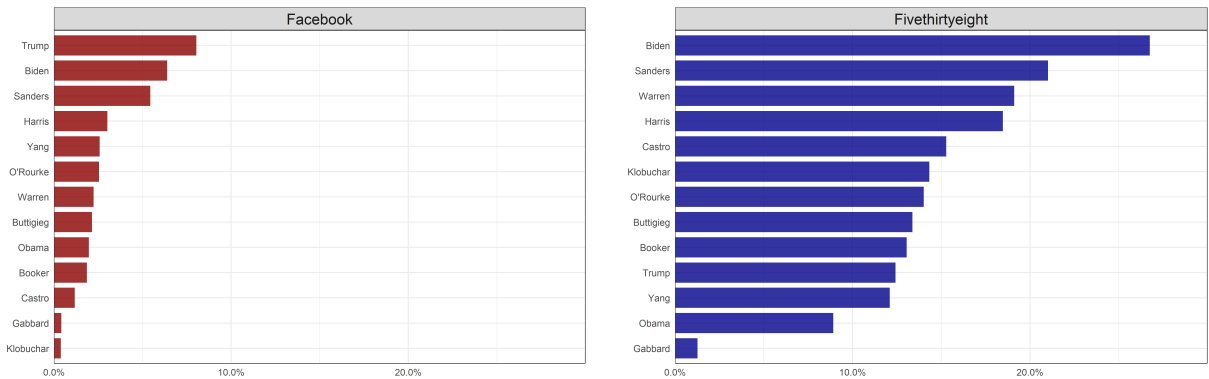


Dictionary Methods for Candidate Mentions

In addition to hand-coding, we use a simple name detection to identify the prevalence of candidates and politicians mentioned on each condition. This information is essential for understanding what viewers may come to think is important during the debate and influence how they evaluate the candidates post-debate, and also allows for some robustness of our results from the main paper. The proportion of comments mentioning a candidate as well as President Donald Trump and former-President Barack Obama are provided in Figure 19.

In the expert chat, the top-polling candidates at the time of the debate, Joe Biden, Bernie Sanders and Elizabeth Warren, were mentioned the most. This observation is consistent with the fact that these three candidates were among the top speaking candidates of the evening. In the social condition, Donald Trump is the most mentioned political figure. This observation is consistent with the fact that President Trump was mentioned— 50 times by name and 10 times referred to as "this president"— during the debate broadcast more than any of the Democratic candidates on the debate stage. Our qualitative analysis of the comments indicates that mentions to Trump were mostly made by republican aligned users who were, or attacking the Democratic candidates in the debate, or just using the streaming chat to show support for Trump. This finding corroborates with our overall picture about the Facebook streaming chat as an environment dominated by high polarized and toxic dialogue.

Figure 19: Share of Comments mentioning the Candidates, Trump, and Obama



Supplemental Results for Republicans

Below we repeat the analyses in the main text for Republican respondents, including leaners.

Figure 20: Republicans: Debate Experience

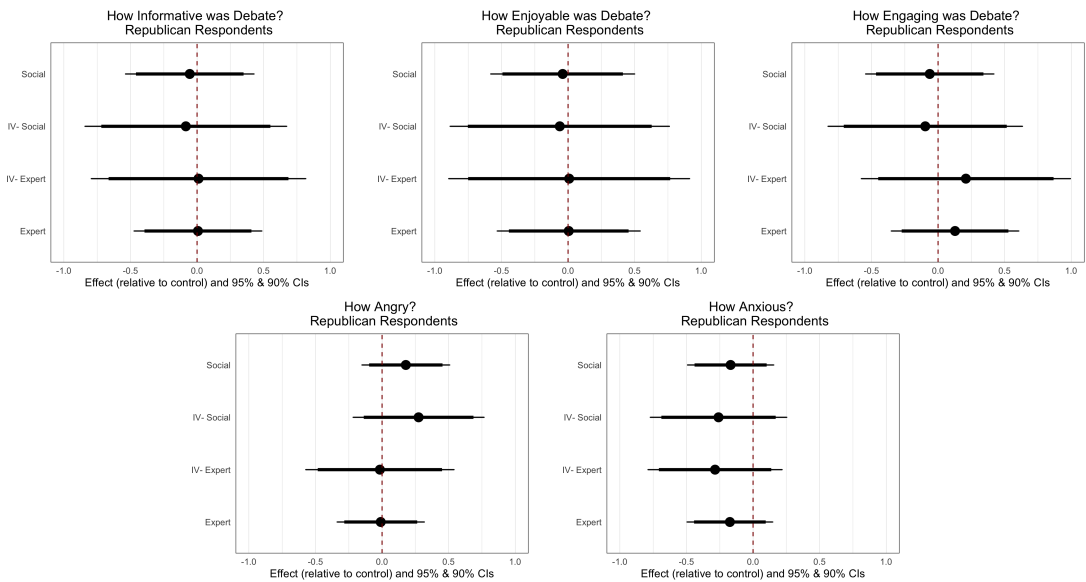


Figure 21: Republicans: Candidate Feeling Thermometers

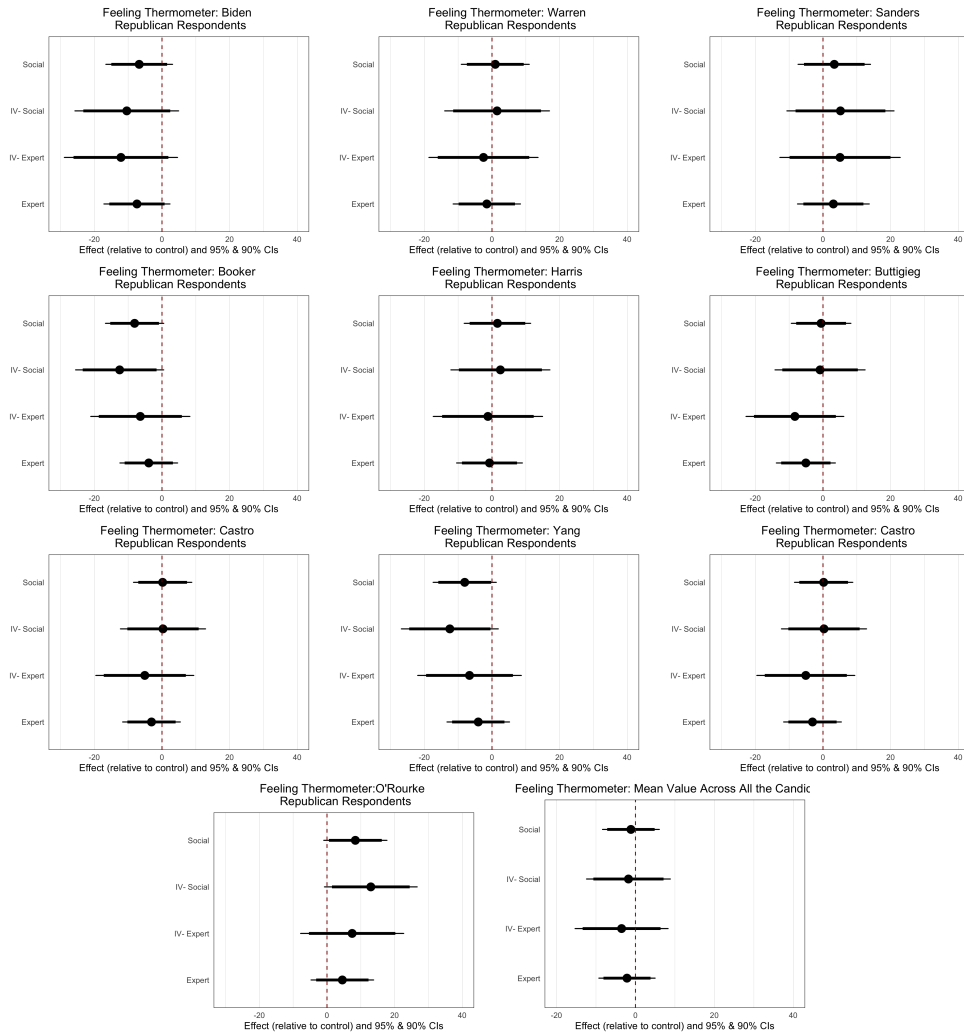
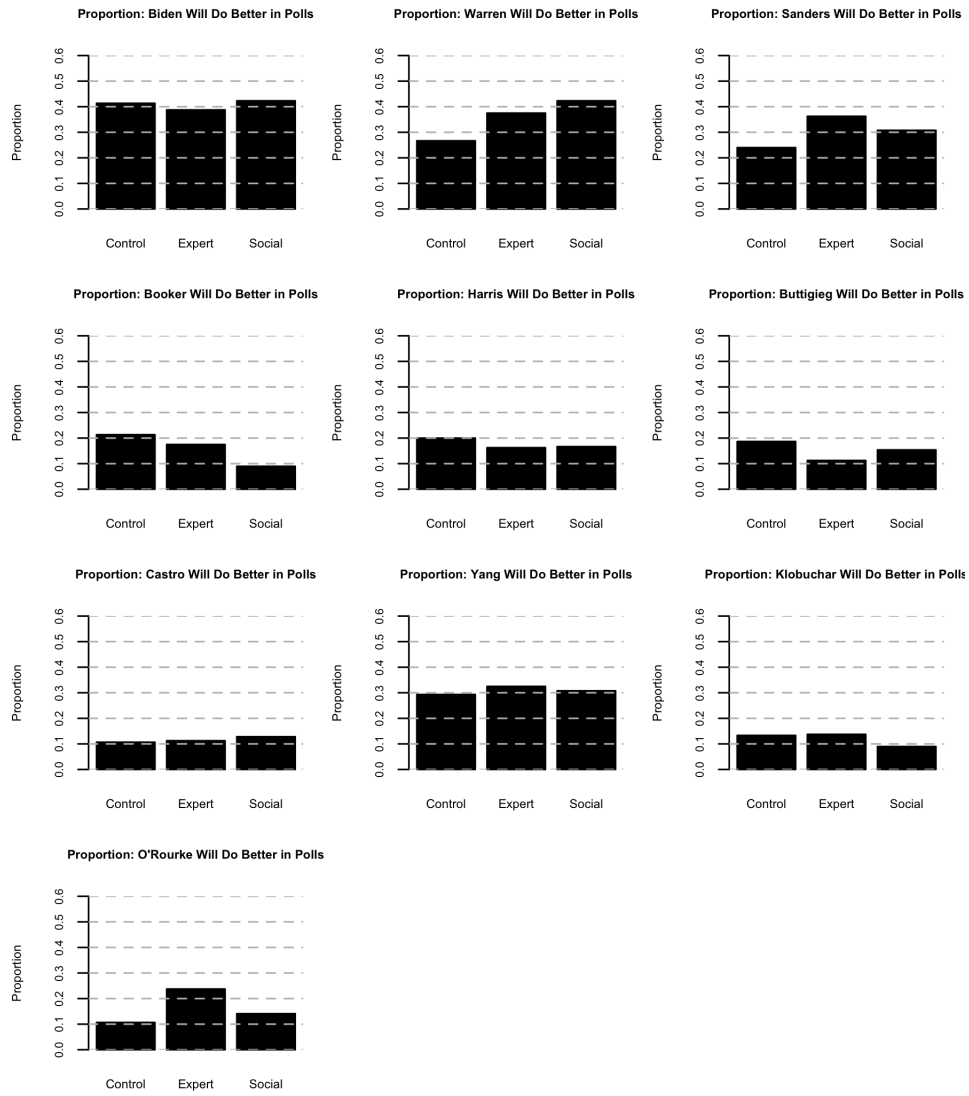


Figure 22: Republicans: Candidate Poll Performance Averages



Survey Question Wording

Wave 1

Demographic Questions *dem_birth* What is your year of birth?

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

- Less than high school degree (1)
- High school graduate (high school diploma or equivalent including GED) (2)
- Some college but no degree (3)
- Associate degree in college (2-year) (4)
- Bachelor's degree in college (4-year) (5)
- Master's degree (6)
- Doctoral degree (7)
- Professional degree (JD, MD) (8)

dem_race Choose one or more races that you consider yourself to be:

- White (1)
- Black or African American (2)
- Latino/Hispanic (3)
- American Indian or Alaska Native (4)
- Asian (5)
- Native Hawaiian or Pacific Islander (6)
- Other (7)

dem_sex What is your sex?

- Male (1)

- Female (2)

Q74 Do you have a Facebook account?

- Yes (1)
- No (2)

dem_income Information about income is very important to understand. Would you please give your best guess? Please indicate the answer that includes your entire household income in (previous year) before taxes.

- Less than \$10,000 (1)
- \$10,000 to \$19,999 (2)
- \$20,000 to \$29,999 (3)
- \$30,000 to \$39,999 (4)
- \$40,000 to \$49,999 (5)
- \$50,000 to \$59,999 (6)
- \$60,000 to \$69,999 (7)
- \$70,000 to \$79,999 (8)
- \$80,000 to \$89,999 (9)
- \$90,000 to \$99,999 (10)
- \$100,000 to \$149,999 (11)
- \$150,000 or more (12)

dem_zip What is your ZIP code?

dem_party Generally speaking, do you usually think of yourself as a Republican, a Democrat, an Independent, or something else?

- Republican (1)
- Democrat (2)

- Independent (3)
- Other (4)
- No preference (5)

Display This Question: If Generally speaking, do you usually think of yourself as a Republican, a Democrat, an Independent,... = Republican

dem_party_str_r Would you call yourself a strong Republican or a not very strong Republican?

- Strong (1)
- Not very strong (2)

Display This Question: If Generally speaking, do you usually think of yourself as a Republican, a Democrat, an Independent,... = Democrat

dem_party_str_d Would you call yourself a strong Democrat or a not very strong Democrat?

- Strong (1)
- Not very strong (2)

Display This Question: If Generally speaking, do you usually think of yourself as a Republican, a Democrat, an Independent,... = Independent Or Generally speaking, do you usually think of yourself as a Republican, a Democrat, an Independent,... = Other Or Generally speaking, do you usually think of yourself as a Republican, a Democrat, an Independent,... = No preference

dem_party_leaning Do you think of yourself as closer to the Republican or Democratic party?

- Republican (1)
- Democratic (2)

dem_ideo Here is a 7-point scale on which the political views that people might hold are arranged from extremely liberal (left) to extremely conservative (right). Where would you place yourself on this scale?

extremely liberal (1 (1) - 7 (7)) extremely conservative

Debate Intention and Viewing Capacity Questions

debate_intention Some of the Democratic candidates for president will be having a debate tonight on Thursday, September 12th from 8:00-11:00 PM ET hosted by ABC.

Do you intend to watch this Democratic presidential primary debate?

- Definitely yes (1)
- Probably yes (2)
- Probably not (3)
- Definitely not (4)

Display This Question: If Some of the Democratic candidates for president will be having a debate tonight on Thursday, Sept... = Definitely yes Or Some of the Democratic candidates for president will be having a debate tonight on Thursday, Sept... = Probably yes

debate_viewing How do you intend to watch the debate?

- On a television (1)
- On a mobile device (2)
- On a computer (3)

Display This Question: If Some of the Democratic candidates for president will be having a debate tonight on Thursday, Sept... = Definitely yes Or Some of the Democratic candidates for president will be having a debate tonight on Thursday, Sept... = Probably yes

debate_viewing_other Do you intend to watch the debate by yourself or with others?

- by yourself (1)
- with others (2)

Display This Question: If Some of the Democratic candidates for president will be having a debate tonight on Thursday, Sept... = Definitely yes Or Some of the Democratic candidates for president will be having a debate tonight on Thursday, Sept... = Probably yes

debate_viewing_comp Regardless of how you intend to watch the debate, is it possible for you to watch the debate on a computer?

- Yes (1)
- No (2)

News Consumption Questions

news_interest How interested, if at all, would you say you are in the news– that is national, international, regional/local news and other topical events accessed via any platform (radio, TV, newspaper or online)?

- Extremely interested (1)
- Very interested (2)
- Somewhat interested (3)
- Slightly interested (4)
- Not interested at all (5)

news_freq Typically, how often do you access news?

- More than once a day (1)
- Once a day (2)
- 4-6 times a week (3)
- 2-3 times a week (4)
- Once a week (5)
- Never (6)

news_source Please select the news sources that you access often: (Check all that apply)

- ABC News (1)
- MSNBC (2)
- Telemundo News (3)
- BBC News (4)

- National Review (5)
- Univision News (6)
- CBS News (7)
- NBC News (8)
- USA Today (9)
- CNN (10)
- New York Times (11)
- Wall Street Journal (12)
- Fox News (13)
- NPR (14)
- Washington Post (15)
- The Guardian (16)
- PBS News (17)
- Washington Times Weekly (18)
- Local television news (19)
- Regional or local newspaper (20)
- Blog sites (please specify) (21)
- YouTube channels (please specify) (22)
- Other (please specify) (23)

news_platform How do you normally get consume news? (Check all that apply)

- Television (1)
- Newspapers (printed) (2)
- Radio (3)

- Social media (4)
- Magazines (5)
- News websites (6)
- Other internet sources (including apps) (7)
- None of these (8)

Q71 How often do you read the comments accompanying news articles, YouTube videos, blog posts, or social media posts?

- Very often (1)
- Often (2)
- Occasionally (3)
- Never (4)

cand_fam How familiar are you with each of the following individuals?

- Joe Biden (1)
- Bernie Sanders (2)
- Elizabeth Warren (3)
- Kamala Harris (4)
- Pete Buttigieg (5)
- Beto O'Rourke (6)
- Cory Booker (7)
- Amy Klobuchar (8)
- Andrew Yang (9)
- Julian Castro (10)

Response code: Extremely familiar (1) Very familiar (2) Moderately familiar (3) Slightly familiar (4) Not familiar at all (5)

Message to respondents who are *ineligible* for Wave 2 Thank you!

Thank you for your participation! You will be paid the base amount listed in the HIT request. You are not eligible for the bonus task. Please click on the arrow below to receive your unique completion code required for payment.

Message to respondents who are *eligible* for Wave 2 Thank you!

Thank you for your participation! In a moment, you will be provided with a unique completion code required for payment.

wave2_interest Earn MORE as a BONUS! You have indicated that you intend to watch the Democratic Primary debate TONIGHT on Thursday September 12th, at 8:00-11:00PM ET) hosted by ABC. We like to invite you to take a second survey after the debate for which you can receive a bonus payment and entry into a \$100 raffle.

If you agree to participate, we will give you a link to view the debate on a specific online platform. Don't worry, ABC is hosting the debate, and the content of the video is the same across platforms. We're interested in understanding how people experience the debate based on the online platform they use to view it. Then, at the end of the debate, we'll send you a link to a follow-up survey (approx. 6-8 minutes) that must be completed by Friday, September 13, 2019 by 12PM ET. Compensation for this second round of participation will be \$1.50 for the 6-8 minute survey and an entry into a raffle for a prize of \$100. Please note that you will be paid for your second-round participation as a bonus and entered in the raffle as long as you complete the follow-up survey.

Are you interested in participating in our bonus study?

- Yes, I will watch the debate using the link you will provide and take the follow-up survey (1)
- No, I would not like to participate in this bonus study (2)

If eligible respondents selected yes for question "wave2_interest," respondent saw one of the following passages indicating to which condition they were assigned.

abcprompt You've been assigned to watch the debate on the ABC News livestream. Please go to this link on Thursday at 8PM ET to view the debate: <https://abcnews.go.com/>. You should see a live video of the debate on your screen. There will be a follow-up survey after the debate. You will be paid \$1.50 for completing this follow-up survey

as a bonus payment on MTurk and entered into a raffle for a chance to win \$100. Bonuses will be issued on Friday, September 13, 2019. We will send a reminder email that includes a link to the follow-up survey to your MTurk account before the debate. Please note, the follow-up survey link will not be live until after the debate has ended. You will have until 12PM ET on Friday, September 13, 2019 to complete the follow-up survey. (Please press the blue button below to receive your unique completion code.)

expprompt Youve been assigned to watch the debate on the FiveThirtyEight livestream. Please go to this link on Thursday at 8PM ET to view the debate: <https://fivethirtyeight.com/>. You should see live comments stream next to the video. There will be a follow-up survey after the debate. You will be paid \$1.50 for completing this follow-up survey as a bonus payment on MTurk and entered into a raffle for a chance to win \$100. Bonuses will be issued on Friday, September 13, 2019. We will send a reminder email that includes a link to the follow-up survey to your MTurk account before the debate. Please note, the follow-up survey link will not be live until after the debate has ended. You will have until 12PM ET on Friday, September 13, 2019 to complete the follow-up survey. (Please press the blue button below to receive your unique completion code.)

socprompt Youve been assigned to watch the debate on the Facebook Livestream. Please go to this link on Thursday at 8PM ET to view the debate: <https://www.facebook.com/ABCNews>. Please log-in to your Facebook account to view the stream and expand the video so that you see live comments stream on the right side of the screen. There will be a follow-up survey after the debate. You will be paid \$1.50 for completing this follow-up survey as a bonus payment on MTurk and entered into a raffle for a chance to win \$100. Bonuses will be issued on Friday, September 13, 2019. We will send a reminder email that includes a link to the follow-up survey to your MTurk account before the debate. Please note, the follow-up survey link will not be live until after the debate has ended. You will have until 12PM ET on Friday, September 13, 2019 to complete the follow-up survey. (Please press the blue button below to receive your unique completion code.)

Wave 2

compliance We understand sometimes things come up and your plans change. Did you watch Thursday's Democratic Presidential Primary Debate ("the debate") hosted by ABC? (Note: Your answer to this question will not affect your bonus compensation.)

- Yes, all of it (1)
- Yes, most of it (2)
- Yes, part of it (3)
- No, I did not (4)

post_reaction Tell us about your reactions to the debate.

- I found the debate informative (1)
- I enjoyed watching the debate (2)
- I found the debate engaging (4)

Response Code: Strongly Agree (1) - Strongly disagree (7) I did not watch the debate (8)

nextdebate Do you plan to watch the next primary debate, which will take place next month?

- Definitely yes (1)
- Probably yes (2)
- Probably not (3)
- Definitely not (4)

post_emotion Sometimes people experience different emotions during the debate. Please indicate if the debate made you feel . . .

- Angry (1)
- Hopeful (2)
- Excited (3)

- Anxious (4)
- Sad (5)

Response Code: Not at all (1), A little (2), Somewhat (3), Very (4) I did not watch the debate (5)

post_polls Which candidates do you think will do better in the election polls this week (after the debate) compared to last week? Joe Biden (1) Bernie Sanders (2) Elizabeth Warren (3) Kamala Harris (4) Pete Buttigieg (5) Beto O'Rourke (6) Cory Booker (7) Amy Klobuchar (8) Andrew Yang (9) Julian Castro (10) No one (11) I did not watch the debate (12)

cand_fam How familiar are you with each of the following individuals?

Joe Biden (1) Bernie Sanders (2) Elizabeth Warren (3) Kamala Harris (4) Pete Buttigieg (5) Beto O'Rourke (6) Cory Booker (7) Amy Klobuchar (8) Andrew Yang (9) Julian Castro (10)

Response Code: Extremely familiar (1) - Not familiar at all (5)

Q92 Survey Questions In this section, you'll be presented with several statements and questions about various issues. All statements and questions refer to the United States and populations within the United States. Please select the answer that comes closest to your views even if no choice is exactly right.

gt_parties How much do you feel that political parties help to make the government pay attention to what the people think?

- A great deal (1)
- A lot (2)
- A little (3)
- None at all (4)

gt_partiescomm How much do presidential primary debates help political parties communicate with voters?

- A great deal (1)
- A lot (2)
- A little (3)

- None at all (4)

gt_choice How much do you agree with the statement below?

”Political parties do not give voters real policy choices”

- Strongly agree (1)
- Somewhat agree (2)
- Somewhat disagree (3)
- Strongly disagree (4)

pol_decisions In general, how much trust and confidence do you have in the wisdom of the American people when it comes to making political decisions?

- A very great deal (1)
- A good deal (2)
- Not very much (3)
- None at all (4)

pol_officials Thinking about elected officials in Washington who share your positions on the most important issues facing the nation:

- They should work with elected officials they disagree with, even if it results in some policies you dont like. (1)
- They should stand up for their positions, even if that means little gets done in Washington. (2)

ft_dem Wed like you to rate how you feel towards the Democratic Party on a scale of 0 to 100. Zero means very unfavorable and 100 means very favorable. Fifty means you do not feel favorable or unfavorable. How would you rate your feeling toward the Democratic Party?

Very unfavorable (0 - 100) Very favorable

ft_rep Wed like you to rate how you feel towards the Republican Party on a scale of 0 to 100. Zero means very unfavorable and 100 means very favorable. Fifty means you do not feel favorable or unfavorable. How would you rate your feeling toward the Republican Party?

Very unfavorable (0 - 100) Very favorable

ft_cand We'd like you to rate how you feel towards each Democratic candidate at the debate on a scale of 0 to 100. Zero means very unfavorable and 100 means very favorable. Fifty means you do not feel favorable or unfavorable. How would you rate your feeling toward each candidate? Very unfavorable (0 - 100) Very favorable

- Joe Biden
- Bernie Sanders
- Elizabeth Warren
- Kamala Harris
- Pete Buttigieg
- Beto O'Rourke
- Cory Booker
- Amy Klobuchar
- Julian Castro
- Andrew Yang

Compliance Check 2

Display This Question: If We understand sometimes things come up and your plans change. Did you watch Thursday's Democratic... != No, I did not
compliance_online Did you watch the debate online?

- Yes (1)
- No (2)

compliance_platform Which online platform did you use? Check all that apply:

- Facebook (1)
- FiveThirtyEight website (2)
- Politico's website (3)

- New York Times' website (4)
- On Twitter (5)
- On the ABC News website (6)
- Other (7)

Display This Question: If Did you watch the debate online? != Yes And We understand sometimes things come up and your plans change. Did you watch Thursday's Democratic... != No, I did not

compliance_howview How did you watch the debate?

Display This Question: If We understand sometimes things come up and your plans change. Did you watch Thursday's Democratic... = No, I did not

compliance_noreason Please share with us why you did not watch the debate. Check all that apply, or add your own reason:

- I could not find where to watch the debate. (1)
- Something came up, and I was not able to watch. (2)
- I was not interested in watching it. (3)
- I did not like the candidates. (4)
- Other (5)

compliance_sm Did you look at any social media content or comments related to the debate?

- Yes (1)
- No (2)

Display This Question: If Did you look at any social media content or comments related to the debate? != No

compliance_smsource What type of social media content and comments related to the debate did you look at? Check all that apply:

- Facebook comments embedded next to the debate livestream (20)
- Facebook comments from your personal feed (21)

- Comments and analysis embedded on the FiveThirtyEight website (22)
- Comments and analysis embedded on a newspaper's website (27)
- Personal Twitter Feed (23)
- Twitter hashtags (24)
- Instagram Stories (25)
- Snapchat (26)

Display This Question: If What type of social media content and comments related to the debate did you look at? Check all t... = Twitter hashtags

compliance_hash Which Twitter hashtags did you follow?

compliance_assign How were you asked to view the debate during the initial survey?

On ABC News website (1) On the FiveThirtyEight website with the streaming comments feed (2) On the FiveThirtyEight website without the the streaming comments feed (3) On Facebook with the streaming comments feed (4) On Facebook without the streaming comments feed (5) Not sure (6)

Q75 You are almost done! In the next section, we'd like for you to tell us something from the debate that stood out to you. Its ok if you cant remember anything for a given candidate.

post_cand_biden What stood out to you about Joe Biden? *post_cand_sanders* What stood out to you about Bernie Sanders? *post_cand_warren* What stood out to you about Elizabeth Warren? *post_cand_harris* What stood out to you about Kamala Harris? *post_cand_buttigieg* What stood out to you about Pete Buttigieg? *post_cand_orourke* What stood out to you about Beto O'Rourke? *post_cand_booker* What stood out to you about Cory Booker? *post_cand_klobuchar* What stood out to you about Amy Klobuchar? *post_cand_yang* What stood out to you about Andrew Yang? *post_cand_castro* What stood out to you about Julian Castro?

Q111 Generally speaking, do you usually think of yourself as a Republican, a Democrat, an Independent, or something else? Republican (1) Democrat (2) Independent (3) Other (4) No preference (5)

Display This Question: If Generally speaking, do you usually think of yourself as a Republican, a Democrat, an Independent,... = Republican

Q113 Would you call yourself a strong Republican or a not very strong Republican? Strong (1) Not very strong (2)

Display This Question: If Generally speaking, do you usually think of yourself as a Republican, a Democrat, an Independent,... = Democrat

Q115 Would you call yourself a strong Democrat or a not very strong Democrat?
Strong (1) Not very strong (2)

Display This Question: If Generally speaking, do you usually think of yourself as a Republican, a Democrat, an Independent,... = Independent

Q117 Do you think of yourself as closer to the Republican or Democratic party?
Republican (1) Democratic (2)

comments Do you have any additional comments about your experience watching the debates?