Cross-Platform State Propaganda: Russian Trolls on Twitter and YouTube during the 2016 U.S. Presidential Election

Yevgeniy Golovchenko¹, Cody Buntain²,³, Gregory Eady¹, Megan A. Brown³, and Joshua A. Tucker³

Abstract
This paper investigates online propaganda strategies of the Internet Research Agency (IRA)—Russian “trolls”—during the 2016 U.S. presidential election. We assess claims that the IRA sought either to (1) support Donald Trump or (2) sow discord among the U.S. public by analyzing hyperlinks contained in 108,781 IRA tweets. Our results show that although IRA accounts promoted links to both sides of the ideological spectrum, “conservative” trolls were more active than “liberal” ones. The IRA also shared content across social media platforms, particularly YouTube—the second-most linked destination among IRA tweets. Although overall news content shared by trolls leaned moderate to conservative, we find troll accounts on both sides of the ideological spectrum, and these accounts maintain their political alignment. Links to YouTube videos were decidedly conservative, however. While mixed, this evidence is consistent with the IRA’s supporting the Republican campaign, but the IRA’s strategy was multifaceted, with an ideological division of labor among accounts. We contextualize these results as consistent with a pre-propaganda strategy. This work demonstrates the need to view political communication in the context of the broader media ecology, as governments exploit the interconnected information ecosystem to pursue covert propaganda strategies.

¹University of Copenhagen, Copenhagen, Denmark
²New Jersey Institute of Technology, Newark, NJ, USA
³New York University, New York, NY, USA

Corresponding Author:
Cody Buntain, New Jersey Institute of Technology, GITC 3420H, Newark, NJ 07102, USA.
Email: cbuntain@njit.edu
Keywords  
propaganda, Internet Research Agency, Twitter, YouTube, social media, foreign intervention

Introduction

Interest in foreign propaganda and disinformation is on the rise, the result of recent and increasing tensions between the West and Russia. Following the 2016 U.S. presidential election, the U.S. Department of Justice accused agents of Russian intelligence services of using fake social media accounts to influence the election in an effort to destabilize the country politically (Department of Justice 2018a; Department of National Intelligence 2017; Nakashima et al. 2017). Relatively little research, however, has sought to examine these recent technological efforts in the context of long-running propaganda strategies. As a consequence, questions remain about whether Russia’s use of social media for propaganda purposes has shown evidence of wholly new strategies, or whether these platforms primarily provide a means to extend existing ones. These questions are further motivated by a renewed interest among political scientists and media scholars in Russia’s use of information campaigns, like those perpetrated in the 2016 U.S. presidential election. Despite more than a quarter century having passed since the collapse of the Soviet Union, the nature of the Kremlin’s current propaganda strategies is less clear than it was during the Cold War.

Current debates concerning Russia’s role in the 2016 election offer two perspectives. On one hand, U.S. authorities have argued that the strategy pursued by the Kremlin, operating through the Russian Internet Research Agency (IRA), was designed to mobilize support for Donald Trump (who promised to improve relations with Russia) and to damage Hillary Clinton (who insisted on sanctioning Russia for its military involvement in Ukraine) (Council on Foreign Relations 2017; Department of Justice 2018a; Department of National Intelligence 2017). On the other hand, others have suggested that the Kremlin was not exclusively trying to support one candidate over another but rather to sow political discord by supporting liberals and conservatives simultaneously (Bloomberg 2018; Shane 2017; Steward et al. 2018). This work evaluates the evidence for these strategies by investigating hyperlinks to news and video content shared by IRA “trolls”—human-operated social media accounts with deceptive identities—on Twitter, one of the most politically relevant social media platforms during the 2016 U.S. presidential election.

While in-depth qualitative analyses of such cases are important for their ability to highlight the nuances of these disinformation campaigns, they are nonetheless limited in their ability to reveal the scope of patterns of disinformation. Anecdotal evidence of IRA trolls criticizing Trump in a few select instances, for example, may be dwarfed in aggregate by shared content that was largely pro-Republican. Our understanding of IRA strategies can therefore benefit from a larger-scale quantitative approach that investigates the extent to which the IRA supported one political ideology over another during the 2016 U.S. presidential campaign.
In this study, we therefore seek to adjudicate between the two perspectives above by analyzing the ideological slant of the behavior of 1,052 Kremlin-tied IRA “troll” accounts. These accounts have been linked to the IRA by Twitter and disclosed by the U.S. Congress as part of its investigation into Russian interference during the 2016 election. We study the ideology of the content shared by these accounts by focusing not on the text of the tweets, but rather on the hyperlinks that IRA trolls shared to other online content. In particular, we focus on links to political news media and to video content on YouTube. We apply both automated and manual coding procedures to these sources to label them by ideological orientation. This coding enables us to examine the extent to which the IRA shared consistent ideological content within and across online media platforms. Using the aforementioned arguments about the IRA’s goals as a point of departure, we investigate how the IRA employed its strategies, and place them in the context of prior propaganda campaigns.

Our examination begins by taking the two main strategies proposed for the Russian IRA’s goals—sowing discord and supporting a particular candidate—and formulating predictions about the IRA’s social media behavior that would be consistent with each strategy. We then examine the extent to which the evidence confirms these predictions within and across platforms by examining links shared by IRA accounts to political news stories and YouTube videos. While mainstream narratives about the IRA’s disinformation campaign present the two aforementioned goals as mutually exclusive, we find evidence consistent with both strategies. This mixed evidence suggests the IRA’s strategy during the 2016 election was both multifaceted and developed around multiple potential outcomes and uses. Despite this empirical support for both strategies and related limitations, when contextualized within the propaganda literature, we posit the evidence nevertheless suggests the IRA was primarily interested in supporting the election of the Republican candidate. Finally, we argue that the introduction of online social platforms has not resulted in a wholly new set of propaganda strategies. Instead, by facilitating anonymity and increased exposure at scale, these platforms have lowered the cost of deploying foreign propaganda strategies.

We contribute to research into news media and political processes in three ways. First, we present substantive findings concerning the Kremlin’s cross-platform propaganda strategies in the context of political news-sharing during the 2016 U.S. presidential election. Second, we take up recent calls to examine the understudied cross-platform aspect of online disinformation campaigns by investigating the links between content shared on Twitter and news media in the broader digital information ecosystem generally, and to YouTube specifically (Bode and Vraga 2018; Tucker et al. 2018). Lastly, we situate this study in the long-standing literature on propaganda and disinformation by demonstrating the historical continuity in the manipulative uses of new media (Bernays 1928; Ellul 1973; Jowett 1987; Lasswell 1938; Taylor 2003). We end by discussing how states can adapt old propaganda techniques to new media environments by exploiting information flows across platforms.
Literature and Hypotheses

Propaganda as a Disinformation Campaign

Researchers across the social sciences have examined various deceptive political uses of information by investigating, for example, fake news (Khaldarova and Pantti 2016; Tandoc et al. 2018) and disinformation (Bennett and Livingston 2018; Hjorth and Adler-Nissen 2019; Mejias and Vokuev 2017). A large portion of this research examines these phenomena domestically, often in the U.S. context (Allcott and Gentzkow 2017; Grinberg et al. 2019; Guess et al. 2019; Pennycook et al. 2018). Increasingly, however, researchers are turning their attention to state-driven campaigns that cross national borders. In these contexts, propaganda is often described as strategic communication (Hallahan et al. 2007; Zerfass et al. 2018) or information warfare (Taylor 2003; Thornton 2015).

A large part of this literature has focused on the Russian government’s strategic use of media in Ukraine and Russia, particularly its use of information warfare (Darczewska 2014; Hoskins and Shchelin 2018; Pomerantsev 2015; Tanchak 2017; van Niekerk 2015). Information warfare in this case refers to manipulative uses of information to achieve military and political goals (Thornton 2015). Lucas and Nimmo (2015) argue that the purpose of Russian disinformation and information warfare is to sow doubt and disagreement. Ramsay and Robertshaw (2018) show, for example, how Russian state-controlled outlets, such as Sputnik and RT (formerly Russia Today), pursue these goals when reaching out to English-speaking audiences. While the literature on information warfare and strategic communication offers important insights, we examine the IRA’s campaign using the concept of “propaganda” to bridge the gap between contemporary digital campaign practices and insights from classical propaganda studies that predate the digital era. In doing so, we place seemingly new techniques in their broader historical context.

Propaganda, according to Jowett and O’Donnell (2014), is “a deliberate, systematic attempt to shape perceptions, manipulate cognition, and direct behavior to achieve a response that furthers the desired intent of the propagandist” (p. 7). One can further distinguish broadly between at least two types of propaganda: “white propaganda,” which relies on truthful information from open sources, and “black propaganda,” which is intentionally misleading or untruthful communication whereby the true source of information is concealed. Black propaganda, in other words, is a form of disinformation that is both systematic and organized. It is also often called “covert propaganda,” to emphasize the fact that it “tends to hide its aims, identity, significance and source” (Ellul 1973: 15). The IRA’s use of trolls is one such example of black propaganda, because it sought to systematically mislead its audience by portraying its social media accounts as ordinary users or news providers while concealing the true source of each account. Nevertheless, the content shared by the IRA was itself not necessarily false or misleading. Yin et al. (2018) find, for instance, that while IRA Twitter accounts shared 50 percent more content from “junk news” (i.e., websites with
misleading news) than typical Twitter users, its overall use of these sources was low (only 6 percent of all links shared)."1

Jowett and O’Donnell (2014: 29–30) describe these disinformation techniques as part of a “Deflective Source Model”: Rather than communicating directly with a targeted audience, information is shared covertly through a source that has no apparent connection to the propagandist. In this way, the audience does not associate the message with the true source (i.e., the IRA or Russian government). The ultimate goal of this common propaganda technique is to increase the credibility of the content (Jowett and O’Donnell 2014).

Many examples of state-driven disinformation campaigns rely on this “Deflective Source Model.” Taylor (2003: 256), for instance, describes how the Soviet Union used the “World Peace Council,” a Western organization, as an “agent of influence” to legitimize North Korea’s invasion of South Korea while spreading disinformation concerning American use of germ warfare during the conflict. The recent use of seemingly “ordinary citizens” as deflective sources in disinformation campaigns, as with the IRA’s troll accounts, is also not new, nor is it limited to the Kremlin. Researchers within media studies and political communication, for example, commonly use the term “astroturfing” to describe the use of online fake accounts as a deflective source to mimic spontaneous grassroots activity (Howard and Kollanyi 2016; Peng et al. 2017; Ratkiewicz et al. 2011).

Finally, to maximize success, authoritarian regimes and liberal democracies have both historically relied on what the propaganda literature refers to as pre-propaganda: propaganda that is not directly related to the political message of the propagandist (Ellul 1973). The main purpose of this strategy is to cultivate legitimacy and make the audience mobilizable to subsequent direct political messaging. For example, a substantial proportion of radio broadcasts transmitted by U.S. authorities into the Soviet Union were of popular music programs or news that were not explicitly anti-Soviet (Uttaro 1982). This content was used to gain credibility and lure audiences toward what would be more explicit anti-Soviet propaganda (Jowett and O’Donnell 2014). Here, we explore these propaganda techniques and how the IRA’s behavior on social media may be interpreted through this lens.

**Cross-Platform Propaganda and Hyperlinks**

The IRA leveraged multiple online platforms in its information campaigns by interlinking content on Facebook, Twitter, YouTube, Instagram, and other platforms. In some cases, the IRA pretended to be the same people or activist groups across multiple platforms (Howard et al. 2018). This cross-platform approach is unsurprising considering the Internet’s interconnected nature but is also strategically important given that social media platforms exist in a broader media ecosystem of hyperlinks, shares, and likes (Bode and Vraga 2018; Gerlitz and Helmond 2013).

Nevertheless, research on political communication and disinformation often focuses on a single platform in isolation. This drawback has recently led to calls for an expansion of cross-platform research (Bode and Vraga 2018; Tucker et al. 2018). We
thus seek to contribute to this evolving research agenda by exploring how state-driven propaganda is connected and disseminated across platforms. To pursue this goal, we focus our analysis on hyperlinks, the unit of information exchange that binds (social) media platforms together.

**Russian Trolls and the U.S. Election**

Google, Twitter, and Facebook have testified in Congress that the IRA reached millions of users in the United States. In 2018, the U.S. Department of Justice charged 13 members of the IRA for being part of a Russian agency engaged in efforts to interfere in U.S. elections (Department of Justice 2018b).

In general, the existing literature within communication research and the computational social sciences has predominantly examined manipulative information campaigns by focusing on bots (Ferrara 2017; Ferrara et al. 2016; Shao et al. 2017; Stukal et al. 2017; Varol et al. 2017). Bots, however, are only one part of online propaganda campaigns. IRA accounts included both Twitter bots (i.e., automated profiles controlled by software) and more sophisticated accounts often called “trolls” or “sock puppets” (Badawy et al. 2018; Morgan and Shaffer 2017; Steward et al. 2018; Zannettou et al. 2018). Unlike bots, IRA troll accounts were those controlled directly by humans posing as seemingly ordinary U.S. social media users. Woolley and Howard (2016, 2017) refer broadly to the use of such strategies in a political context as “computational propaganda.”

The reach of these troll accounts is evidenced by Twitter’s internal data that suggest that at least 1.4-million users in the United States directly engaged with IRA-controlled accounts during the 2016 elections, either by quoting, mentioning, liking, retweeting, or replying to these accounts (Twitter 2018). While Hindman and Barash (2018) note that relatively few IRA accounts gained a large number of followers, even a small subset of trolls is capable of deceiving ordinary citizens and trained journalists given their highly complex behavior and relatively authentic appearance. Lukito et al. (2020) show, for instance, that the IRA accounts were quoted by at least 71 news media organizations, including the *Washington Post*, *The Guardian*, and *Huffington Post*. The extent of online propaganda campaigns by foreign actors can, in other words, be substantial.

It is useful to note that journalists have reported on the IRA and its tactics long before the 2016 presidential election. The Russian newspaper, *Novaya Gazeta*, for example, provided investigative coverage of the IRA as far back as 2013 (Garmazhapova 2013). Using leaked information and conversations with IRA recruitment personnel, the newspaper reported that the IRA was recruiting Russians to operate fake accounts on a regular basis. Employees would receive a regular salary for posting a specific amount of online content each day while supervised by other staff members. They were to engage with Russian-speaking audiences in order to weaken Russian opposition and to strengthen the government. According to Lyudmila Savchuk, who claims to have worked for the IRA, the agency had a specialized unit to engage with English-speaking audiences (Myers and Evstatieva 2018). One of their goals was to sow social
division in the United States. Another former employee claimed in an interview that the IRA’s “foreign desk” was not attempting to turn U.S. audiences toward Russia. Instead, its goal was to destabilize the United States by setting “Americans against their own government: to provoke unrest and discontent, and to lower Obama’s support ratings” (Meduza 2017). While these accounts of IRA goals are anecdotal, they suggest that the IRA was highly organized, with a division of labor among IRA trolls, and supervision from managers who sought to fulfill explicit political goals. One goal of the research presented herein is to examine whether these organizational objectives and divisions of labor manifest in the IRA’s behavior in social media.

The Kremlin’s Strategy during the 2016 Elections: Two Views

Journalists and scholars in security studies have argued that the Kremlin has sought to exploit the extremes of the ideological right and left to politically weaken western societies (Beitéane 2015; European Union Institute for Security Studies 2016; Michel and Goldenberg 2017). According to an official assessment by the U.S. Office of the Director of National Intelligence, the Kremlin’s strategy during the 2016 U.S. election was not only “to undermine public faith in the US democratic process,” but also to damage Hillary Clinton and support Donald Trump. The methods of investigation by the intelligence services, however, remain opaque, and it is unclear how information regarding IRA behavior and its goals were collected and analyzed. In line with its assessments, the U.S. Department of Justice indicted thirteen Russians and three companies for a covert campaign to damage the Democratic Party and support Trump’s election campaign (Department of Justice 2018a).

This same indictment, however, also points in a different direction. The Department of Justice accuses Russian entities of organizing protests against Trump, such as the “Charlotte Against Trump” protest on November 19, 2016, in North Carolina (Department of Justice 2018a: 23). The Department of Justice thus also accused the Russia of “spread[ing] distrust towards the candidates and the political system in general” (Department of Justice 2018a: 6). Furthermore, analysis by journalists of political Facebook advertisements bought by the IRA suggests that some of the ads also promoted Bernie Sanders together with campaigns against Trump (Shane 2017). These findings suggest the IRA’s propaganda strategy was not solely to strengthen the electoral prospects of Donald Trump.

In general, social media offers various opportunities for foreign actors to engage with legitimate users by reducing the cost of interaction and to sow political discord by supporting opposing parties and candidates. The literature in political communication offers examples showing, for instance, how political retweet networks resemble polarized “echo-chambers”—that might be targeted by unscrupulous foreign actors—where users primarily share content from like-minded individuals (Golovchenko et al. 2018; Steward et al. 2018). The evidence that social media polarizes the public by minimizing ideologically cross-cutting information is mixed (Tucker et al. 2018), but Barnidge (2017) finds that individuals perceive more political disagreement on social media than they do in their face-to-face interactions. Alternatively, Boxell et al. (2017)
find little correlation between social media usage and polarization. Recent field experimental work suggests, however, that exposure to social media and cross-cutting partisan information does cause polarization (Allcott et al. 2020; Bail et al. 2018). The literature increasingly suggests, moreover, that political elites and inter-party conflict may fuel political polarization (for a review, see Tucker et al. 2018). The U.S. context thus appears to offer fertile ground for foreign actors to increase polarization merely by amplifying existing antagonistic messages from opposing political parties.

Past research on IRA behavior has examined it both qualitatively and quantitatively. Zannettou et al. (2018), for example, analyzed twenty-seven thousand tweets from IRA trolls and found they had a minor effect on propagating hyperlinks to news on Twitter, Reddit, and 4chan, with the exception of links to Russia Today (Zannettou et al. 2018). They show that at least 10.3 percent of trolls explicitly portrayed themselves as Trump supporters by writing either “trump” or “maga” in their profile names or self-descriptions (Zannettou et al. 2018: 4). It is unclear, however, how many of the remaining accounts associated themselves with the political left. A recent study by Lukito (2020) suggests that the IRA’s activity was coordinated across multiple platforms, with Reddit potentially serving as a testing ground for content that would later be used on Twitter.

From a more qualitative perspective, Linvill and Warren (2018) suggests right-wing IRA trolls on Twitter produced more content than left-wing trolls. Of the 1,311 troll accounts they analyze, 617 were manually coded by the authors as right-wing and 230 as left-wing. The remaining 464 were classified into other categories such as “Fearmonger,” “Hashtag Gamer,” and “News Feed.” Although left-wing trolls were more active on average, the IRA generated approximately 64 percent more tweets using right-wing accounts than left-wing accounts in aggregate (716 compared to 437 tweets per day on average respectively). Linvill and Warren (2019) have since updated this qualitative assessment using Twitter’s official data release of 3,613 IRA troll accounts, finding 454 right- and 228 left-leaning trolls, with many more “Fearmonger” accounts (Linvill and Warren 2019). Neither Linvill and Warren (2018) nor Linvill and Warren (2019) specify what proportion of these tweets are related to politics, however. Where Zannettou et al. (2018), Linvill and Warren (2018), and Linvill and Warren (2019) focus on the individual characteristics of troll accounts, Steward et al. (2018) examine IRA trolls through a network perspective. They identify ninety-six trolls that posted tweets related to both Black Lives Matters and mass shootings. Steward et al. (2018) show, moreover, that IRA accounts successfully infiltrated both left-leaning and right-leaning parts of the retweet network, potentially amplifying preexisting polarization between the two groups. However, they also argue that the fake accounts were slightly more prevalent in the left-leaning cluster of retweets (Steward et al. 2018). Similarly, Bastos and Farkas (2019) find a group of IRA accounts that engaged with the Black Lives Matter topic on Twitter. They argue that the goal of these messages may have been to discourage African-Americans from voting for Clinton, or from voting altogether.

Similar to our research, Badawy et al. (2018) analyzed tweets from a small number of 221 Russian trolls disclosed by the U.S. Congress and the tweets from users who interacted with them online. Using an automated approach to estimate account ideology, they
find that 107 of the trolls were “liberal,” and 108 were “conservative.” This finding supports the hypothesis that the IRA sought to influence both sides of the ideological spectrum. The authors also show that conservative trolls were more active because they posted 844 original tweets, while liberal trolls posted only 44. This finding appears consistent with a strategy of supporting one preferred candidate over the other. Furthermore, Badawy et al. (2018) argue that the trolls had a “mostly conservative, pro-Trump agenda” based on text analysis of the words most frequently used by IRA troll accounts. They show that conservative users were 31 times more likely to retweet trolls than were liberal users. Nevertheless, these results do not show the extent that the trolls themselves amplified conservative content, nor do they reveal whether their strategic use of ideology changed over time. The results we present below thus differ from Badawy et al. (2018) to the extent that we analyze a much larger set of IRA troll accounts and seek to evaluate how their ideology changes over time. Our temporal analysis provides additional insight into whether the IRA sought to support the Republican Party by artificially increasing its apparent online support or by seeking to shift liberal user in a conservative direction. The former would be suggested by trolls having a consistent ideology over time, whereas the latter would be suggested by trolls’ ideologies beginning as primarily liberal and then becoming increasingly conservative.

More recently, Howard et al. (2018) analyzed a range of datasets across platforms of content posted by the IRA, arguing that it did target both liberals and conservatives on Facebook and Instagram. They argue that “the IRA sought to energize conservatives around Trump’s campaign and encourage the cynicism of other voters in an attempt to neutralize their vote” (Howard et al. 2018: 32). In other words, the IRA engaged with both liberals and conservatives, but the goal was the same: to support Donald Trump’s presidential campaign. In their analysis of the IRA’s presence on Twitter, Howard et al. (2018) find that the IRA targeted conservatives more than liberals in early 2015. This gap in targeting closed, however, later that year. Liberal and conservative activity by trolls were similar throughout the presidential campaign until early 2017, after which the authors observe a surge in activity targeting conservatives (Howard et al. 2018). Howard et al. (2018) argue that the IRA accounts on Twitter heavily relied on links to YouTube in their attempt to manipulate audiences in the United States (Howard et al. 2018). According to an analysis for the U.S. Senate Intelligence Committee, carried out by the research firm New Knowledge, 96 percent of content on 1,107 videos uploaded on the seventeen fake YouTube channels by the IRA were thematically related to Black Lives Matter and police brutality (DiResta et al. 2019). These data are not yet publicly available, nor do they capture the IRA’s use of YouTube content generated by ordinary users outside of the IRA. Indeed, little research exists, to the best of our knowledge, on how the IRA exploited the Twitter–YouTube link to support either the liberal or conservative side of the election.

Differences in the conclusions arrived at in the literature may result from variance among the research questions posed, as few researchers explicitly focus on the propaganda strategies underlying trolls’ behavior. Overall, current assessments of the Kremlin’s use of ideological polarization in their propaganda efforts during the U.S. presidential election lack (1) contextualization of this behavior with prior propaganda
strategies, (2) analysis of how these behaviors evolved on social media across time, and (3) cross-platform investigation of how troll accounts exploited the link between Twitter and YouTube. Below, we seek to address these omissions and examine the propaganda strategies that explain IRA behavior during the 2016 U.S. presidential election.

**Hypotheses and Predictions**

Given the two potential goals of the IRA during the 2016 U.S. presidential election as identified by journalists, researchers, and the U.S. government, we begin by formulating two general hypotheses, one for each of the two schools of thought:

**Hypothesis 1 (H1):** IRA Twitter accounts sought to increase the vote for Donald Trump at the expense of Hillary Clinton during the 2016 U.S. presidential election.

**Hypothesis 2 (H2):** IRA Twitter accounts sought to sow discord by promoting both conservative and liberal arguments and candidates during the 2016 U.S. presidential election.

While the overall goals of supporting a preferred candidate and sowing discord in the United States are not mutually exclusive, one should nevertheless expect different behaviors based on how strongly these goals are prioritized. That is, if the IRA’s primary goal was to support Donald Trump, one would expect that IRA accounts would share more conservative content than liberal content (H1). Conversely, if the IRA’s primary goal was to sow discord rather than support one candidate over another, one would expect relative balance in the ideological content shared by these accounts (H2).

Our primary approach for testing these hypotheses is to investigate the ideological slant of the hyperlinks shared by IRA accounts, which we describe below in further detail. Because our hypotheses are broad by definition, we assess four pairs of sub-hypotheses that draw upon prior literature concerning propaganda by focusing, in turn, on (1) the ideology of the individual hyperlinks, (2) the ideology of troll accounts, (3) the consistency of the accounts’ ideological orientation, and (4) patterns in the ideology of shared links across time. For each pair of hypotheses, the first hypothesis concerns our expectations if the primary goal of the IRA was to increase support for Donald Trump, and the second hypothesis our expectations if the primary goal was to sow discord by sharing partisan content from both sides of the ideological spectrum.²

**Link ideology**

**Hypothesis 1.1 (H1.1):** Assuming H1, IRA accounts should produce more links to conservative content than liberal content.

**Hypothesis 2.1 (H2.1):** Assuming H2, IRA accounts should produce equal amounts of links to liberal and conservative content.
Account ideology

**Hypothesis 1.2 (H1.2):** Assuming H1, the number of Russian IRA-controlled Twitter accounts that share primarily conservative content should be larger than the number of accounts that share primarily liberal content.

**Hypothesis 2.2 (H2.2):** Assuming H2, the number of Russian IRA-controlled Twitter accounts that share primarily conservative content should be approximately equal to the number of accounts that share primarily liberal content.

Ideological consistency

**Hypothesis 1.3 (H1.3):** Assuming H1, IRA accounts that primarily share conservative links should not share liberal content, whereas accounts that share primarily liberal content should share more conservative content.

**Hypothesis 2.3 (H2.3):** Assuming H2, IRA accounts should be consistent in the ideological content they share (i.e., a single account should not alternate between ideological extremes).

Temporal patterns of ideological consistency

**Hypothesis 1.4 (H1.4):** Assuming H1, the ideology of IRA Twitter accounts overall should consistently skew toward conservative content over time.

**Hypothesis 2.4 (H2.4):** Assuming H2, the ideology of IRA Twitter accounts overall should appear moderate with little fluctuation or swing equally from liberal to moderate.

The first pair of hypotheses captures expectations concerning the overall ideology of political content shared by IRA accounts, whereas the second pair focuses on the account as the unit of analysis. Importantly, the hypotheses in the third pair are motivated by prior literature on pre-propaganda strategies, wherein the target audience is fed content that is primarily consistent with their own ideology, with occasional nudges toward the propagandist’s desired outcome. Such a strategy could be seen as a way to mobilize liberals to support Donald Trump (H1), but less useful if the goal were to sow discord (H2). The final pair of hypotheses addresses temporal dynamics and suggests that, over time, the content shared by the IRA would show consistent patterns based on whether they were supporting one candidate or both parties equally.

Data and Measurement

**Shares of Media Organizations and YouTube Videos on Twitter**

To examine the ideological behavior of IRA trolls, we rely on a large collection of politically oriented tweets collected during and after the 2016 U.S. presidential election. The data were collected using Twitter’s streaming API between November 5, 2015, and December 31, 2018, and include 1,257,131,765 unique tweets mentioning
any of a large set of keywords related to the election. The keywords were designed to capture a wide array of tweets across the political spectrum, the list of which we present in Table A2 of the Supplementary Material. We note Twitter has released a related dataset of content authored by IRA accounts as part of its “Election Integrity” initiative.3 Twitter’s official data release contains a large collection of tweets, but many of these messages are not clearly related to the U.S. presidential election targeted by our dataset. As a robustness check, we have compared our election-oriented dataset to Twitter’s official release and find significant overlap between the two sets.

For analysis, we extract all tweets and retweets authored by any IRA account that was identified in the report to the Senate Intelligence Committee (U.S. House of Representatives 2018). This collection yielded 108,781 tweets from 1,052 IRA accounts. From these tweets, we then extracted 30,662 unique URLs sent by IRA members across 2,002 unique domains. In all, 10,450 of the URLs link to news stories by national media organizations, and 855 link to 499 unique YouTube videos across 315 YouTube channels.4

**Measurement Strategy**

To test our hypotheses, we first require estimates of (1) the ideology of the links sent by Russian trolls and (2) a measure of how each troll presents themselves ideologically on social media. For the latter, the ideology of each troll does not represent the “true” ideology of the actors who operate IRA accounts. Instead, it represents the ideology of the accounts as they appear as if they were ordinary users. Such a measure is useful because the goal of those who ran the IRA accounts was to falsely represent themselves to the public as ordinary social media users engaged in standard political discourse. Assessing the ideological presentation of IRA users in this way thus allows us to investigate the extent to which IRA accounts sought to appear as if they were liberal, conservative, or otherwise.

As our empirical goal is to investigate the role of link-sharing in the propaganda strategies of the IRA, we use links themselves as data to estimate the ideology of news media and IRA accounts. Our measurement strategy proceeds in two steps. First, to estimate the ideology of the news media links shared by the trolls, we use a recent method developed for scaling media domains based on the sharing of links to political news stories by politicians and ordinary users. Second, to measure the ideology of YouTube videos and channels, we use manual coding of videos posted by IRA accounts.

*Inferring ideology from URLs.* Our primary measurement task is to infer the ideology of media organizations from the URLs contained in tweets, which we can then use to infer the ideology of the troll accounts sharing these links. This process proceeds in two steps: To be begin, we rely on a recent method proposed by Eady et al. (2019) that simultaneously estimates the ideology of politicians, users, and news organizations using data from hyperlinks to news media domains (e.g. foxnews.com, nytimes.com) shared on social media. We then aggregate these news-media ideology scores to come
to come up with a troll-specific ideology estimate based on its sharing behavior. The intuition behind the method is straightforward: Social media users are assumed to be more likely to share links to news organizations that are ideologically close to them than they are those that are ideologically distant. A user who shares news stories from msnbc.com or democracynow.org, for instance, is more likely to be ideologically liberal than a user who generally shares stories from, for example, foxnews.com or breitbart.com.

The statistical model is a latent-space model that uses as data the frequency of the news media domains tweeted by each user. The model is specified as follows. Let $y_{img}$ denote the count of the news media domain $m = 1\ldots M$ tweeted by social media user $i = 1\ldots N$ who belongs to group $g \in \{R, D, U\}$ (Democrat, Republican, Unaffiliated). To concretize this model, Table 1 presents an example input to this model, containing the frequencies with which media organizations were tweeted by well-known political actors. These data make up an $N \times M$ count matrix, such that each cell $y_{img}$ represents the number of times that a given social media user has tweeted a story from a given news source.

The model works by introducing two latent variables, $\theta_i$ and $\zeta_m$, that denote the political ideology of social media users and news media organizations respectively. The probability that a user tweets the observed number of links to a media domain ($y_{img}$) is modeled as a function of the distance between that user’s own political ideology and the ideology of the media domain as follows:

$$y_{img} \sim \text{NegBin}\left(\pi_{img}, \omega_m\right)$$

$$\pi_{img} = \exp\left(\alpha_i + \gamma_m - \| \theta_i - \zeta_m \|^2\right)$$

where $\alpha_i$ denotes a user intercept, $\gamma_m$ denotes a media domain intercept, and $\omega_m$ denotes a media organization dispersion parameter. Substantively, the parameter $\alpha_i$
represents the relative degree to which a given user tweets links to news media, and $\gamma_m$ denotes the relative degree to which a media organization’s domain is tweeted. We then estimate the ideology of news media domains using Twitter timeline data from a large set of ordinary users and national political actors (e.g. Members of Congress, governors, the current and former Presidents). We exclude links shared by the IRA trolls in this step because, as Jessee (2016) shows, the structure that underlies political ideology can be influenced by the sample of actors included in the model. Because the sharing behavior of IRA trolls is unlikely to follow the same underlying ideological structure as ordinary citizens and politicians—rather, trolls are likely to use it strategically—we exclude the URLs shared by trolls here. The model is fit in a Bayesian context, placing priors on the parameters as necessary for model identification. Details of this estimation procedure are provided in Eady et al. (2019). As they show, ideology estimates using this model for Members of Congress have considerable convergent validity when NOMINATE scores are used as a baseline measure of ideology ($\rho = 0.96$, $\rho_{Rep} = 0.63$, $\rho_{Dem} = 0.62$).

We use these ideology estimates of media organizations to estimate the ideology of the IRA trolls. Each troll’s ideology is calculated as the mean ideology of the links to media organizations that it shares, weighted by the frequency with which stories from each media organization is tweeted. Finally, to ease interpretation of our results, we code news media organizations into discrete ideology categories representing “Liberal,” “Moderate,” and “Conservative.” Media organizations with a ideology score that is as liberal as, or more liberal than, The Washington Post are labeled “liberal”; those with a score as conservative as, or more conservative than, The Wall Street Journal are labeled “conservative”; and those in between are labeled “moderate.” A list of the 145 news domains and their inferred ideologies is provided in Table A4 in the Supplementary Material.

The ideology of YouTube videos. Like Twitter, YouTube is a complex ecosystem of content producers and consumers where videos are grouped into channels devoted to a diverse range of cultural and political topics. Unlike Twitter, however, the properties of YouTube’s underlying social network (e.g., follower networks, user interactions, etc.) are difficult to measure. While many media organizations maintain YouTube channels, many popular channels are devoted to celebrities and pop culture, which mix in occasional political content. These issues—along with the inherently visual content on YouTube—make it difficult to develop and apply automated techniques for inferring the political ideology of individual YouTube videos.

To address these difficulties, we therefore turn to human annotation to measure the ideology of each video shared by IRA accounts. While for news media we use links to infer ideology at the domain level of the media organization, for YouTube videos we perform manual annotation at the level of each individual video. Labeling videos is a manageable task for manual classification because the number of videos shared by IRA accounts is not overwhelming. For this manual classification task, we follow a two-step coding scheme in which annotators first code each video for whether it
contains politically relevant content; second, if the video is coded as political, it is further coded for ideology, as “Liberal,” “Moderate,” “Conservative,” or “Unclear.”

To facilitate this process, we employed three annotators selected from a set of student research assistants. In collaboration with these research assistants, we iteratively developed a code book defining these classifications, as we discuss in the Supplementary Material. For cases in which annotators disagreed on labels for a given YouTube video, we assign that video the majority class label. If no majority exists, the video is discarded from the analysis. Following several iterations of code book development and annotator training, our coding procedure results in a Krippendorff’s $\alpha$ reliability value (Krippendorff 1980) of 0.93 for the coding of whether the videos are political or not, and 0.79 for the coding of videos’ political ideology. Further details are provided in the Supplementary Material, as is the list of YouTube videos and their ideological classifications.

Results

Aggregate Link Data

In this section, we examine how the IRA used Twitter to amplify ideological content by sharing hyperlinks to news media and YouTube videos and channels.

Our first finding is simply that the use of hyperlinks in tweets played an important role in the IRA’s communication strategy. Among the 108,781 tweets posted by the 1,052 IRA accounts, 35,549 (33 percent) in our collection contain links to sites other than Twitter. Of these tweets, 10,450 (10 percent) link to news stories from media organizations for which we are able to calculate an ideology score. In addition, 3 percent of the hyperlinks shared by IRA trolls were to YouTube (1,205 tweets). As we show in Figure 2, YouTube is the second-most frequently linked domain in our data, suggesting that YouTube played a nontrivial role in the IRA’s social media strategy.

Next, we focus on the links sent by IRA trolls to news domains. As we illustrate in Table 2, IRA troll accounts link to conservative news sources more frequently than to liberal ones, accounting for 34 and 24 percent respectively, of the 10,450 links to news media websites. This is consistent with prediction H1.1 to the extent that conservative content is shared more frequently by IRA accounts than liberal content. Furthermore, a considerable proportion of the hyperlinks amplify articles from moderate sources, accounting for 42 percent of shared links.

Figure 1 illustrates how these distributions change over time and shows that IRA accounts became more active in the summer and fall of 2016. Patterns of ideological sharing alternate between conservative and moderate, with the mean ideology of shared links consistently between moderate and conservative throughout the election cycle. For a brief period during the summer of 2016 (around the national conventions), news from liberal and moderate sources is shared by trolls more frequently than conservative news. This time frame is the only relatively long period during which sharing from liberal and moderate news sources outpaces that of conservative news, barring periods when there was little news media shared at all. September 2016 sees a shift in
mean ideology toward more conservative news media, when Hillary Clinton fell ill during a campaign event, and the week of the Access Hollywood video leak, in which Donald Trump was recorded making derogatory comments about women. This skew toward conservative content over time is consistent with H1.4 to the extent that IRA accounts increasingly favored moderate and conservative content over liberal content over time.

As shown in Figure 2, the list of most frequently used domains is not only dominated by alternative media, such as Breitbart News, but also reputable mainstream media such as the *Washington Post* and *The Hill*. Of the ten domains, seven refer to news sources; the remaining three domains in Figure 2 indicate IRA accounts heavily share YouTube videos, have attempted to monetize their content sharing via viid.me (a link shortener that forces the user to watch an advertisement prior to redirecting to the actual destination link), and have relied on automated Twitter posting via Twibble.io (an automated Twitter posting service that injects links back to twibble.io).

These results contradict the popular notion that the IRA relied exclusively on hyper-partisan sources or “fake news.” Instead, the data suggest that the IRA deployed a diverse range of news sources, which may indicate that it sought to amplify both conservative and liberal sides of the presidential campaign. At the same time, the shift
toward a more moderate tone during the summer of 2016, followed by an increase in ideological sharing in the fall, may be evidence of an effort to first gain an audience among moderate voters and later pull that audience toward the Republican candidate (i.e., consistent with pre-propaganda techniques).

Unlike links to news media, those to YouTube were overwhelmingly conservative. As shown in Table 2, over 75 percent of links to YouTube were to conservative content. As with news media sharing, Figure 3 shows that IRA accounts increased their sharing of YouTube videos in the lead-up to election day, with an additional spike in sharing in early 2016. Furthermore, similar to news media sharing, the mean ideology of YouTube videos consistently leaned conservative.

Investigating link-sharing behavior further, we also examine the distribution of the type and ideology of YouTube videos shared during this time. First, as we show in Figure 4, the vast majority of shared videos contained political content (96 percent). As Table 2 shows, only 25 percent of shared political YouTube videos were ideologically moderate or liberal. The frequency of conservative YouTube videos is further
The ten most frequently shared videos account for 23 percent of links shared to YouTube videos by the IRA accounts (Table 3). All ten of these videos are pro-Republican. The most frequently shared video by IRA trolls, for example, illustrated in Figure 5, showing that IRA accounts overwhelmingly linked to conservative videos.

Instead of sharing a diverse range of YouTube content, the IRA promoted a narrow selection of highly popular videos. The ten most frequently shared videos account for 23 percent of links shared to YouTube videos by the IRA accounts (Table 3). All ten of these videos are pro-Republican. The most frequently shared video by IRA trolls, for example,
is a song accompanied by a compilation of pro-Trump images and titled “Trump About You.” Its description is Donald Trump’s campaign slogan, “Make America Great Again!” Similarly, Table 4 shows that the ten most frequently shared YouTube channels account for 40 percent of channels the IRA shared. Each of these channels was also coded as conservative by annotators, suggesting that IRA accounts were consistent in sharing conservative YouTube videos.

**Ideological Sharing across Twitter and YouTube**

In this section, we examine differences in the IRA’s use of news media and YouTube videos. To begin, we calculate the mean ideology of media domains shared by troll accounts. This metric provides a summary measure of the ideology of link-sharing by each account. The ideological distribution of troll accounts based on their news-media sharing behavior is presented in Figure 6. Of the 379 troll accounts that posted links to news media, 26 percent are liberal, 45 percent are moderate, and 29 percent are conservative. The higher prevalence of conservative accounts relative to liberal accounts is consistent with prediction H1.2, although the difference is slight. Furthermore, Figure 7 illustrates the news-sharing behavior of IRA accounts in the lead-up to the election. These time series data suggest that troll behavior is relatively consistently throughout the election cycle with respect to the sharing of news media content, and thus consistent with prediction H2.3.

Also evident in these time series data is that, while conservative troll accounts are outnumbered by moderate ones, conservative accounts are significantly more active, posting nearly twice as many links as moderate accounts. Although consistent ideological sharing and similar numbers of troll accounts on the ideological left and right may support the hypothesis that troll accounts support both sides (H2.1, H2.3), the increase in conservative troll activity without a matched increase from moderate and liberal trolls suggests stronger support for the Republican candidate (H1.1, H1.3).

We now examine whether individual troll accounts were ideologically consistent in their sharing of YouTube and news media. Among IRA accounts in our data, 14 percent (147) linked to both YouTube and news websites, representing 39 percent of the
Table 3. Most Shared Election-Oriented YouTube Videos and Their Channels.

<table>
<thead>
<tr>
<th>YouTube Video Title</th>
<th>YouTube Channel</th>
<th>Ideology</th>
<th>Share Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trump About You</td>
<td>American Right Now</td>
<td>Conservative</td>
<td>63</td>
</tr>
<tr>
<td>Ben Carson, CNN and Ted Cruz</td>
<td>Connect the Political Dots</td>
<td>Conservative</td>
<td>18</td>
</tr>
<tr>
<td>Jimmy Carter: I Would Choose Trump</td>
<td>Ted Cruz</td>
<td>Conservative</td>
<td>13</td>
</tr>
<tr>
<td>Ted Cruz with Neil Cavuto after the #GOPDebate</td>
<td>Ted Cruz</td>
<td>Conservative</td>
<td>12</td>
</tr>
<tr>
<td>Rigging the Election—Video II: Mass Voter Fraud</td>
<td>Project Veritas Action</td>
<td>Conservative</td>
<td>11</td>
</tr>
<tr>
<td>African Americans for Donald Trump 2016—Part 2</td>
<td>Mark F</td>
<td>Conservative</td>
<td>8</td>
</tr>
<tr>
<td>The Truth About the New York Bombing</td>
<td>Paul Joseph Watson</td>
<td>Conservative</td>
<td>7</td>
</tr>
<tr>
<td>WOMEN VOTING FOR TRUMP! “Standing Strong”</td>
<td>Mark F</td>
<td>Conservative</td>
<td>6</td>
</tr>
<tr>
<td>Full Speech: Donald Trump Speaks To Fired Up Crowd in Las Vegas (12-14-15)</td>
<td>Right Side Broadcasting Network</td>
<td>Conservative</td>
<td>6</td>
</tr>
<tr>
<td>NO JOKE: FIDEL CASTRO JUST ENDORSED BERNIE SANDERS</td>
<td>The Next News</td>
<td>Conservative</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 4. Most Shared Election-Oriented YouTube Channels.

<table>
<thead>
<tr>
<th>YouTube Channel Title</th>
<th>Ideology</th>
<th>Shares</th>
</tr>
</thead>
<tbody>
<tr>
<td>American Right Now</td>
<td>Conservative</td>
<td>63</td>
</tr>
<tr>
<td>Ted Cruz</td>
<td>Conservative</td>
<td>44</td>
</tr>
<tr>
<td>Right Side Broadcasting Network</td>
<td>Conservative</td>
<td>33</td>
</tr>
<tr>
<td>The Alex Jones Channel</td>
<td>Conservative</td>
<td>23</td>
</tr>
<tr>
<td>The Next News Network</td>
<td>Conservative</td>
<td>18</td>
</tr>
<tr>
<td>Connect the Political Dots</td>
<td>Conservative</td>
<td>18</td>
</tr>
<tr>
<td>Paul Joseph Watson</td>
<td>Conservative</td>
<td>17</td>
</tr>
<tr>
<td>Mark F</td>
<td>Conservative</td>
<td>17</td>
</tr>
<tr>
<td>H. A. Goodman</td>
<td>Conservative</td>
<td>14</td>
</tr>
<tr>
<td>Project Veritas Action</td>
<td>Conservative</td>
<td>12</td>
</tr>
</tbody>
</table>

379 accounts that shared links to news media. This subset of troll accounts were also those that were highly active in their volume of posts in general. On average, troll accounts within this subsample posted 350 election-related tweets, compared to 64 tweets posted, on average, from the remaining IRA accounts.
Lastly, in Figure 8, we show the relationship between the ideology of YouTube videos and news media shared by individual troll accounts. Each point represents an IRA account, where the size is proportional to the number of tweets it posted. The horizontal axis represents the mean ideology score of the news websites shared by a given IRA account, and the vertical axis, the mean ideology score of the YouTube videos it shared. The histograms show the respective distributions of news media and YouTube ideology. As the figure shows, the correlation between the ideology of news media and YouTube sharing is relatively moderate ($\rho = .47$). Trolls who have high (i.e. conservative) ideology scores for news website sharing also share more links to right-wing YouTube videos. Interestingly, troll accounts that shared more liberal news media appear more likely to cross ideological lines than conservative accounts when sharing YouTube videos. That is, accounts that share liberal news media also share conservative YouTube videos, whereas IRA accounts that share conservative news rarely share liberal videos. Although many trolls shared liberal and moderate news media—potentially as pre-propaganda—YouTube video content was primarily used to promote pro-Republican, pro-Trump content.

**Robustness of Ideological Classification**

This article relies on aggregating the ideology of news media domains to infer how each IRA account sought to present itself ideologically to ordinary Twitter users. This connection between media domains and ideology has been previously established for politicians by Pew Research Center (Messing et al. 2017; Mitchell and Weisel 2014), as these works show politicians’ ideologies are predictive of the news domains they share. This approach is limited, however, to the extent that it does not use other text content in each tweet. A troll account might, for example, express liberal views in the text of tweets, but share links to conservative news media, or share such links to comment on them with criticism or ridicule. Fortunately, other researchers have evaluated
Figure 7. IRA account news media sharing behavior by ideology: (a) liberal IRA accounts, (b) moderate IRA accounts, and (c) conservative IRA accounts.

Note. These figures show the sharing behavior of IRA accounts over time, decomposed by their mean ideologies. Results suggest IRA accounts behave consistently over time. IRA = Internet Research Agency.
the ideology of these accounts using qualitative coding of text content for many of the IRA accounts in our data (Linvill and Warren 2018, 2019). This evaluation allows us to evaluate the robustness of our link-based classifications by comparing them with the text-based qualitative classifications by Linvill and Warren (2018). We focus on the classifications used in Linvill and Warren (2018) rather than Linvill and Warren (2019) since the more recent analysis presents only hashed usernames, making it difficult to match across classifications. Compared to Linvill and Warren (2018), differences between the text- and link-based measures are minimal: Classifications based on link-sharing agree with those by Linvill and Warren (2018) in identifying liberal and conservative trolls with 87- and 96-percent precision respectively, as shown in Table 5.6.

In analyzing instances where our approach fails this robustness check, however, we find an interesting case with the IRA troll account “@BleepThePolice.” Rather than this account sharing a domain and criticizing the article’s content, the @BleepThePolice account changed both its identity (it was originally called @Chad_Shorttw) and its patterns of link-sharing. While both identities maintained the same unique Twitter account ID number, this earlier @Chad_Shorttw identity shared more than twice the number of media links, many of which were conservative (e.g., links to Breitbart.com

Figure 8. Mean ideology for YouTube and news website links.
and RedState.com). Conversely, the @BleepThePolice shared many fewer links to media organizations and adopted a more liberal position both in moniker and content shared. We have since examined the other accounts in our dataset and find 17 percent of troll accounts change their screen names, though many of them maintain a similar identity (e.g., @AtlantaBreaking changed to @Atlanta_Online). Since these accounts maintain the same Twitter account ID, and other users will observe all of the content from these accounts regardless of their screen name when viewing each account’s timeline, we do not treat these accounts as separate entities.

**Discussion**

In this article, we investigated the IRA’s use of ideologically liberal and conservative news media and YouTube videos during the 2016 U.S. presidential election. We examined two broad hypotheses concerning the extent to which the IRA took ideological sides during the campaign, and we explored the empirical implications of our hypotheses by examining the link-sharing behavior of IRA troll accounts.

Our results suggest that the IRA had a multifaceted approach to ideological news and YouTube content. On one hand, the IRA’s consistent use of ideologically diverse sources, overall consistency in ideology, and the large number of moderate and liberal accounts relative to conservative accounts could suggest that the IRA sought to sow ideological division. This evidence could be interpreted as an attempt to increase polarization in the United States and weaken its domestic political fabric. On the other hand, the prevalence of conservative content and YouTube videos, highly active conservative accounts, and conservative lean of the IRA sharing over time are all consistent with attempts to support the Trump campaign.

While the two strategies of sowing discord and supporting one candidate over the other can be thought to exist in isolation, it is reasonable to consider that they may also
co-exist, especially in light of the previous propaganda literature. That is, a propaganda campaign can target conservatives and moderates/liberals with different objectives, but with an overarching goal of supporting the Republican campaign. As shown in Figure 6, a relatively clear ideological division of labor exists among IRA accounts, with some sharing primarily conservative content, and others liberal content. In this interpretation, the first set of accounts, comprised of trolls who primarily linked to conservative news media, is tasked with spreading false consensus and making conservative talking points appear more widely accepted than they actually are. Critically, these conservative-leaning accounts need not subvert or lure their audience. The second set of accounts, comprised of trolls who link primarily to moderate- and liberal-leaning news media, instead connects with more liberal audiences, sharing news media that signals membership in this group to engage in pre-propaganda, but occasionally linking to conservative media to drive support away from Hillary Clinton. Under this interpretation, differences among liberal and conservative IRA accounts in sharing YouTube content—with conservative accounts linking to conservative content, and liberal accounts linking to both liberal and conservative content—are especially consistent with theories of pre-propaganda.7

More specifically, the conservative YouTube videos are not only propagated by trolls who link to conservative news media, but also by those who link to liberal and moderate news websites. This pattern can be theoretically interpreted as a form of pre-propaganda (Ellul 1973), where the ideologically diverse news media is used to gain a broad audience and credibility. This behavior, in turn, can be used to inject conservative YouTube content into the feed of users who are normally drawn to liberal or moderate news sources.

We advance this theoretical interpretation for two reasons: First, the theory explains why a substantive subgroup of seemingly “liberal” or “moderate” IRA accounts, based on links to news outlets, share links to conservative YouTube outlets and how this behavior is not present among conservative accounts. Second, pre-propaganda is a classical and common tactic that state and non-state actors use to promote specific messages without appearing overtly political or partisan. It is important to note that pre-propaganda does not imply that trolls appear liberal or moderate in one point in time before shifting to consistently conservative behavior. On the contrary, propaganda often takes place by continuously mixing direct propaganda with moderate content to gain credibility—a common and well-documented strategy that predates social media (Jowett and O’Donnell 2014). Our study suggests that cross-platform links between Twitter and YouTube may potentially play an important role in these pre-propaganda strategies. As mentioned, pre-propaganda is not historically new, nor is it limited to Russia. However, our study shows nations may adapt the classical technique in new ways to leverage the increasingly complex media environment, where users constantly shift between different social media platforms. We refer to this technique as “cross-platform pre-propaganda.” Here, pre-propaganda is carried out by exploiting the interconnected nature of the different platforms—in this case, the Twitter–YouTube link. We must, however, emphasize
that further research is needed in order to confirm whether IRA’s reliance on conservative YouTube videos is indeed intentional and strategic.

A key motivation for this additional research is the possibility that YouTube itself contains a large quantity of conservative content, and the Russian IRA’s use of the Twitter–YouTube link to convey right-wing messages could be a reflection of that reality. Several commentators have claimed that right-wing communities dominate YouTube (Whyman 2017), while Kaiser and Rauchfleisch (2018) argue that the platform is the “backbone for the far-right in the United States.” In particular, commentators and journalists alike have raised concerns that YouTube’s recommendation algorithms themselves may have been biased in favor of pro-Trump content in the context of the 2016 U.S. elections (Lewis and McCormick 2018). From this perspective, the sheer quantity of easily available, conservative content on YouTube may make it the preferred source for foreign propaganda campaigns who wish to utilize right-wing content.

Future research could explore this question of ideological distribution in YouTube in a number of ways. One such path could replicate the method we describe above for inferring ideology from news domains, replacing or augmenting news domains with YouTube channels. This method would project YouTube channels into a similar ideological space, which we could then use to estimate the distribution of liberal and conservative channels shared by politicians in Twitter. Such an approach would be necessarily biased toward videos shared within a platform like Twitter, but, given YouTube’s prevalence in Twitter, this limitation may not be critical. Alternatively, one could sample the space of YouTube IDs (as suggested in Zhou et al. 2011) and manually assess the political and ideological components of these samples, thereby decoupling this evaluation from sharing dependencies. From these distributions, however, one could then evaluate whether the distribution of conservative YouTube content shared by Russian IRA accounts was consistent with randomly sampling from the platform or evidence of a more strategic choice. Alternatively, the YouTube domain may carry less ideological signal than the New York Times or Breitbart, making it ideal for exposing individuals to pre-propaganda content that might otherwise conflict with their ideological preferences.

In addition, our analysis is limited to content shared by the Russian IRA that appeared in our collection of political tweets during the 2016 U.S. election, and more research is needed to provide insight into the overall ideological distribution of YouTube videos shared by the IRA trolls. Furthermore, the propaganda patterns presented in this study are limited to content directly related to the 2016 elections. It is possible that the IRA has a more vocal support for liberal views through links to both news media and YouTube videos in other political arenas, such as lesbian, gay, bisexual, and transgender (LGBT) rights or the politics of race and identity. Thus, further research is needed that explicitly examines whether Russian covert propaganda campaigns use content from YouTube to support different ideological sides depending on the topic, or whether the strategy remains the same across different political issues. Here, we contribute to the ongoing discussion by highlighting the important role of the cross-platform link between Twitter and YouTube in the IRA’s attempt to promote
It is highly plausible that the Russian agency would use similar cross-platform techniques in other contexts outside of the 2016 U.S. presidential election. These findings could therefore offer an important starting point for researchers who want to explore the strategies behind Russian covert campaigns in other countries, such as France, Germany, or the United Kingdom, as well as in subsequent elections in the United States.

It is also important to note that the success of the conservative candidate was not a foregone conclusion leading up to the election. As such, the Russian IRA, having established an audience, might have been hesitant to out themselves as overwhelmingly pro-conservative before the election, as these accounts could have been intended for use in a campaign to delegitimize a Hillary Clinton presidency if she won. Indeed, perhaps that was the plan all along. Under such a scenario, the consistent ideologies of Russian IRA accounts also makes sense and lends support to the idea that the Russian IRA employed a diverse set of strategies to cover several possible outcomes (or that they never anticipated Trump would win the election).

Conclusion

This article presents an analysis of the Russian IRA’s ideological sharing behaviors of links to political content and YouTube videos on Twitter and contextualizes these observations with prior literature on propaganda. It extends existing research on the behavior of IRA trolls during the 2016 U.S. presidential election into the realm of the links shared by these accounts to news media domains and YouTube. We test two competing sets of hypotheses drawn from existing research, one set based on the idea that troll accounts were primarily trying to help increase support for Donald Trump and conservative candidates and the other that IRA accounts were agnostic to providing support for one side versus the other and instead were trying to enflame partisan divisions. While we find some support for both sets of predictions, on balance our empirical evidence is more consistent with the hypothesis that the Russian IRA was primarily interested in supporting the Republican election efforts during the 2016 U.S. presidential campaign.

We also find evidence of what the propaganda literature has called “pre-propaganda,” which may explain the prevalence of liberal-leaning Russian IRA accounts and why these liberal accounts are more likely to share conservative YouTube videos. This mixed evidence suggests the IRA’s strategy during the 2016 election was consistent but multifaceted and developed to account for multiple potential outcomes and uses. As such, we suggest that the introduction of online social platforms has not resulted in a wholly new set of propaganda strategies; instead, these strategies have evolved to take advantage of new technology. This paper also demonstrates the advantages of analyzing (and conceptualizing) political communication in the context of a broader media ecology: States can exploit the interconnected nature of the modern information ecosystem to pursue covert propaganda strategies, which means scholars and analysts should be doing thinking this way too.
Declaration of Conflicting Interests
The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research was carried out as part of NYU’s Center for Social Media and Politics, directed by Joshua A Tucker, Jonathan Nagler, and Richard Bonneau. The Center is supported by funding from the Knight Foundation, Charles Koch Foundation, Craig Newmark Philanthropies, Hewlett Foundation, Siegel Family Endowment, Gates Foundation, and the National Sciences Foundation (SES-1756657). In addition, the research was carried out as part of the ‘Digital Disinformation’ project, directed by Rebecca Adler-Nissen and funded by the Carlsberg Foundation (project number CF16-0012).

ORCID iD
Cody Buntain https://orcid.org/0000-0003-4797-3726

Supplemental Material
Supplemental material for this article is available online.

Notes
1. While it is empirically difficult to distinguish between different “shades” of (gray) propaganda, one could alternatively argue that the Internet Research Agency (IRA) is on the darker end of the spectrum, especially in instances where it used fake accounts to share truthful content.
2. It is of course possible that different accounts were pursuing different strategies, or that the same accounts pursued different strategies at different times. Diverse behaviors could be part of a complex overall strategy, or simply the result of decentralization and/or a lack of an overall strategy. However, here we are focusing on highlighting hypotheses that would be consistent with the pursuit of one of the two general strategies.
4. Many URLs in tweets are shortened by link shorteners (e.g. bit.ly, ow.ly). As a preprocessing step, we unshorten all links using the Python library urlExpander.
5. The model is fit using the Bayesian inference engine Stan (Carpenter et al. 2017).
6. Our comparisons focus on Linvill and Warren’s (2018) “LeftTroll” and “RightTroll” categories. The authors do not code a “moderate” category.
7. Moreover, it is important to remember that even when linking to a liberal or moderate media source, it is always possible that the content shared by that moderate or liberal source could still weaken support for Clinton (e.g., a Washington Post discussion of Clinton’s emails).

References


**Author Biographies**

**Yevgeniy Golovchenko** is a research assistant at the Department of Political Science at the University of Copenhagen and a member of the Digital Disinformation research group and the Copenhagen Center for Social Data Science (SODAS). His research interests include disinformation, propaganda, censorship and political participation on social media.

**Cody Buntain** is an assistant professor in the Informatics Department at the New Jersey Institute of Technology. Prior, Dr. Buntain was a postdoctoral researcher at New York University’s Social Media and Political Participation (SMaPP) lab and was an Intelligence Community Postdoctoral Fellow.

**Gregory Eady** is an assistant professor of political science at the University of Copenhagen and a former postdoctoral fellow at the SMaPP lab.

**Megan A. Brown** is a research scientist at the Center for Social Media and Politics at New York University. Her research focuses on the information environment on social media.

**Joshua A. Tucker** is a professor of politics, co-director of the Center for Social Media and Politics and SMaPP lab, and director of the Jordan Center for the Advanced Study of Russia at New York University.