INFORMATION, COMMUNICATION & SOCIETY
https://doi.org/10.1080/1369118X.2018.1499793

Routledge
Taylor & Francis Group

Check for updates

Who is responsible for Twitter's echo chamber problem? Evidence from 2016 U.S. election networks

Lei Guo, Jacob A. Rohde and H. Denis Wu

College of Communication, Boston University, Boston, MA, USA

ABSTRACT

This study examined the echo chamber phenomenon and opinion leadership on Twitter based on the 2016 U.S. presidential election. Network analysis and 'big data' analytics were employed to analyze more than 50 million tweets about the two presidential candidates, Donald Trump and Hillary Clinton, during the election cycle. Overall, the results suggested that Twitter communities discussing Trump and Clinton differed significantly in the level of political homogeneity and opinion leadership, and that certain opinion leaders were responsible of creating homogeneous communities on Twitter. This study made a theoretical contribution to the literature by linking opinion leadership and Twitter's network structure and shedding light on what may have caused the echo chamber problem to happen in an emerging media landscape.

ARTICLE HISTORY
Received 26 December 2017
Accepted 3 July 2018

KEYWORDS
Twitter; presidential election; opinion leader; political homophily; network analysis; big data

Extreme partisanship has become a primary concern in many democratic countries for its role in hindering any constructive, forward-looking legislative or policy-making endeavor. Many blame social media, such as Twitter, for exacerbating the problem. As researchers have observed, users on Twitter tend to interact more with those who share the same political ideology, thus forming fragmented, yet homogeneous communities (Colleoni, Rozza, & Arvidsson, 2014). Members within these Twitter communities merely echo each other's viewpoints, reinforcing existing perspectives. Although this 'echo chamber' effect has been extensively researched (e.g., Colleoni et al., 2014; Freelon, Lynch, & Aday, 2015; Himelboim, McCreery, & Smith, 2013), the crucial questions as to how and what have caused this to happen remain unanswered.

This study aims to address this issue by investigating the characteristics of Twitter communities involving political candidates and opinion leaders who could have influenced the degree of opinion diversity in their circles. The main goal of this investigation is to determine whether opinion leadership on social media affects the echo chamber phenomenon. That is, can certain individuals or institutions trigger a more homophilous or homogeneous community on Twitter? Explicating the potential connection between opinion leadership and network structure of Twitter sheds light on the diversity of online political

CONTACT Lei Guo guolei@bu.edu College of Communication, Boston University, 640 Commonwealth Ave., Boston, MA 02215, USA

discourse and provides implications for facilitating a more democratic public sphere on this social networking platform.

The 2016 U.S. presidential election provides an ideal case for this analysis. Some believe that the partisan divide on political values contributed to the success of Donald Trump, a non-conventional politician (Bump, 2017). Among all media platforms, Twitter played an outsized role in this election. Donald Trump as well as Hillary Clinton treated the platform as a battleground for their respective electoral campaigns, resulting in more than one billion tweets during the election cycle (Coyne, 2016). To what extent did the echo chamber effect take place on Twitter during the 2016 election? Did Twitter communities present a unique feature with respect to the discussion of Trump? More importantly, who led the conversation on Twitter, and who contributed to the echo chamber phenomenon on this platform? Using the 2016 U.S. election as a case study, this paper empirically examines opinion leadership and the echo chamber effect on Twitter and the link between the two. Results of the study will not only enrich our understanding of the role Twitter played in the 2016 election, but also make a theoretical contribution to the two important areas of network research.

To provide comprehensive, empirical evidence for our research questions, network analysis and 'big data' analytics were used in this study. More than 50 million tweets from two large Twitter datasets were analyzed. Overall, our results showed that conversations about the two political candidates differed significantly in their levels of political homogeneity and certain opinion leaders did have the power to influence the nature of Twitter communities.

Literature review

Twitter and the echo chamber effect

Twitter, with its many social networking affordances, greatly increases citizens' exposure to political discussion and confrontation. The question of whether this social platform facilitates the formation of an open online public sphere or leads to political polarization has attracted considerable scholarly attention (e.g., boyd & Ellison, 2007; Kwak, Lee, Park, & Moon, 2010). Although mixed findings have been presented, a predominant view suggests that Twitter functions as an echo chamber that reinforces its users' established perspectives and opinions (Colleoni et al., 2014; Freelon et al., 2015; Himelboim et al., 2013).

The level of homophily within a circle has been linked to the echo chamber effect. Homophily is 'the principle that a contact between similar people occurs at a higher rate than among dissimilar people' (McPherson, Smith-Lovin, & Cook, 2001, p. 416). This tendency to associate with like-minded others can be explained by cognitive dissonance theory (Festinger, 1962). According to the theory, individuals tend to enjoy exposing themselves to congruent messages that support their existing beliefs; on the contrary, they will experience negative feelings when presented with divergent opinions.

To avoid cognitive dissonance, individuals will purposively choose among media channels and content that best align with their political perspectives – a process of selective exposure. Twitter is an open social networking platform where users can choose whoever to follow or interact with. This provides a fertile ground for selective exposure and, as a consequence, may engender a high degree of political homophily. A number of recent studies have shown that like-minded people associate with one another on Twitter in various political communication contexts (Colleoni et al., 2014; Freelon et al., 2015; Himelboim et al., 2013).

From a network perspective, political homophily can be best observed in *clusters* within a large network. Clusters are subgroups in a network in which nodes are substantially more connected to one another than to nodes from outside the cluster (Wasserman & Faust, 1994). On Twitter, clusters about a political candidate are formed by *retweeting*, *mentioning* other users in posts, or *replying to* others' messages. These Twitter clusters, or communities, can be based on real or imagined interpersonal relationships (Gruzd, Wellman, & Takhteyev, 2011), and can exist either over an extended period of time or temporally (Freelon et al., 2015). More importantly, the nature of Twitter communities may vary with subjects of discussion (Himelboim et al., 2013). That is, people may interact with ideological adversaries when speaking on certain topics but only choose to connect with allies on others. As such, people would form sub-publics on Twitter when discussing different political candidates (McKelvey, DiGrazia, & Rojas, 2014) and may present varied levels of political homophily. Therefore, we hypothesize:

> H1: Twitter communities about Trump and Clinton will differ in their levels of political homophily.

Another concept that can be used to examine the echo chamber effect is political homogeneity, which refers to the balance between divergent views in a community (Williams, McMurray, Kurz, & Lambert, 2015). While homophily focuses on nodes (i.e., users), homogeneity examines the nature of edges (i.e., messages). For example, Williams et al. (2015) showed that Twitter communities demonstrated a highly homogeneous distribution of attitudes toward climate change; most communities were dominated by either activist or skeptic views, with few communities having a mix of varied perspectives. Likewise, Himelboim et al. (2013) analyzed the proportions of liberal, conservative, and neutral-oriented political messages in Twitter communities, and found that users in a community are unlikely to be exposed to cross-ideological content. Notably, political homogeneity provides a different angle than homophily when examining political communication on Twitter. For example, it is possible that like-minded people interact with each other but exchange divergent views about a political candidate. Especially in the case of the 2016 election, both candidates were involved in a number of scandals, which could have resulted in some of their respective supporters both endorsing and criticizing their behaviors. As such, a homophilous community does not presuppose a homogeneous community. Therefore, we further hypothesize:

> H2: Twitter communities about Trump and Clinton will differ in their levels of political homogeneity.

Opinion leadership on social media

Research suggests that Twitter communities form around prominent opinion leaders, which are often in the form of popular individuals, celebrities, or organizations (Gruzd et al., 2011). On Twitter, a critical mass of followers usually exist beforehand and are waiting to be led by opinion leaders (Tremayne & Minnie, 2013; Watts & Dodds, 2007). In political communication, campaign managers often work with opinion leaders to help disseminate key information to their potential followers and ultimately win elections (Nisbet & Kotcher, 2009). As such, understanding social media's opinion leaders is a crucial step in explicating the formation of different communities on Twitter.

The notion of opinion leadership has its roots in the theory of two-step flow of information (Lazarsfeld, Berelson, & Gaudet, 1948), which asserts that new ideas usually flow from the mass media to opinion leaders. Opinion leaders then spread that information to their respective circles where it eventually reaches the general public. Traditionally, opinion leaders are conceptualized as those who have more access to news and information, are better educated, and of higher socioeconomic status (Rogers, 1983). Because of their advantageous positions in the information networks, opinion leaders, such as politicians and scholars, serve to 'collect, diffuse, filter, and promote the flow of information' between media and personal agendas (Brosius & Weimann, 1996, p. 564).

One critique of the two-step model asserts that the process of influence is more complex than what the model describes. This notion led to the emergence of the multistep flow paradigm (Weimann, 1982). This extended model suggests that opinion leaders communicate their ideas to followers who, in turn, spread those ideas to others. In addition, the information exchange is not unidirectional in that opinion leaders can be influenced by their followers and that media coverage can be influenced by audiences. Nonetheless, news media are still considered the center of communication in this paradigm.

In a more recent modification of the model, influential individuals and media organizations are seen as embedded in a multidimensional network (Ognyanova & Monge, 2013). In this network, especially in the context of social media such as Twitter, several assumptions of opinion leaders should be revisited. First, opinion leaders are not necessarily those of higher socioeconomic status. Twitter allows anyone to send information to the public, which, to some extent, lowers the bar of what constitutes an 'opinion leader' (Park, 2013). Avid individuals, such as bloggers and activists, without a recognizable social status can also be influential on Twitter. Second, opinion leaders are not necessarily mediators between the news media and the public. Rather, they may promote their own agendas to the public or even compete with the news media for public attention (Dubois & Gaffney, 2014). In this light, the new media environment is more decentralized, in which the news media are no longer the core of a given information network. Therefore, we suggest that the distinction between opinion leader and news media should be discarded and the term 'opinion leader' should be used broadly to refer to any individuals or entities including the media that have become influential in a certain network. In other words, opinion leaders should not be pre-determined, but instead explored in different contexts.

This opens the question as to who opinion leaders on social media are and who can exert more influence on this platform. Chang and Ghim (2011) showed that the majority of the most popular tweets produced in South Korea were created by ordinary users instead of traditional opinion leaders, such as politicians or scholars. On the other hand, researchers have also suggested that established groups of individuals and institutions, such as journalists and media organizations, continue to dominate the discourse on Twitter, extending their offline influence to social media (Bruns, 2012). Likewise, Xu, Sang, Blasiola, and Park (2014) found that organizations were more likely to influence Twitter-based political activism than individual users. Beyond what has been examined, it is reasonable to expect that individuals and institutions with different political orientations may exhibit different levels of impact on Twitter. Given the inconsistent findings and questions that remain unanswered, we first ask:

RQ1: Who are the opinion leaders in Trump's and Clinton's Twitter networks, respectively?

Just like the nature of communities may vary by topics, opinion leaders are also likely to be issue-specific (Nisbet & Kotcher, 2009). In a political election, varied types of opinion leaders may exert influence on Twitter discussion networks about different candidates. Thus, we hypothesize:

H3: Types of opinion leaders will vary in Twitter communities about Trump and Clinton.

Finally, and most importantly, opinion leaders may influence the nature of Twitter communities. On Twitter, opinion leaders can influence others by drawing their attention to certain topics and stimulating responses to messages they post (Bruns, 2012; Gruzd et al., 2011; Xu et al., 2014). It stands to reason that tweets from some opinion leaders may invite discussions that involve diverse viewpoints, while others would likely lead to partisan echo chambers. Thus, we assume that opinion leaders with certain characteristics would trigger their communities on Twitter to be more homophilous or homogeneous than others.

H4: Types of opinion leaders affect the level of political homophily in Twitter communities.

H5: Types of opinion leaders affect the level of political homogeneity in Twitter communities.

Method

This study examined opinion leadership and the echo chamber effect on Twitter during the 2016 U.S. presidential election. The analyses were based on two datasets: *Dataset I*: Network data of tweets collected from the network analysis application NodeXL through Twitter's public Application Programming Interface (API; Hansen, Shneiderman, & Smith, 2010); and *Dataset II*: Tweets collected using a hybrid method including scraping Twitter search results.

Network data collection

The network data (*Dataset I*) were collected using the built-in Twitter API search tool from NodeXL. Twitter's public API allows users to gather data about a popular hashtag or search term from public Twitter accounts. The professional version of NodeXL, used here, allows for retrieving up to 18,000 edges (i.e., tweets) per search.

Two groups of networks were retrieved to represent Twitter conversations surrounding the two final candidates of the 2016 U.S. election: Hillary Clinton and Donald Trump. Search terms 'Clinton' and 'Trump' were established as the criteria for whether or not a tweet was included in each network. The collection process began on 17 July 2016. This marked the first day of the Republican National Convention and also symbolized the final stretch of the 2016 election season where Democrat and Republican parties unified around their respective nominees. To retrieve a representative sample of Twitter networks over time, data were gathered once every Monday at around 5pm (Eastern Standard Time) until the day before the election on 8 November 2016. That is, up to 18,000 tweets and their network relationships were retrieved on Monday of each week during the sampled time period. In total, we collected 34 separate networks (17 for Trump and 17 for Clinton).

Once collected, the data were cleaned so that networks only contained tweets exclusively about each party's respective candidate. Instances where tweets mentioning both Trump and Clinton were removed. A total of 69,132 tweets about Clinton, and 159,181 about Trump were included in *Dataset I*.

Identify Twitter communities

To explore the nature of Twitter networks, the unit of analysis for this study was a Twitter community, or a *cluster* using the network analysis term (Wasserman & Faust, 1994). Two types of relationships that connect Twitter users were retrieved through NodeXL. Using the symbol @, Twitter user A can mention any other user B in a post, which indicates a tie from user A to B. Retweeting is included as a part of this relationship because when Twitter users retweet, they will have to mention the original sender of the tweet.

Twitter also allows its users to reply to other users' messages, which represents another type of relationship. When Twitter user A replies to user B's message, a tie from A to B is established. Based on these relationships, NodeXL offers a number of methods to breakdown each network into distinct intra-network clusters. This study used the Clauset-Newman-Moore algorithm because of its ability to handle large datasets (Himelboim et al., 2013). For this study, the output of this method often resulted in networks with hundreds of identified clusters, with many clusters only containing a handful of Twitter users. Little published research has provided an appropriate cutoff for analyzing clusters in a large network. As such, this study chose to only analyze prominent clusters that contained at least one percent of a network's total nodes (i.e., Twitter users). We believe this cutoff was appropriate as the aggregate of all prominent clusters resulted in capturing at least one-third of a network's total users.

Identify opinion leaders

The next step was to identify opinion leader(s) in each of the identified clusters. Both 'mention' and 'replies to' indicate influence of a Twitter user (Himelboim et al., 2013). When user A mentions/retweets or replies to user B's message, it indicates that user A pays attention to B's message and acknowledges its informational value. That is, user B influences user A. Therefore, to measure a Twitter user's influence in a cluster, we used in-degree centrality, which represents the amount of network ties directed toward a user (Wasserman & Faust, 1994). Previous research has defined an opinion leader as having at least 10% of the overall in-degree power within a network (Valente & Pumpuang, 2007). However, given the large size of some of the Twitter networks, this study instead used a 1% threshold. This criterion ensured that we can detect a wide range of users – as there would have been fewer opinion leaders that can meet the criteria of having at least 10% in-degree centrality. In doing use, a total of 189 unique opinion leaders were identified.

To examine the characteristics and types of opinion leaders, two graduate students were instructed to classify the identified opinion leaders based on their Twitter profile status and available tweets. For each opinion leader, the coders first decided whether the Twitter user was a person or an entity, and the user's partisanship (i.e., liberal, conservative, or unknown). For person-type users, the coders continued to classify them as politicians, journalists, entertainment celebrities, or bloggers or average Twitter users. Entity-type users were classified as traditional news media (e.g., *CNN*, *the New York Times*), online-only, emerging media (e.g., *Buzzfeed*, *Huffington Post*), non-media organizations (e.g., NRA, DNC/RNC, etc.), or other/unknown. Intercoder reliability was assessed on a random sample of 10% of the data (i.e., 19 Twitter users), reaching an average of 0.96 Krippendorf's alpha (α).

Sentiment analysis

This study used a supervised machine learning approach to measure political homophily and homogeneity. We classified twitter users as candidate supporters, critics, or neutral based on their tweets, which were coded as positive, neutral, or negative.

To build our supervised machine learning models, we first manually labeled a sample of tweets. Two coders were trained to classify tweets based on objective instances of mutually exclusive positive, negative, or neutral sentiment. An initial random sample of 1000 tweets from both Clinton and Trump networks were selected and 10% were classified by both coders to measure intercoder reliability (Clinton $\alpha = .82$; Trump $\alpha = .81$). After the remaining tweets were classified, they were fed into two separate machine learning models (i.e., one for Clinton-mentioning tweets and one for Trump-mentioning tweets). The Support Vector Machine (SVM), a machine learning algorithm considered superior for text analysis (Collingwood & Wilkerson, 2012), was used to classify tweets.

It turned out that the initial sample of 1000 tweets was not enough to train a model with satisfactory performance. It was especially challenging to reach satisfactory accuracy for classifying Trump-mentioning tweets, mainly due to the struggles with coding for Twitter sarcasm. Therefore, more tweets were manually labeled. After several rounds of iterations, a final sample of 2100 labeled tweets were used to build a model to classify Clinton-mentioning tweets, and 3600 labeled tweets were used to build a model to classify Trump-mentioning tweets. See Table 1 for the two models' performance.

Political homogeneity

Political heterogeneity, the opposite of homogeneity, is defined as the overall balance between members holding skeptic and activist views in a given community (Williams et al., 2015). This study adjusted Williams et al.'s (2015) formula and measured political heterogeneity in each cluster as

$$H = 1 - \left|\frac{p - n}{p + n}\right|$$

Table 1. SVM model performance.

	Clinton ($N = 2100$)		Trump ($N = 3600$)	
	Positive	Negative	Positive	Negative
Recall	.92	.83	.80	.77
Precision	.91	.83	.86	.78
F-measure	.91	.82	.82	.77

where p is the observed frequency of tweets that expressed positive sentiment toward a political candidate and n is the observed frequency of tweets that held negative sentiment toward the candidate. The formula gives a value on a linear scale that ranges from 0 (completely homogenous) to 1 (completely heterogeneous).

Political homophily

Political homophily is measured by examining the degree to which individuals interact with like-minded others. This is a more complex measurement than political homogeneity because it requires estimating *users*' political ideologies. The challenge is that one cannot determine a Twitter user's political ideology based on one tweet in a cluster. For example, it was not unusual for Clinton supporters to critique her in certain scenarios (e.g., the scandals she was involved in). Therefore, more data analyses were needed to decide whether a Twitter user was a candidate supporter or a critic.

Following Vargo, Guo, McCombs, and Shaw (2014), we determined a Twitter user's political orientation based on a sample of tweets posted by the user. To do so, our analysis drew upon another Twitter dataset *(Dataset II)* that contained 51,767,174 tweets mentioning candidates' names (Trump and Clinton) throughout the election cycle. The list of unique Twitter user handles was extracted from *Dataset I* and used to match users in *Dataset II*. The analysis was conducted separately for Trump- and Clinton-mentioning tweets. All matched users in *Dataset II* who tweeted three times or more about Trump were included in the analysis of political homophily in Trump-mentioning networks; users who tweeted at least three times about Clinton were included in the analysis regarding Clinton. Thus, a total of 3,272,683 tweets mentioning Trump and 1,569,926 tweets mentioning Clinton were considered. In doing so, the analysis inevitably had to exclude Twitter users who tweeted less frequently. The decision, however, ensured that only active Twitter users were included and increased the accuracy to predict the user's political orientation.

For each candidate-mentioning dataset, the aforementioned SVM models were used to predict the sentiment of tweets posted by the identified Twitter users. For users who tweeted three or more times, a random sample of three tweets were drawn and used to predict the user's political ideology. For each candidate, if a Twitter user expressed positive attitude in two out of three tweets, the user was categorized as the candidate's supporter. Conversely, if a user expressed negative attitude in two out of three tweets, the user was categorized as the candidate's critic. All other users were considered to hold a neutral or unknown attitude toward the candidate.

Once all identified Twitter users were categorized, we then used the formula in Colleoni et al. (2014) to calculate political homophily in each cluster, which is defined as the number of outbound ties directed to users who share political orientation (i.e., candidate supporter to candidate supporter, candidate critic to candidate critic), divided by the overall number of outbound ties. The homophily rate also ranges from 0 to 1, where 0 means the cluster is completely heterophilous and 1 means the cluster is completely homophilous. Only clusters with at least 10 pairs of categorized Twitter users were included in the analysis.

Data analysis and visualization

H1-2 asked about the difference between Trump and Clinton's Twitter communities in their level of homophily and homogeneity. Considering that the size of the Twitter community might be a confounding variable, a one-way ANCOVA was conducted to examine whether the two candidates' communities significantly varied controlling for the number of tweets in each community. A Chi-square test for independence was conducted to examine H3. H4-5 expected that types of opinion leaders predict the level of homophily and homogeneity in the Twitter communities of Clinton and Trump. Four blocks of variables were included in the OLS regression models: (1) number of nodes and opinion leaders in each cluster (control variables); (2) gender of opinion leaders; (3) five types of opinion leaders – politician, journalist, celebrity, emerging media, and non-media entity; (4) number of opinion leaders in each ideological camp. Predictors (e.g., traditional media, bloggers) that were found to have no significant correlation with the dependent variables in pre-tests were not included in the models. Each of the predictor blocks were entered into the model in the above sequence and their net contribution to the dependent variable's variance was measured.

Finally, networks about Trump and Clinton, respectively, sampled on 7 November 2016 (the day before the election) were visualized using NodeXL. In Trump's network, users who expressed positive attitude toward the candidate were color-coded as red, users who expressed negative attitude toward him were color-coded as blue, and all other users were color-coded as gray. Opposite colors were used to visualize Clinton's network: blue indicates positive, red indicates negative, and gray indicates neutral. Opinion leaders in each cluster are represented in a diamond shape and their Twitter handles are specified below.

Results

The study examined the network structure of Twitter communities about the two U.S. presidential candidates, Trump and Clinton, and how it related to opinion leadership. Overall, the results showed that Twitter communities about the two candidates differed significantly in a number of aspects and that certain individuals and organizations did affect the level of homogeneity of Twitter communities.

In testing H1, Twitter communities discussing Trump had a slightly higher level of political homophily ($M = .61$, $SD = .19$) than those discussing Clinton ($M = .59$, $SD = .21$), but the difference was not significant after controlling for the number of tweets in each community: $F\ (1, 391) = 2.398$, $p = .122$.

In addressing H2, the results showed that political views in the Twitter communities about Trump ($M = .25$, $SD = .30$) were significantly more heterogeneous than those about Clinton ($M = .10$, $SD = .21$): $F\ (1, 579) = 32.970$, $p < .01$. Although Twitter communities about both Trump and Clinton were rather homogeneous, those discussing Trump were more likely to be exposed to mixed attitudes; there is a higher level of both positive and negative viewpoints shared in the Twitter communities about Trump. Figures 1 and 2 visually present the networks about Trump and Clinton, respectively, sampled on 7 November 2016. As Figure 1 shows, Clinton's communities were homogenous in the sense that the majority of tweets expressed negative attitude toward her (i.e., dominated by red nodes). In contrast, Twitter communities discussing Trump (see Figure 2) included users that held both positive and negative viewpoints toward him.

RQ1 asked who the opinion leaders were in the discussion of Trump and Clinton, respectively. Table 2 presents opinion leaders that appeared most frequently in Twitter

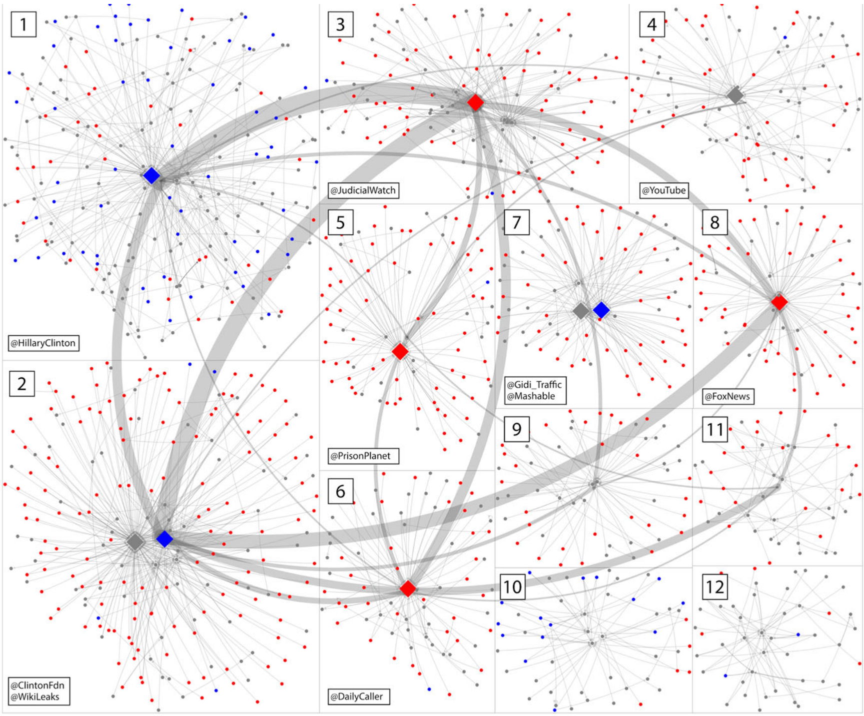

Figure 1. Twitter network about Clinton sampled on 7 November 2016.

conversations about the two candidates. Notably, many of Clinton's top opinion leaders were not liberal-oriented. For instance, @WikiLeaks and @YouTube – considered neutral – appeared as opinion leaders 11 and 8 times throughout Clinton's networks, respectively. Trump's top opinion leaders were far different in comparison. Aside from the candidate himself, @CNN and @Fahrenthold (a *Washington Post* reporter) were the most recurring opinion leaders, appearing 7 and 3 times throughout Trump's networks, respectively. It should be noted that @CNN was the only opinion leader that was influential in both Clinton and Trump's Twitter communities. Also noteworthy, many of Trump's opinion leaders recurred very infrequently. In fact, the majority of his opinion leaders only appeared in 1-2 networks total as demonstrated in Table 2. Clinton's opinion leaders, on the other hand, were more likely to maintain their influence over time, with many being present in 4 or more networks. This may suggest that Trump's Twitter networks were more dynamic, with prominent voices changing on a week-to-week basis.

H3 focused on the different types of opinion leaders between Twitter communities about Trump and Clinton. The results showed that influential individual users were more likely to lead the Twitter conversation about Trump, whereas entities (i.e., media or non-media organizations) were more likely to lead the conversation about Clinton: X^2 (1, $N = 246$) = 8.03, $p < .01$. For both entity and individual opinion leaders, Trump- and Clinton-mentioning communities varied significantly on political orientation of their opinion leaders: X^2 (2, $N = 246$) = 13.09, $p < .01$. Specifically, communities about Trump had significantly more liberal-oriented opinion leaders (54%) compared with communities about Clinton (33%).

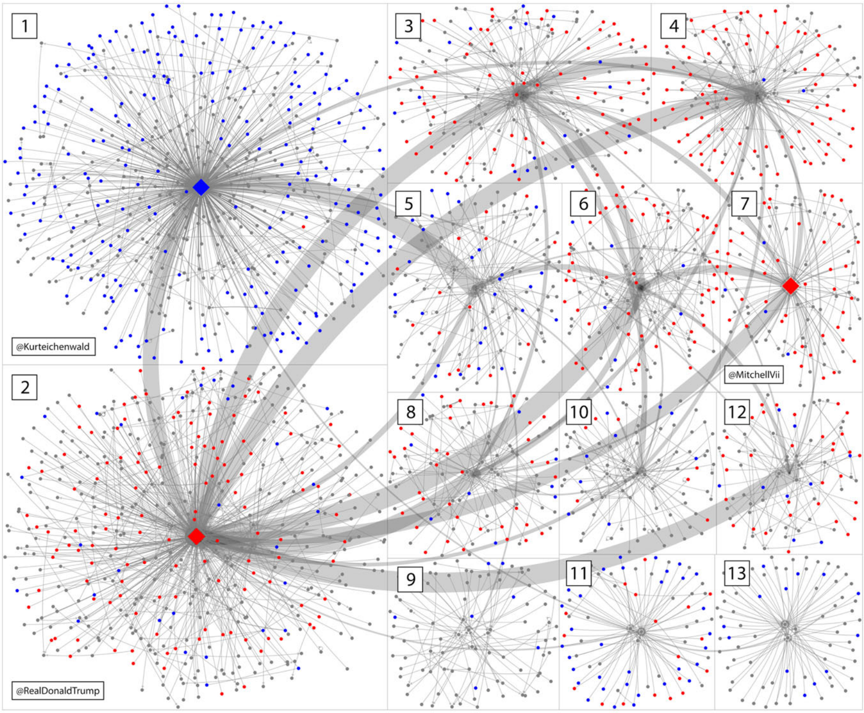

Figure 2. Twitter network about Trump sampled on 7 November 2016.

Table 2. Top 10 opinion leaders in the Twitter communities.

	Number of times present
Clinton communities	
Hillaryclinton	16
Wikileaks	11
Youtube	8
Foxnews	7
Loudobbs	7
Thehill	5
Cnn	4
Cnnpolitics	4
Judicialwatch	4
Prisonplanet	4
Trump communities	
Realdonaldtrump	17
Cnn	7
Fahrenthold	3
Danscavino	2
Ddale8	2
Donaldjtrumpjr	2
Joyannreid	2
Kurteichenwald	2
Lindasuhler	2
Nytimes	2

Note: Opinion leaders presented in terms of Twitter handles (@).

Among Twitter communities that had influential individuals as opinion leaders, the gender of opinion leaders differed: X^2 $(1, N = 140) = 4.91, p < .05$. Females were more likely to be opinion leaders in communities about Clinton (34%), whereas males were more likely to lead the conversation about Trump (83%). In terms of the type of individual opinion leaders, journalists were found to be the most common in Clinton's Twitter communities (45%), whereas politicians were the most common in Trump's communities (34%). However, this difference was not statistically significant.

In addition, no significant differences were found in the types of influential entity opinion leaders between Clinton- and Trump-mentioning communities. Descriptive analysis showed that traditional news media (e.g., @CNN, @FoxNews) was the most common type of entity opinion leader across both Clinton (48%) and Trump (75%) communities, whereas non-media organizations (e.g., @WikiLeaks, @ClintonFDN) and unknown entities were the least present. The presence of emerging media opinion leaders (e.g., @YouTube, @JudicialWatch, etc.) varied between Clinton (30%) and Trump (17%) communities. See Table 3 for the distribution of the types of opinion leaders in the Twitter communities.

Finally, H4-5 examined whether types of opinion leaders would affect the echo chamber effect in Twitter communities. It appeared that opinion leadership did not affect the level of political homophily, but did predict political heterogeneity of Twitter communities (see Table 4). As for Twitter communities about Trump, type of opinion leaders block explained 8.9% of the variance in the community's heterogeneity ($\Delta R^2 = .089$, $p < .01$). The political orientation of opinion leaders further contributed 6.8% of the model variance ($\Delta R^2 = .068$, $p < .01$). Specifically, conservative-oriented opinion leaders significantly increased the political heterogeneity of Twitter communities ($\beta = .510$, $p < .01$). The overall model, including all blocks in the analysis, accounted for 18.1% of the total variance in political heterogeneity of Trump's communities.

When it came to Twitter communities discussing Clinton, gender of opinion leaders explained 7.4% in predicting political heterogeneity ($\Delta R^2 = .074$, $p < .01$). In this block,

Table 3. Types of opinion leaders present in the Twitter communities.

	Clinton communities ($N = 164$)	Trump communities ($N = 82$)	χ^2
Type of opinion leader			8.03**
Individual	85 (52%)	58 (71%)	
Entity	79 (48%)	24 (29%)	
Partisanship			13.09**
Liberal	54 (33%)	44 (54%)	
Conservative	71 (43%)	31 (38%)	
Other/unknown	39 (24%)	7 (8%)	
Gender			7.03*
Male	54 (64%)	48 (83%)	
Female	28 (33%)	10 (17%)	
Other/unknown[a]	3 (3%)	0 (0%)	
Person-type users			2.23
Politicians	23 (27%)	20 (34%)	
Journalists	38 (45%)	18 (31%)	
Entertainment celebrities[a]	4 (5%)	7 (12%)	
Bloggers/average users	18 (21%)	13 (22%)	
Other/unknown[a]	2 (2%)	0 (0%)	
Entity-type users			3.08†
Traditional news media	38 (48%)	18 (75%)	
Online-only, emerging media	24 (30%)	4 (17%)	
Non-media organizations[a]	16 (20%)	2 (8%)	
Other/unknown[a]	1 (2%)	0 (0%)	

[a]Category excluded from Chi-square analysis due to low expected cell size.
†$p < .1$; *$p < .05$; **$p < .01$.

Table 4. Prediction of Twitter community heterogeneity.

	Clinton (N = 293)	Trump (N = 289)
Block 1: Network statistics		
Number of nodes	.073	−.022
Number of opinion leaders	.184	−.171
ΔR^2	.021*	.017
Block 2: Opinion leader gender		
Number of males	−.162	.151
Number of females	.030	.094
ΔR^2	.074**	.007
Block 3: Type of opinion leaders		
Number of politicians	.307**	−.186
Number of journalists	.062	−.152
Number of celebrities	.143*	−.059
Number of emerging media	−.121	−.049
Number of non-media	−.165*	.062
ΔR^2	.101**	.089**
Block 4: Opinion leader ideology		
Number of liberal opinion leaders	−.043	−.014
Number of conservative opinion leaders	−.192	.510**
ΔR^2	.013	.068**
Total model R^2	.208**	.181**

Note: Cells are final-entry OLS standardized Beta (β) coefficients.
*$p < .05$; **$p < .01$.

although individual variables did not exert significant influence, it did appear that female opinion leaders contributed to a more heterogeneous conversation about Clinton while male leaders were more likely to decrease the level of heterogeneity. The block of type of opinion leaders accounted for 10% of the variance in explaining heterogeneity ($\Delta R^2 = .101$, $p < .01$). The results showed that politicians ($\beta = .307$, $p < .01$) and entertainment celebrities ($\beta = .143$, $p < .05$) were positively correlated with political heterogeneity, whereas non-media organizations negatively predicted political heterogeneity ($\beta = -.165$, $p < .05$). The entire model explained 20.8% of the total variance of the level of heterogeneity in Clinton's communities.

Discussion

This study examined the echo chamber phenomenon and opinion leadership on Twitter during the 2016 U.S. presidential election. Overall, the results suggest that Twitter communities about Trump and Clinton differed significantly and that certain opinion leaders were responsible of creating homogeneous communities on Twitter. This study made a theoretical contribution to the literature by linking two important areas of network research and shedding light on what may have caused the echo chamber problem in an emerging mediascape. Specific findings and contributions of the study are discussed below.

First, it is interesting to find that Twitter communities discussing Trump showed a higher level of heterogeneity than those about Clinton. An additional analysis based on the Janis Fadner coefficient (1943) revealed that Twitter's conversations about Clinton ($M = -.49$, $SD = .32$) were significantly more negative than those about Trump ($M = -.27$, $SD = .25$) after controlling for the size of the community: F (1, $N = 579$) = 80.275, $p < 0.01$. Taken together, these results indicate that Twitter users were more likely to be exposed to diverse political sentiment, both positive and negative, toward Trump, whereas those discussing Clinton tended to be exposed to negative sentiment. As Figures 1 and 2 demonstrated, on the day before the election, Twitter communities about Trump carried a mixture of political views (i.e., red and blue), whereas those about Clinton were dominated by negative views (i.e., red). This is likely given that Clinton's email controversy and other 'scandals' unfolded near the election day – WikiLeaks and the Clinton Foundation were found to be two opinion leaders in one major Twitter community about Clinton (see Figure 1). The finding is also in line with other analyses that showed Trump better leveraged social media to both reach his audience and stimulate conversations (Andrews, 2016).

The role opinion leaders play in social media networks is worth more scholarly attention. The revised definition of the concept in a social media context and the influence opinion leaders have yielded in opinion diversity in their respective Twitter communities are two fruitful areas this paper contributes to. Specifically, we have suggested that the term 'opinion leader' should refer to any influential individual or institution on Twitter in that all kinds of Twitter users would compete with each other in setting the public's agenda. However, our results suggest that those with established reputation offline were still influential on Twitter. For Twitter communities about Clinton, the top 10 opinion leaders were mostly news organizations or media professionals. It is worthwhile to note that *WikiLeaks* was found to be an important opinion leader in Clinton's communities, only second to Clinton herself. This confirms Darwish, Magdy, and Zanouda's (2017) assessment that the global whistleblower on Twitter might play a key role in Clinton's failure in the 2016 election. It is equally notable to report that *Prison Planet*, a right-wing, conspiracy-driven online news website, was also quite influential in leading the discussion about Clinton on Twitter. The recurring influence of fake news websites was not found in the Twitter communities discussing Trump. This may indicate that fake news sites indeed had a large following on Twitter but were mainly responsible for spreading negative news about Clinton, a finding consistent with Allcott and Gentzkow's (2017) analysis. In addition, readers should use caution when interpreting the finding that YouTube also served as an opinion leader in Clinton's Twitter communities. Unlike other emerging media organizations such as *Huffington Post*, YouTube is a social networking platform that might host someone else's opinion. Future research could consider further identifying individuals or institutions who produced the YouTube videos and their cross-platform opinion leadership.

Unlike opinion leaders in Clinton's communities, those who led the discussions about Trump changed on a week-to-week basis. The most recurring ones included Trump himself and @CNN, which was also an influential opinion leader in Clinton's Twitter communities. This finding illustrates the agenda-setting power of CNN, a traditional media organization, in the Twittersphere. It also suggests that different stakeholders engaged in a competitive contest in framing the discussion of Trump, a controversial, non-conventional candidate. The relationship between personal characteristics of political candidates and the diversity of opinion leaders warrants more systematic investigation.

Also noteworthy, out of the top ten most recurring opinion leaders in Trump's discussion, six were from media organizations with a liberal ideology, and most of those were news reporters. The statistical analysis also showed that liberal-oriented opinion leaders were more influential in Twitter communities about Trump than Clinton. This is an important finding because it indicates that opinion leaders are not only active in leading the conversations about their 'own' political candidates but can be more influential in the opposite camp, which is another promising area of future research.

The most important contribution of this study is that it theoretically links the concept of opinion leadership and the echo chamber effect on social media, an area unexamined in the previous literature. Our analyses empirically demonstrated that types of opinion leaders did affect the level of political heterogeneity of Twitter communities. Specifically, we found that politicians and entertainment celebrities tended to trigger a more heterogeneous conversation about Clinton. However, when some organizations such as political groups served as the opinion leader, the conversation would become more homogeneous. For Twitter communities discussing Trump, conservative-oriented opinion leaders significantly increased the degree of political heterogeneity. Based on the findings, we contend that perhaps one way to tackle the echo chamber problem is to work with opinion leaders who are more likely to stimulate a dynamic conversation on social media.

Two hypotheses that were not supported are also worth attention. First, Twitter communities about Trump and Clinton did not differ significantly in terms of political homophily, but did in political homogeneity. Second, opinion leadership seemed to have no impact on the political homophily on Twitter, but did on political homogeneity. These findings reveal that political homophily and political homogeneity are two related, but different constructs. In this study, while certain individuals and institutions had the power to influence the diversity of tweet *content* within Twitter communities, they could not determine how Twitter *users* of different political orientations interacted with one another. Based on the finding, we suggest that the concept of echo chamber should be revisited with these two dimensions in future research.

To conclude, we argue that the current echo chamber critique is premature in that it neglects the nuances found in the online discussion networks. Our findings that Clinton's and Trump's Twitter networks differed significantly and that different opinion leaders made varied contributions to the network nature of Twitter communities indicate that a sweeping conclusion about all Twitter networks is not instrumental for a better understanding of political communication on a social media platform. Results from this study shed light on the nuances of Twitter networks and call for more discreet inspections on the involvement of opinion leaders in the echo chamber effect.

Methodologically, the study contributes to Twitter research by developing a systematic procedure to identify communities and opinion leaders on Twitter. Sampling Twitter networks every week for a four-month period prior to the election day also provides a unique longitudinal dataset for such an analysis. Future research might consider applying this sampling approach in other political or cultural contexts.

This study is limited in several aspects. First, the tweets we analyzed were collected via multiple approaches including the use of Twitter's public API. Thus, results from this study do not reflect the entire picture of the Twitter networks. Second, our calculation of political homophily required at least three tweets from each user, thus inevitably excluding users who were less active on Twitter from the final analysis. Within the active users, there is no guarantee the analysis was based on political expressions from actual human beings given the influence of social bots in online political discourse (e.g., Grimme, Preuss, Adam, & Trautmann, 2017). In fact, nearly one-fifth of all election-related tweets were sent by social bots (Bessi & Ferrara, 2016). The possibility of social bots participating in the candidates' networks, along with the finding of a fake news website as an opinion leader, point to the crucial role misinformation and disinformation play in facilitating the echo chamber phenomenon. Researchers could consider adding the investigation of fake news to the analysis of opinion leadership and the echo chamber phenomenon. Lastly, the candidate networks were built based on retweet and reply, without considering another relationship on Twitter: following. Although previous research found that networks based on interactions proved to be more influential in driving Twitter usage than simply following (Huberman, Romero, & Wu, 2008), analyzing following networks may shed light on user behaviors of lurkers on Twitter. In fact, research has shown that Twitter networks based on different types of relationships varied in their levels of homophily (Williams et al., 2015). Given that nearly half of registered Twitter users have never sent a tweet (Koh, 2014), who they follow and listen to warrants more academic attention. Despite these limitations, this study helps enrich our understanding of the role Twitter played in a controversial election and of the process of electoral discourse on this social platform overall. Not only academic scholars but also practitioners in political communication can benefit from our research findings.

Acknowledgement

The authors would like to thank Chris Vargo for providing the Twitter dataset (Dataset II) for this analysis.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Lei Guo (Ph.D., The University of Texas at Austin) is an assistant professor of emerging media studies at Boston University. She is also an affiliated faculty member at Boston University's Department of Computer Science, and a junior faculty fellow at the Hariri institute. Her research focuses on the development of media effects theories, computational social science methodologies, and emerging media and democracy in the United States and China. Dr. Guo's research has been published in a number of leading peer-reviewed journals such as *Journal of Communication*, *Communication Research*, and *New Media & Society*. Her co-edited book *The Power of Information Networks: New Directions for Agenda Setting* (2015) introduces a new theoretical perspective to understand media effects in this emerging media landscape [email: guolei@bu.edu].

Jacob A. Rohde earned his M.A. degree from the Division of Emerging Media Studies, Boston University, in 2016. He is currently a Ph.D. student at the School of Media and Journalism, The University of North Carolina at Chapel Hill [email: jarohde@live.unc.edu].

H. Denis Wu (Ph.D., The University of North Carolina at Chapel Hill) is a professor of mass communication at Boston University. His research focuses on international communication and political communication. He has won numerous research prizes, grants, and published widely in respected journals, including *Journal of Communication* and *Journalism & Mass Communication Quarterly*. Additionally, he has co-authored three books – *Media, Politics, and Asian Americans* (2009), *The Media are Ill: The Symptoms and Solutions of Taiwanese News Environment* (2013), and *Image and Emotion in Voter Decisions: The Affect Agenda* (2015). Dr. Wu was the head of the Mass Communication and Society division of the Association for Education in Journalism and Mass Communication (AEJMC). Currently, he sits on the editorial boards *of Journalism & Mass Communication Quarterly*; *Journalism and Communication Monographs*; and *Journal of Political Marketing* [email: hdw@bu.edu].

References

Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *The Journal of Economic Perspectives*, *31*(2), 211–236. doi:10.3386/w23089

Andrews, N. (2016, November 10). How some social media data pointed to a Donald Trump win. *The Wall Street Journal*.

Bessi, A., & Ferrara, E. (2016). Social bots distort the 2016 us presidential election online discussion. *First Monday*, *21*(11). doi:10.5210/fm.v21i11.7090

boyd, d., & Ellison, N. B. (2007). Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, *13*(1), 210–230. doi:10.1111/j.1083-6101.2007.00393.x

Brosius, H.-B., & Weimann, G. (1996). Who sets the agenda? Agenda-setting as a two-step flow. *Communication Research*, *23*(5), 561–580. doi:10.1177/009365096023005002

Bruns, A. (2012). How long is a tweet? Mapping dynamic conversation networks on Twitter using Gawk and Gephi. *Information, Communication & Society*, *15*(9), 1323–1351. doi:10.1080/1369118X.2011.635214

Bump, P. (2017, January 19). America's deep partisan rift may be setting Donald Trump up for success in advance. *The Washington Post*.

Chang, D., & Ghim, G. (2011). The structure and dynamics of the Korean Twitter network. *Journal of Communication Research*, *48*(1), 59–86. doi:10.22174/jcr.2011.48.1.59

Colleoni, E., Rozza, A., & Arvidsson, A. (2014). Echo chamber or public sphere? Predicting political orientation and measuring political homophily in Twitter using big data. *Journal of Communication*, *64*(2), 317–332. doi:10.1111/jcom.12084

Collingwood, L., & Wilkerson, J. (2012). Tradeoffs in accuracy and efficiency in supervised learning methods. *Journal of Information Technology & Politics*, *9*(3), 298–318. doi:10.1080/19331681.2012.669191

Coyne, B. (2016, November 7). *How #Election2016 was Tweeted so far*. Retrieved from https://blog.twitter.com/2016/how-election2016-was-tweeted-so-far

Darwish, K., Magdy, W., & Zanouda, T. (2017). Trump vs. Hillary: What went viral during the 2016 US presidential election. In *International conference on social informatics* (pp. 143–161). Cham: Springer.

Dubois, E., & Gaffney, D. (2014). The multiple facets of influence identifying political influentials and opinion leaders on twitter. *American Behavioral Scientist*, *58*(10), 1260–1277. doi:10.1177/0002764214527088

Festinger, L. (1962). *A theory of cognitive dissonance*. Stanford, California: Stanford University Press.

Freelon, D., Lynch, M., & Aday, S. (2015). Online fragmentation in wartime: A longitudinal analysis of tweets about Syria, 2011–2013. *The ANNALS of the American Academy of Political and Social Science*, *659*(1), 166–179. doi:10.1177/0002716214563921

Grimme, C., Preuss, M., Adam, L., & Trautmann, H. (2017). Social bots: Human-like by means of human control? *Big Data*, *5*(4), 279–293.

Gruzd, A., Wellman, B., & Takhteyev, Y. (2011). Imagining Twitter as an imagined community. *American Behavioral Scientist*, *55*(10), 1294–1318. doi:10.1177/000276421140937

Hansen, D., Shneiderman, B., & Smith, M. A. (2010). *Analyzing social media networks with NodeXL: Insights from a connected world*. Burlington, MA: Morgan Kaufmann.

Himelboim, I., McCreery, S., & Smith, M. (2013). Birds of a feather tweet together: Integrating network and content analyses to examine cross-ideology exposure on Twitter. *Journal of Computer-Mediated Communication*, *18*(2), 40–60. doi:10.1111/jcc4.12001

Huberman, B. A., Romero, D. M., & Wu, F. (2008). Social networks that matter: Twitter under the microscope. ArXiv preprint arXiv:0812.1045. Retrieved from https://arxiv.org/pdf/0812.1045.pdf

Janis, I. L., & Fadner, R. H. (1943). A coefficient of imbalance for content analysis. *Psychometrika*, *8*(2), 105–119. doi:10.1007/BF02288695

Koh, Y. (2014). Report: 44% of Twitter accounts have never sent a Tweet. *The Wall Street Journal*. Retrieved from https://blogs.wsj.com/digits/2014/04/11/new-data-quantifies-dearth-of-tweeters-on-twitter/?mod=WSJBlog&utm_source=twitterfeed&utm_medium=twitter

Kwak, H., Lee, C., Park, H., & Moon, S. (2010). *What is Twitter, a social network or a news media?* Proceedings of the 19th international conference on World Wide Web (pp. 591–600). doi:10.1145/1772690.1772751.

Lazarsfeld, P. F., Berelson, B., & Gaudet, H. (1948). *The people's choice*. New York, NY: Columbia University Press.

McKelvey, K., DiGrazia, J., & Rojas, F. (2014). Twitter publics: How online political communities signaled electoral outcomes in the 2010 US house election. *Information, Communication & Society*, *17*(4), 436–450. doi:10.1080/1369118X.2014.892149

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, *27*(1), 415–444. doi:10.1146/annurev.soc.27.1.415

Nisbet, M. C., & Kotcher, J. E. (2009). A two-step flow of influence? Opinion-leader campaigns on climate change. *Science Communication*, *30*(3), 328–354. doi:10.1177/1075547008328797

Ognyanova, K., & Monge, P. (2013). A multitheoretical, multilevel, multidimensional network model of the media system: Production, content, and audiences. *Annals of the International Communication Association*, *37*(1), 67–93. doi:10.1080/23808985.2013.11679146

Park, C. S. (2013). Does Twitter motivate involvement in politics? Tweeting, opinion leadership, and political engagement. *Computers in Human Behavior*, *29*(4), 1641–1648. doi:10.1016/j.chb.2013.01.044

Rogers, E. M. (1983). *Diffusion of innovations*. New York: Free Press.

Tremayne, M., & Minnie, M. (2013). Opinion leadership on gun control in social networks: Preferential attachment versus reciprocal linking. *American Communication Journal*, *15*(4), 34–52.

Valente, T. W., & Pumpuang, P. (2007). Identifying opinion leaders to promote behavior change. *Health Education & Behavior*, *34*(6), 881–896. doi:10.1177/1090198106297855

Vargo, C., Guo, L., McCombs, M., & Shaw, D. L. (2014). Network issue agendas on Twitter during the 2012 US presidential election. *Journal of Communication*, *64*(2), 296–316. doi:10.1111/jcom.12089

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge: Cambridge University Press.

Watts, D. J., & Dodds, P. S. (2007). Influentials, networks, and public opinion formation. *Journal of Consumer Research*, *34*(4), 441–458. doi:10.1086/518527

Weimann, G. (1982). On the importance of marginality: One more step into the two-step flow of communication. *American Sociological Review*, *47*(6), 764–773.

Williams, H. T., McMurray, J. R., Kurz, T., & Lambert, F. H. (2015). Network analysis reveals open forums and echo chambers in social media discussions of climate change. *Global Environmental Change*, *32*, 126–138. doi:10.1016/j.gloenvcha.2015.03.006

Xu, W. W., Sang, Y., Blasiola, S., & Park, H. W. (2014). Predicting opinion leaders in Twitter activism networks: The case of the Wisconsin recall election. *American Behavioral Scientist*, *58*(10), 1278–1293. doi:10.1177/0002764214527091