Article

new media & society
1–22

DOI: 10.1177/1461444816634054
nms.sagepub.com

Political rumoring on Twitter during the 2012 US presidential election: Rumor diffusion and correction

Jieun Shin and Lian Jian
University of Southern California, USA

Kevin Driscoll
Microsoft Research, USA

François Bar
University of Southern California, USA

Abstract
Social media can be a double-edged sword for political misinformation, either a conduit propagating false rumors through a large population or an effective tool to challenge misinformation. To understand this phenomenon, we tracked a comprehensive collection of political rumors on Twitter during the 2012 US presidential election campaign, analyzing a large set of rumor tweets ($n = 330,538$). We found that Twitter helped rumor spreaders circulate false information within homophilous follower networks, but seldom functioned as a self-correcting marketplace of ideas. Rumor spreaders formed strong partisan structures in which core groups of users selectively transmitted negative rumors about opposing candidates. Yet, rumor rejecters neither formed a sizable community nor exhibited a partisan structure. While in general rumors resisted debunking by professional fact-checking sites (e.g. Snopes), this was less true of rumors originating with satirical sources.

Keywords
Elections, political communication, partisan homophily, rumor, satire, social media

Corresponding author:
Jieun Shin, Annenberg School for Communication & Journalism, University of Southern California, 3502 Watt Way, Los Angeles, CA 90089, USA.
Email: jieunshi@usc.edu

Introduction

In recent years, politically interested publics have increasingly used social media sites as a platform to consume, share, and discuss relevant information with others (Bruns and Burgess, 2011; Rainie and Smith, 2012). As social media provides unprecedented access to diverse information sources, scholars and pundits (Kim, 2011; Papacharissi, 2009) have discussed social media's enabling role in increasing exposure to divergent political views and facilitating cross-cutting discussion. However, despite the potential, a growing body of research suggests that social media often functions as an echo-chamber that reinforces individuals' pre-existing attitude. Users tend to transmit ideologically congenial information and associate with like-minded others, at least, in the political context (Barbera et al., 2015; Conover et al., 2011; Himelboim et al., 2013; Jacobson et al., 2015).

The current political landscape of social media has particularly negative implications for the spread of misinformation. Experiments have shown that high segregation and clustering within a community increase the polarization of opinions in rumors as well as belief perseverance (DiFonzo et al., 2013, 2014). Therefore, it is critical to understand how political rumors diffuse and are debunked in a contemporary media environment where users can easily filter and choose their information sources. Despite the real consequences of rumors on electoral decisions (Weeks and Garrett, 2014) and the vast amount of misinformation circulating on the web (World Economic Forum, 2014), there has been little research on the political rumoring phenomenon that takes place in social media. Previous studies on political rumors (Cacciatore et al., 2014; Nyhan and Reifler, 2010; Weeks and Garrett, 2014) mostly focused on ascertaining the psychological mechanisms of rumor acceptance and correction at the individual level, rather than the media systems within which such rumors circulate. Exceptionally, Garrett (2011) and Rojecki and Meraz (2014) contextualized political rumors in the new media environment and provided some empirical evidence that communication technologies (e.g. websites and emails) can facilitate rumor diffusion.

In this study, we investigate diffusion patterns of rumors and debunking messages on Twitter by tracking a comprehensive collection of political rumor tweets during the 2012 US presidential election. While political discourse on Twitter typically revolves around a core of highly motivated partisan individuals and political elites (e.g. political bloggers) who use the platform as a dissemination tool (Bekafigo and McBride, 2013), the US general election seemed to attract even broader participation. During the 2012 US election season, major political events repeatedly broke Twitter's "most heavily tweeted" records (McKinney et al., 2014).

While taking an exploratory approach in this article, we focus on three research questions. First, we explore the underlying network structures of both political rumor-tellers and debunkers to ascertain the extent of polarization within these communities. Second, we examine the unique communication practices of rumor spreaders to understand the social relationships they form, focusing on three Twitter features: *retweet, hashtag*, and *mention*. Last, we investigate how resilient rumors are in the face of corrections by professional fact-checking organizations. Although we cannot directly test the effectiveness of rumor rebuttals on people's perception due to limitations of observational data, we explore whether major fact-checking sites contributed to slowing down the spread of debunked rumors.

To this end, we content-coded and analyzed a large set of tweets (n=330,538) containing 57 political rumors that circulated during the 15 months preceding the 2012 presidential election in the United States. Our data collection included rumors mostly targeting leading presidential candidates, such as one suggesting that Mitt Romney's campaign slogan was identical to Ku Klux Klan's (KKK's), or another asserting that Barack Obama's birth records were sealed.

To the best of our knowledge, this is the first large-scale analysis of content-analyzed political rumoring on a social network site. Our analyses revealed that rumor spreaders were clearly divided based on the target of the rumor: anti-Obama and anti-Romney. However, we did not find such partisan community structures among rumor debunkers. We also found that rumors were mainly shared with the spreader's own followers rather than interest-oriented publics organized around hashtags. Last, our analysis showed rumors were relatively resilient to debunking and continued to propagate despite the emergence of countervailing information. Yet, rumors originating with satirical websites were more vulnerable to debunking than others.

Literature review

Definition of rumor and political dimension

Rumor is commonly defined as unverified information that contains instrumentally valuable statements or public concerns for a certain group (DiFonzo and Bordia, 2007). They circulate "through the grapevine" via unofficial communication channels, as they are not publicly confirmed by an official source (Allport and Postman, 1945, 1947). Therefore, rumor contrasts with news, which is often reported by authorized outlets or media. Rumor tends to involve public matters, whereas gossip pertains to private relationships (DiFonzo and Bordia, 2007). Although the precise definition of rumor varies among scholars, rumors are generally defined by a lack of veracity.

Pioneered by Allport and Postman (1945), rumor scholarship has a long history and produced a substantial body of research. Although insightful, the existing rumor literature is limited in two respects. First, most prior rumor research did not specifically investigate political rumors. Rather, it jointly examined different types of rumor—such as crime or commercial rumors. Therefore, it was hard to discern whether the unique diffusion pattern of a political rumor can be attributed to it being a rumor or a political statement. Political rumors have distinctive properties, which set them apart from other rumors. For instance, Rojecki and Meraz (2014) coined the term "FIBs" (Factitious Information Blends) to describe politically motivated rumors. They argue that, unlike rumors, FIBs are deliberately planted and spread to discredit opposing politicians and parties. Therefore, for political rumors, "verification is less important than the coherence and integrity of claims sustained within a particular information environment" (p. 4).

Second, previous rumor research largely ignores the larger social and communication context in which rumor-tellers are embedded and instead focuses on individual differences in adopting rumors. Shibutani (1966) criticized rumor researchers early on for not grounding their work in the social context in which rumors arise. Therefore, examining political rumors within a social networking site fills a gap in the current literature.

Viewing the site as a social system, researchers can unobtrusively explore political rumoring in a social community that may influence the speed or scope of the rumor spread. Furthermore, by examining what technological features rumoring uses, researchers can better infer rumor-tellers' motivation and intended audience.

Theory of homophily and rumor membership

Homophily theory holds that "birds of a feather flock together" as people form connections with others who share their interests and characteristics, such as political attitude (McPherson et al., 2001). One mechanism driving homophily is individuals' desire to avoid cognitive dissonance (Festinger, 1957). To avoid psychological discomfort from situations that challenge their beliefs, people selectively choose interaction partners and information based on similarity (Festinger, 1957). This argument also aligns with the selective exposure thesis, which proposes that people seek out attitude-consistent messages and avoid conflicting messages to reduce dissonance.

Although some studies contested selective exposure, arguing that people may not systematically avoid attitude-challenging messages (Garrett, 2009a, 2009b; Knobloch-Westerwick and Kleinman, 2012), research findings agree that individuals generally prefer congenial information over discordant information (Knobloch-Westerwick et al., 2015). For instance, US political bloggers were found to be more likely to include hyperlinks to information sources on their side of the political spectrum than on the other side (Adamic and Glance, 2005; Jacobson et al., 2015). Twitter users also retweet and follow those with similar political attitudes more than those having conflicting attitudes (Boutet et al., 2012; Colleoni et al., 2014; Conover et al., 2011).

Given the partisan nature of political rumors (Rojecki and Meraz, 2014) and the new "pull" media environment, rumor-sharing patterns may show an extreme case of political homophily. Ease of online information search and low cost of online organizing allow individuals with obscure interests to easily form communities based on common identity (Benkler, 2006; Sunstein, 2009). Homophily can also manifest indirectly through individuals' media consumption patterns whereby like-minded people orbit around the same information source or agenda (Farrell, 2012). Thus, a rumor target may serve as a divisive topic that pulls different population segments apart, demarcating social boundaries between in-group and out-group. Therefore, we explore whether political rumors show homogeneous memberships along party lines, expecting those who believe anti-Romney rumors not to spread anti-Obama rumors and vice versa. We apply the same logic to rumor rejecters and examine membership polarization among those who debunk rumors. We explore the following research questions:

RQ1. Are believers and rejecters polarized based on the target of rumors?

Rumor diffusion patterns on Twitter

On Twitter, users associate with one another through technical features such as *retweet, hashtag*, and *mention*. Communication networks defined through these architectural features exhibit varying degrees of homophily (boyd et al., 2010; Kwak et al., 2010; Larsson and Moe, 2011). For instance, *retweet al*lows users to pass along existing information—often from sources they themselves follow—to their own followers. Prior research has shown *retweet* networks to be highly homophilous because retweeting often indicates endorsement or friendship toward the original author (Bouet et al., 2011; Conover et al., 2011). Furthermore, the primary audience for retweets seems to be one's followers, who many Twitter users perceive as their supporters or "fan base" (Marwick and boyd, 2011: 7).

By contrast, *hashtag* helps Twitter users address an audience beyond their followers (Bruns and Burgess, 2011). Using keywords prefixed with a hashtag symbol (#), users carve out a channel-like environment where people sharing similar interests gather and collectively streamline information. Some long-lasting hashtags reflect interest-based homophily, like #p2 (Progressives 2.0) and #tcot (Top Conservatives On Twitter). Hashtags are public, easy to search and accessible to the entire Twitter sphere, allowing users who do not emotionally belong to a particular community to follow or participate in a hashtag thread (Bode et al., 2015; Smith et al., 2014). Bode et al. (2015) found that some users intentionally use hashtags associated with an out-group to provoke the community. Therefore, while hashtags may be aimed at a community of interest, they potentially include users from across the political spectrum.

Last, *mentions* link a message to a particular user, adding "@username" to a tweet. Depending on a user's intention, homophily patterns in mention networks can vary widely. For instance, Bruns and Highfield (2013) found that among Australian Twitter users, while the politician network showed a strong partisan structure as mentions remained mainly within party lines, the ordinary citizen network did not show partisan homophily, as citizens tended to mention both "political friends and enemies alike." Other studies have also consistently observed a lack of political homophily patterns in mention networks (Conover et al., 2011: Williams et al., 2015).

By examining these communication patterns, we seek to understand the type of social relationships rumors primarily rely on. Depending on how users utilize Twitter, the platform can function as a public sphere where users actively interact with each other across political lines or as an echo-chamber in which users merely circulate content consistent with their views (Colleoni et al., 2014; Kwak et al., 2010). Therefore, we examine three modes of rumor communication (retweet, hashtag, and mention) and compare them with non-rumor political discourse, to understand how the platform was used for political rumor activity:

> *RQ2*. What types of social relationships do rumor-tellers exhibit in their tweets? Are they geared toward one's follower network, a broader community of interest, or a specific user?

Rumor's responses to verification

Rumor control and misinformation management have been examined by a few scholars focusing on the effects of rumor debunking, with mixed results. In particular, research finds rebuttals to political rumors produce effects ranging from "boomerang" or "backfire" (Byrne and Hart, 2009; Nyhan and Reifler, 2010) to "uniformly effective" (Weeks and Garrett, 2014). For instance, an experiment by Nyhan and Reifler (2010) demonstrated that debunking backfired, strengthening participants' belief in a false claim when they encounter a correction inconsistent with their view. By contrast, other experimental (Weeks, 2015; Weeks and Garrett, 2014) and survey (Garrett, 2011) studies reported that rebuttal of false rumors increased the accuracy of the claims without backfiring. Weeks and Garrett (2014) argued that "exposure to either strong or repeated rumor rebuttals can temper the likelihood that a backfire effect will occur" (p. 5).

Relatively few studies empirically explore corrections of rumor or misinformation on the web where various social factors interplay. Using survey data collected from individuals living in the United States, one study (Garrett, 2011) found that getting political information online was associated with more exposure to both rumors and their rebuttals. While the sources of misinformation have increased online, the channels that debunk rumors have also become increasingly available. Another survey-based study (Gottfried et al., 2013) showed positive effects of political fact-checking sites on the accurate perception of the presidential candidates.

Although the topic was not political, two recent studies found that rumor debunking had limited or no effect on social networking platforms (Friggeri et al., 2014; Hannak et al., 2014). Using a large sample of rumor posts on Facebook, Friggeri et al. (2014) found that rumor debunking had a transient effect on the number of reshares (i.e. reposts) but in the long term it had minimal effects on the volume of reshares of the original rumor post. With Twitter data, Hannak et al. (2014) observed that most rumor corrections occur among strangers, but these tweets tend to be ignored compared to friends' corrections.

Since previous studies either relied on self-report measures or used non-political rumors as data, more research is needed to understand the role of corrections in political rumor diffusion dynamics. Here, we investigate whether professional fact-checking organizations help slow down rumor propagation and accelerate rumor debunking. Although we cannot directly test the effects of their corrections on people's attitude change, we explore whether professional debunking affects subsequent rumor diffusion. We analyze qualitative and quantitative changes in rumor diffusion after rumor debunking and ask:

> *RQ3*. How do rumors respond to corrections issued by professional rumor debunking websites?

Method

Data collection

This project studies political rumor diffusion on Twitter, a micro-blogging platform, by analyzing a large collection of political tweets (n=298,894,327) collected during the 15-month period (October 2011–December 2012) leading up to the 2012 presidential election in the United States. Tweets in the dataset were collected in real-time, using the Gnip PowerTrack service to extract a subset of the comprehensive Twitter "firehose" matching a custom set of filtering rules.[1] The rules used in this research, first compiled in October 2011, included 208 keywords and phrases related to the election, such as names of likely candidates and issue-specific terminology. Rules were updated manually based on emerging discursive trends during the election cycle, for example, new campaign themes or candidates entering the race. By Election Day, 6 November 2012, the rules included 427 unique keywords and phrases. Records in our dataset included the content of the tweet, metadata about the post and the author, and a snapshot of the author's profile information at the time the message was sent. Data collection ran continuously for 13 months.

The analysis was organized around all the political rumors circulating in the traditional news media or on social media sites during the same period of time. In May 2013, we identified 82 rumors investigated by at least one of the three main rumor checking websites—Factcheck.org, Snopes.com, and About.com's "Urban Legends" page—during 2012. We eliminated image-based rumors as well as rumors that started circulating on Twitter before our data collection period.[2] Our final collection contained 57 rumors, such as a false claim about Obama wearing an earpiece that fed him answers during the third presidential debate and a conspiracy theory about Romney's son owning a company that manufactured voting machines. A detailed description of all rumors is included in Appendix 1.

Constructing the rumor dataset

We retrieved all tweets from the political tweet set that contained matching keywords ($n=439{,}556$) for each of these 57 rumors. The retrieval task involved three steps. First, we pre-processed the entire Twitter corpus ($n=298{,}894{,}327$) by converting all characters into lower case, extracting words, and reducing each word to its stem. Second, based on the description of each rumor offered by the rumor debunking sites, we assembled an initial set of keywords with a goal of identifying as many rumor-related tweets as possible, at the risk of making some false-positive match. Last, we refined our queries by combining keywords into logical expressions, repeatedly testing them on the actual dataset, and having two authors manually inspect a random sample of the retrieved results. For example, the query "romney" AND ("america" OR "american") AND ("kkk" or "klan") accurately retrieved tweets related to a rumor (debunked by Snopes) that Romney's campaign slogan was identical to a catchphrase once used by the KKK.[3]

We took an additional measure by human-coding each tweet for two variables to ensure high accuracy. The first indicated whether the message was actually about the rumor, to eliminate tweets that happened to match our keywords but were not rumor-relevant. If a tweet was confirmed as rumor-relevant, we subsequently classified its author's attitude as endorsing, rejecting, or unclear. Tweets repeating or confirming a rumor were coded as "endorsing." Tweets denying the rumor or citing those who debunked the rumor were coded as "rejecting." All other tweets were coded as "unclear." For example, "That awkward moment when Mitt Romney's campaign slogan was the slogan for the KKK used in 1922" was coded "endorsing"; "Stop saying Romney used the KKK slogan 'Keep America American'. He didn't even say that. He said 'Keep America, America'." was coded "rejecting"; and "Did Romney adopt KKK slogan?" was coded "unclear."

Due to a high proportion of retweets, 59% of the retrieved tweets were duplicates.[4] Therefore, we hand-coded only unique tweets (n = 180,300). This extensive content coding lasted 11 months and involved four pairs of independent undergraduate coders, with two coders coding the same sets of messages. The first author (J.S.) trained each coding team to understand the rumor's context and to evaluate tweets for the two variables. During the initial training period for each rumor, coders practiced the protocol with a subset of the retrieved tweets independently. When coders produced sufficiently reliable data, they were given a batch of tweets containing the rumor for which they had been trained. To reduce coders' fatigue, we asked coders to work no more than 10 hours/week and no more than 200 tweets/hour. We also encouraged coders to code every day rather than to code for 1 or 2 days a week. Coders worked on one rumor at a time, and the order of rumors was randomly determined.

Inter-coder reliability was measured by Krippendorff's alpha. For the rumors that were coded over an extended period of time (longer than 2 weeks), inter-coder reliability was checked at multiple points in time. In general, reliability stayed above .75 and increased as the coders gained more experience.[5] For each rumor, remaining disagreements were resolved by randomly adopting one of the values assigned by two coders. Following Riffe et al.'s (1998) recommendation, some of the large rumors (n = 48,442) were partially single coded. More specifically, for the rumors (n = 9) which had over 10,000 tweets, we randomly selected 1–5% of the tweets for each rumor and asked two coders to code them. The rest were single coded only when intercoder-reliability (alpha > .8) was high. According to this rule, three rumors were partially singled coded.

In addition to coding tweets, we coded each of the 57 rumors as "True," "False," and "Unverifiable," based on the analysis of aforementioned fact-checking websites: typically, fact-checking sites report a rumor's accuracy using descriptions such as "Fake," "No. There is no evidence," and "Yes, True." We coded rumors as "false" if any of the sites described it to be "unlikely" or "not true."[6] Out of the 57 rumors, 11 rumors (19.30%) were true, while 44 rumors (77.19%) were false. True rumors included a 1991 promotional booklet from Obama's literary agency listing Obama as having been born in Kenya. False rumors included suspicion that Obama's wedding ring bears the inscription of "There is no god but Allah." Two rumors (3.51%) were classified as unverifiable by the fact-checking sites—one claiming the caps for Romney's campaign were made in China, the other that John F. Kennedy's daughter Caroline called Obama a liar.

The majority of rumors (n = 51) were about the leading presidential candidates for each major political party—38 rumors about Barack Obama and 10 about Mitt Romney. Three of the rest referred to both candidates such as a rumor that Tim Tebow, an American football quarterback, publicly predicted a Romney victory over Obama. Although our dataset included four times more Obama-related rumors than Romney-related rumors, Romney-related rumor tweets (n = 187,791) overall exceeded Obama-related rumor tweets (n = 108,376). The other rumor targets included Rick Santorum (Republican primary candidate), Paul Ryan (vice presidential candidate), and George Bush (former President).

Of all the tweets preliminarily identified as relevant to these 57 rumors (n = 439,556) via keyword matching, our coders found 75.20% (n = 330,538) were relevant, showing a relatively high precision in our initial keyword search. Additionally, all 57 rumors, but one, were negative about their target. The sole exception, a positive and true rumor about Romney, claimed he helped a Bain Capital colleague locate his missing teenage daughter.

Results

Overall, there were few tweets rejecting any rumor in our dataset, be it true or false. For each rumor, we calculated the percentage of tweets coded as endorsing or rejecting. Rumors with fewer than 200 tweets ($n=12$) were excluded from this analysis to eliminate the influence of marginal rumors on the mean rate of rejection and endorsing. Out of 33 false rumors, we found an average rejection rate of 3.37%, while we observed an average rejection rate of 0.06% for the 10 true rumors. The two unverifiable rumors had an average rejection rate of 8.43% and endorsing rate of 89.37%.

Membership overlap between rumors

To explore community structure and membership patterns among rumors, we view the relationship between a rumor and its tellers as a form of affiliation. In Social Network Analysis (SNA), affiliation networks are two-mode networks that depict the relationship between two distinct sets of nodes (Wasserman and Faust, 1994). In our dataset, the first mode was a set of Twitter users, and the second mode was a set of rumors these users have tweeted about. In order to obtain user data (the first mode), we identified the authors of all tweets for each rumor in our dataset. For instance, the rumor about Romney's campaign slogan included 94,623 endorsing and 535 rejecting tweets, corresponding to 91,675 unique believers and 525 unique rejecters.

For this analysis we focused only on false rumors, as there were not enough true rumors, and true rumors were seldom rejected. We repeated this process for each false anti-Obama ($n=21$) and false anti-Romney rumors ($n=7$) that involved more than 200 users—a threshold meant to exclude rumors of small size.[7] We then constructed 2 two-mode network matrices (i.e. 28 rumors by 176,149 users for the believer matrix; 27 rumors by 4688 users for the rejecter matrix) where users occupy the rows, and rumors the columns.[8] Cells in each of these two-mode matrices were set to 1 if a user tweeted about the rumor, 0 otherwise.

To examine user overlap across rumors, we transformed the two-mode affiliation matrices (both believer and rejecter matrix) into one-mode rumor-by-rumor matrices. This transformation is commonly used to understand the relationships among one set of nodes (rumors) connected via the other set of nodes (tweet authors). By reducing two-mode data to a one-mode (rumor-to-rumor), we were able to infer the extent to which two rumors attracted similar people. In the converted matrix, each cell contains the number of users shared by two rumors.

To identify sub-clusters in the rumor co-membership network, we used the Infomap algorithm, one of the best-performing community detection methods (Rosvall and Bergstrom, 2008), which can efficiently handle weighted edge values (here, the number of shared members), using the concept of random walks on the network. Figure 1 shows the two communities (modularity = 0.48) of rumor believers, corresponding exactly to

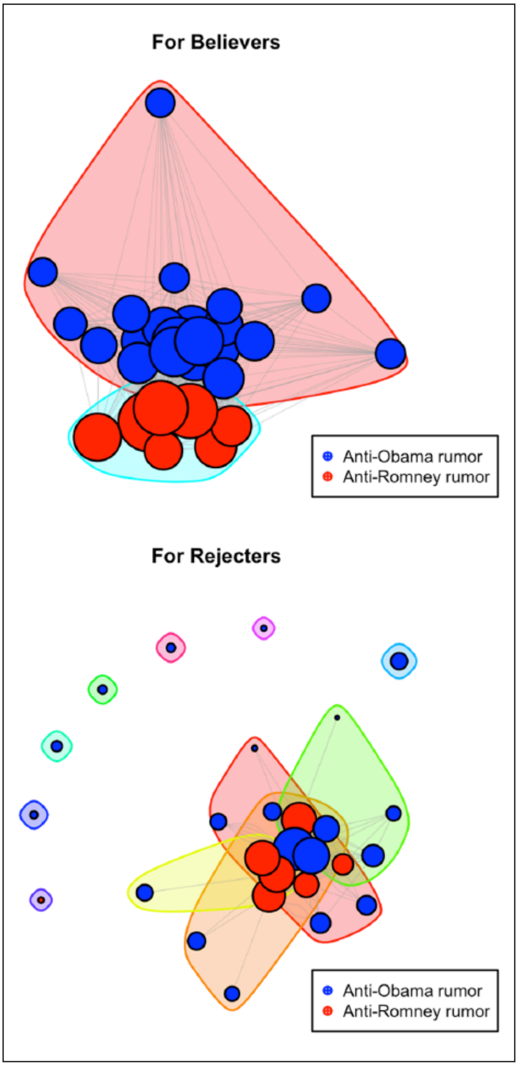


Figure 1. Community structure among anti-Obama and anti-Romney Rumors.
The shaded areas represent communities identified by Infomap algorithm. In the believer community, nodes with the same color correspond to the identified two communities. However, color of nodes does not function as a predictor of community among rejecters. Node size is proportional to the number of believers and rejecters in each rumor.

anti-Obama and anti-Romney rumors, and 13 communities (modularity = 0.57) of rumor rejecters with mixtures of anti-Obama and anti-Romney rumors.

In addition, we compared the density of the two believer communities, to examine the strength of connections among rumors. The density measure is defined by the sum of existing ties divided by the number of possible ties, which reflects the cohesiveness of the group (Wasserman and Faust, 1994). The analysis revealed that the network density of anti-Obama rumors was much higher (0.58) than that of anti-Romney rumors (0.06).

Rumor Tweeters' communicative behavior

To investigate the underlying social relationships among the rumor spreaders, we examined systematic differences between rumor and non-rumor tweets by comparing rumor tweets that endorsed 57 rumors with a random sample of the same size from our larger collection of political tweets (Figure 2). Rumor tweets (n=315,647) contained a much higher proportion of retweets (65.48%) than non-rumor tweets (44.43%). The proportion of retweets, across both false and true rumors, seems unusually high compared with previous studies (Nagarajan et al., 2010; Tonkin et al., 2012) which reported 27–48% of retweets in social movement discussions on Twitter.[10] A closer look at the data revealed that the retweet ratio was particularly high among the subset of tweets pertaining to false rumors (67.24%).

Additionally, we compared hashtag adoption ratios of the rumor set and random non-rumor set on original tweets (n=108,951). We excluded retweets for this analysis to examine users' own expression rather than the original authors'. A much lower proportion of rumor tweets included a hashtag (19.79%) compared to non-rumor tweets (30.74%). This pattern was relatively consistent across our combined set (19.79%), false rumor set (19.48%), and true rumor set (20.55%). We did not find any differences in the nature of hashtags between the rumor and non-rumor tweets. The most frequently used hashtags in both sets included "#tcot," "#p2," "#obama," "#gop," and "#romney."

Last, similar to the hashtag analysis, we compared ratios of tweets that mentioned a user (i.e. @user) between the original rumor tweets and the original non-rumor tweets (n=108,951). The two sets were not significantly different in terms of ratio of mentions (34.97% for rumor tweets and 31.43% for non-rumor tweets). Proportions of mentions were similar across the false rumor set (34.68%) and true rumor set (35.02%).

Changes in rumor trends after debunking

To examine changes in the rumor propagation trends upon fact-checking, we identified the date on which each rumor was first debunked by one of three fact-checking websites. We split each rumor tweet set into two groups—before and after the date of rumor debunking. Comparing proportions of rejecting rumor tweets (both true and false rumors) between these two groups, we observed a small increase in rejection rate after debunking. However, for the 33 false rumors with more than 200 tweets, the average rejection rate decreased slightly from 3.47% to 3.08%. The average endorsing rates before and after debunking were 91.81% and 91.79%, respectively. Additionally, the average rates of unclear tweets before and after debunking were 4.72% and 5.13%, respectively.

A series of chi-square tests (see Appendix 2) were run to assess differences in the proportion of endorsing and rejecting tweets before and after each rumor was debunked. Of the 33 false rumors that had more than 200 tweets, 16 showed no significant difference in attitude toward the rumor (endorsing vs rejecting) before and after. In 12 of the remaining 17 rumors, the proportion of rejecting tweets significantly increased after

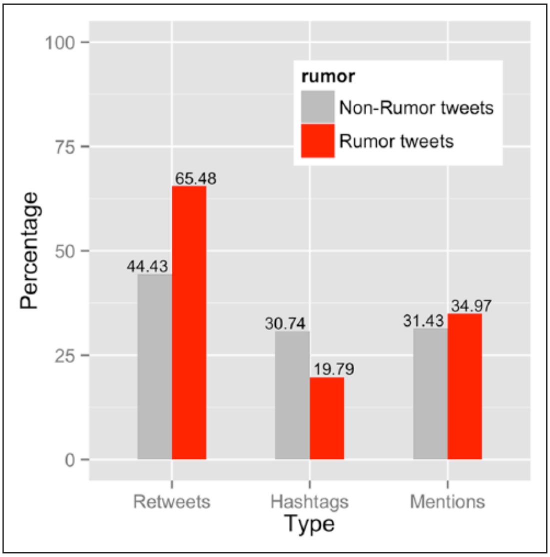


Figure 2. Percentage of each tweet type between rumor and non-rumor tweets.
The comparison of proportion of tweets containing a hashtag and a mention was on original tweets ($n = 108{,}951$), while the analysis on retweets was on the entire tweets including both original and retweets ($n = 315{,}647$).

debunking, while that of endorsing tweets decreased. Nevertheless, such attitude changes were small, given that the majority of tweets were endorsing both before and after the publication of debunking information.

However, we did notice that rumors originating with satirical news websites tended to show relatively strong responses to corrections. In five such cases out of eight, when the debunking sites clarified that the rumor was based on a fake story, the proportion of tweets endorsing the rumor significantly decreased, whereas that of rejecting tweets increased (all $p < .001$). For instance, after Snopes highlighted the satirical origin of a rumor quoting Romney saying, "I can relate to black people" because his ancestors once owned slaves, the number of tweets mentioning either "fake" or "satire" increased from 3 (0.2%) to 241 (3.1%). Although the difference was not statistically significant, the two other satire-based rumors also showed similar trends after debunking.

Since there could be many factors influencing the spread of a rumor, we sought out evidence of debunking beyond the three fact-checking sites in the original research design. First, we examined whether there were rumor-correcting tweets that mentioned other fact-checking sites (e.g. Politifact) or rumor-debunking sites (e.g. truthorfiction.com), but we did not find any in our rumor dataset. Second, in order to check whether mainstream media debunked the rumor during our data collection period, we searched Google with the same set of keywords used for retrieving rumor tweets. Among the 33 false rumors, 4 received mainstream media coverage. However, in three cases out of four, mainstream media did not directly verify the rumor and instead passively reported about the event without adjudication. Such coverage led to more tweets about these rumors, including a large number linking back to the mainstream media. In one case, MSNBC and The Washington Post publicly apologized for having reported on the rumor about Romney's campaign slogan without verification. We observed that although these apologies tamed the rumor temporarily, tweets endorsing the rumors surfaced again within a month.

To get a clear picture of how rumor debunking took place on Twitter, we separately analyzed the tweets that cited or paraphrased the fact-checking sites' debunking notices. A total of 858 tweets were extracted using keyword searches. Of these, 91.14% ($n=782$) had been coded as "rejecting" the rumor. Our codebook instructed coders to call a tweet "rejecting" if it repeated rumor-debunking sites even when they did not add their own comments denying the rumor. A closer look at these "rejecting" tweets citing fact-checking sites revealed that the majority of them (71.23%, $n=557$) were in fact repeating the false claim, even though they linked to the fact-checking site's correction. For instance, the tweet, "RT @snopes: Did Mitt Romney say he can relate to black people because his ancestors once owned slaves? http://t.co/VZPuZVjf"; does not clearly reject the rumor but re-states the rumor itself, as summarized in the Snopes headline for the article debunking it. Unless readers clicked on the link, they would only be exposed to the rumor, not to the correction. In this sense, only 28.77% ($n=225$) of tweets that referenced fact-checking sites were truly debunking by clearly indicating that the rumor was false.

Moreover, we found that despite the correction published by fact-checking sites, 5.01% of these tweets ($n=43$) still endorsed the rumor and cast doubt on the objectivity of the fact-checking entities. One user shared a link to Factcheck.org's webpage that debunked an Obama-related rumor and commented that "See how Annenberg's 'Fact Check' tries to cover for Obama … Lol!!" The remaining tweets, 3.85% ($n=33$), were categorized as "unclear" because they still expressed uncertainty about the rumor.

Discussion and conclusion

Previous scholars have observed that Twitter's unique features and users provide a medium for broad-based, free flowing, and instantaneous discussion of political claims (Colleoni et al., 2014; Park, 2013). Our investigation into political rumoring during the 2012 US presidential election season revealed that Twitter also served as a platform for partisans to selectively share unsubstantiated claims with their followers and accelerate rumor virality. First, our network analysis of political rumor tweets revealed that rumor publics selectively transmitted rumors about opposing candidates and formed polarized communities based on the rumors' target. People who spread one anti-Obama rumor also tended to spread other anti-Obama rumors, but did not spread anti-Romney rumors. The same applied to users who shared anti-Romney rumors. These results suggest that the circulation of rumors occurs within "echo chambers" defined by political homophily. Such echo-chamber effects, which amplify existing beliefs, can be detrimental to a marketplace of ideas, the notion that ideas should compete for acceptance based on truth and merit rather than ideological attitudes (Sunstein, 2009).

Interestingly, however, we did not find a partisan community structure for rumor rejecters. There were no sizable, cohesive groups that actively participated in debunking rumors about the candidate they support. Rather, many rejecters engaged in debunking both anti-Obama anti-Romney rumors. The different network structures and divergent political homophily between rumor-tellers and rejecters may imply that the partisan selective exposure phenomenon depends on the nature of the political activity. Alternatively, the fact that the size of rumor rejecters (2.7%) was only a fraction of rumor spreaders may mean that those who transmitted diverse views belonged to a minority. It could also simply be a reflection of the subversive pleasure of partisan rumoring: debunking is dull compared to spreading an attention-grabbing rumor.

We further examined various types of social relationships on which rumor spreaders rely. We found that rumors diffuse through homogeneous follower–followee relationships using the retweet feature rather than through public hashtag communities. The proportion of rumors that contained a hashtag was much lower than that of non-rumor tweets. By including trendy hashtags, users can go beyond their follower network and potentially expose much larger audience to their tweets. However, within rumor tweets, we observed more transmission of rumors within follower networks than engagement in dialogue with larger political discussion communities.

We also found that rumor-tellers do not signal strong motivations to discuss the merit of a rumor nor to seek accurate information. Even though fact-checking sites were available for verification, only a small fraction of rumor tweets either mentioned or retweeted such sites. This finding echoes Rojecki and Meraz's (2014) description of political rumors, where information seeking motivations seem largely absent from political rumor diffusion. Yet, while small in terms of the absolute numbers, we observed significant shifts in rumor diffusion trends in 12 out of 33 rumors: the proportion of endorsing tweets decreased and rejecting tweets increased after debunking. In particular, satire-based rumors responded to fact-checking sites' corrections, suggesting that Twitter users may not want to humiliate themselves by believing in obviously fake stories.

A closer examination of rumor-rejecting tweets also revealed that fact-checking sites might design their debunking notices more effectively for Twitter spreading. Their current headlines often state a false claim without clearly indicating their verdict. Users retweeting such headlines simply repeat a false statement which, according to Nyhan and Reifler (2012), makes the claim easy to remember, thus eventually leading to an association between the rumor and the target. This may or may not be a conscious tactic deployed by partisans to spread rumors. Therefore, future research needs to evaluate whether factcheckers could more effectively debunk rumors on social media by clearly stating their verdict within their headlines or tweets.

This study has a number of limitations. First, since retweeting is common on Twitter in general, and since the high retweet proportion in our rumor dataset may be partly explained by the inclusion of a few extremely popular rumors, systematic comparisons controlling for the nature and popularity of a topic would be beneficial.[11] Second, the unique features of Twitter may engender communicative behaviors different from other social media platforms or face-to-face interactions. Unlike Facebook, Twitter relationships are asymmetrical, meaning that one user can follow another without their permission. Such architecture creates a type of social space that is at once both public and private (Colleoni et al., 2014). Therefore, comparisons between Twitter and other platforms would further explain the effects of communication channel on rumor diffusion. Additionally, we are limited by our ability to observe only the manifested behaviors of Twitter users, and we may have missed the rumor debunking effects that occurred at the perceptual level. Finally, the activity we observe on Twitter represents just a slice of the larger discursive scene within which rumors unfold. It is possible, indeed likely, that rumor telling and debunking are shaped by different media contexts and non-mediated relationships.

In sum, this article systematically examines political rumor diffusion patterns on Twitter during the 2012 US presidential election. Although Twitter has the potential to challenge the flow of misinformation, our analysis showed that the platform mainly served as a conduit for rumors to spread through partisan communities rather than as a self-correcting system involving a larger political discussion community. Future research is encouraged to explore rumor spreaders' motivations and social media tactics that could help develop efficient debunking strategies.

Acknowledgements

The authors thank the Annenberg Innovation Lab at the Annenberg School for Communication and Journalism for providing the data. They are grateful to Kjerstin Thorson for her helpful comments.

Declaration of Conflicting Interests

The author(s) 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: This research received financial support from the Annenberg School of Communication, the Annenberg Innovation Lab, and the Undergraduate Research Associates Program (2014) at the University of Southern California.

Notes

1. For a comparison between Gnip PowerTrack and the freely accessible Twitter Streaming API, see Driscoll and Walker (2014).
2. Due to difficulty in identifying rumor-relevant tweets using keyword searches.
3. We may have missed tweets related to the rumor but using none of these keywords.
4. Not all retweets are duplicates of the original tweet. We treated retweets as unique tweets when users modified the original tweet even by one character.
5. Whenever Krippendorff's alpha fell below .75, coders re-coded the tweets after discussion with the author.
6. A total of 25 rumors were vetted by several fact-checking sites, which disagreed only in one case.
7. Accounts held by fact-checking sites were excluded from the analyses.
8. One rumor did not have any rejecters.
9. We classify tweets as retweets if they contain one of several retweet conventions (Boyd et al., 2010), for example, RT @username.

10. Our non-rumor tweets include a higher proportion of retweets than in previous studies. This may be due to undocumented biases resulting from their reliance on the Twitter Search API (Driscoll and Walker, 2014; González-Bailón et al., 2012).
11. Retweeting is also easier to automate with bots to form artificial grassroots campaigns (i.e. astroturfing).

References

Adamic LA and Glance N (2005) The political blogosphere and the 2004 U.S. election: divided they blog. In: *Proceedings of the 3rd international workshop on link discovery*, New York, 21 August.

Allport GW and Postman LJ (1945) The basic psychology of rumor. *Transactions of the New York Academy Sciences* 8(2): 61–81.

Allport GW and Postman LJ (1947) An analysis of rumor. *The Public Opinion Quarterly* 10(4): 501–517.

Barbera P, Jost JT, Nagler J, et al. (2015) Tweeting from left to right: is online political communication more than an echo chamber? *Psychological Science* 26: 1531–1542.

Bekafigo MA and McBride A (2013) Political participation of Twitter users during the 2011 gubernatorial elections. *Social Science Computer Review* 31(5): 625–643.

Benkler Y (2006) *The Wealth of Networks: How Social Production Transforms Markets and Freedom*. New Haven, CT: Yale University Press.

Bode L, Hanna A, Yang J, et al. (2015) Candidate networks, citizen clusters, and political expression: strategic hashtag use in the 2010 midterms. *The Annals of the American Academy of Political and Social Science* 659(1): 149–165.

Boutet A, Kim H and Yoneki E (2012) What's in Twitter: I know what parties are popular and who you are supporting now. In: *Proceedings of the 2012 international conference on advances in social networks analysis and mining*, New York, pp. 129–132. Available at: http://www.cl.cam.ac.uk/~ey204/pubs/2012_ASONAM.pdf

Boyd D, Golder S and Lotan G (2010) *Tweet, Tweet, Retweet: Conversational Aspects of Retweeting on Twitter. HICSS-43*. Kauai, HI: IEEE.

Bruns A and Burgess (2011) The use of Twitter hashtags in the formation of ad hoc publics. *Paper presented at the European consortium for political research conference*. Available at: http://eprints.qut.edu.au/46515/1/The_Use_of_Twitter_Hashtags_in_the_Formation_of_Ad_Hoc_Publics_%28final%29.pdf

Bruns A and Highfield T (2013) Political networks on Twitter. *Information, Communication & Society* 16(5): 667–691.

Byrne S and Hart PS (2009) The boomerang effect: a synthesis of findings and a preliminary theoretical framework. *Communication Yearbook* 33: 3–37.

Cacciatore M, Yeo SK, Scheufele DA, et al. (2014) Misperceptions in polarized politics: the role of knowledge, religiosity, and media. *PS: Political Science & Politics* 27: 654–661.

Colleoni E, Rozza A and 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.

Conover MD, Ratkiewicz J, Francisco M, et al. (2011) Political polarization on Twitter. In: *Fifth international AAAI conference on weblogs and social media*. Available at: https://www.aaai.org/ocs/index.php/ICWSM/ICWSM11/paper/viewFile/2847/3275

DiFonzo N and Bordia P (2007) *Rumor Psychology: Social and Organizational Approaches*. Washington, DC: American Psychological Association.

DiFonzo N, Bourgeois M, Suls J, et al. (2013) Rumor clustering, consensus, and polarization: dynamic social impact and self-organization of hearsay. *Journal of Experimental Social Psychology* 49(3): 378–399.

DiFonzo N, Suls J, Beckstead JW, et al. (2014) Network structure moderates intergroup differentiation of stereotyped rumors. *Social Cognition* 32(5): 409–448.

Driscoll K and Walker S (2014) Working within a black box: transparency in the collection and production of big Twitter data. *International Journal of Communication* 8: 1745–1764.

Farrell H (2012) The consequences of the internet for politics. *Annual Review of Political Science* 15: 35–52.

Festinger L (1957) *A Theory of Cognitive Dissonance*. Evanston, IL: Peterson.

Friggeri A, Adamic LA, Eckles D, et al. (2014) Rumor cascades. In: *Proceedings of the eighth international AAAI conference on weblogs and social media*, pp. 101–110. Available at: https://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/viewFile/8122/8110

Garrett RK (2009a) Echo chamber online? Politically motivated selective exposure among Internet news users. *Journal of Computer Mediated Communication* 14(2): 256–285.

Garrett RK (2009b) Politically motivated reinforcement seeking: reframing the selective exposure debate. *Journal of Communication* 59(4): 676–699.

Garrett RK (2011) Troubling consequences of online political rumoring. *Human Communication Research* 37(2): 255–274.

González-Bailón S, Wang N, Rivero A, et al. (2012) *Assessing the Bias in Communication Networks Sampled from Twitter*. Available at: http://arxiv.org/abs/1212.1684

Gottfried JA, Hardy BW, Winneg KM, et al. (2013) Did fact checking matter in the 2012 presidential campaign? *American Behavioral Scientist* 57(11): 1558–1567.

Hannak A, Margolin D, Keegan B, et al. (2014) Get back! You don't know me like that: the social mediation of fact checking interventions in Twitter conversation. In: *Proceedings of the eighth international AAAI conference on weblogs and social media*. Available at: http://www.brianckeegan.com/papers/ICWSM_2014.pdf

Himelboim I, McCreery S and 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.

Jacobson S, Myung E and Johnson SL (2015) Open media or echo chamber: the use of links in audience discussions on the Facebook pages of partisan news organizations. *Information, Communication & Society*. Epub ahead of print 23 October. DOI: 10.1080/1369118X.2015.1064461.

Kim J (2011) The contribution of social network sites to exposure to political difference: the relationships among SNSs, online political messaging, and exposure to cross-cutting perspectives. *Computers in Human Behavior* 27: 971–977.

Knobloch-Westerwick S and Kleinman SB (2012) Preelection selective exposure: confirmation bias versus information utility. *Communication Research* 39(2): 170–193.

Knobloch-Westerwick S, Johnson BK and Westerwick A (2015) Confirmation bias in online searches: impacts of selective exposure before an election on political attitude strength and shifts. *Journal of Computer Mediated Communication* 20(2): 171–187.

Kwak H, Lee C and Moon S (2010) What is Twitter, a social network or a news media. In: *Proceedings of the 19th world-wide web (WWW) conference*, Raleigh, NC, 26–30 April.

Larsson AO and Moe H (2011) Studying political microblogging: Twitter users in the 2010 Swedish election campaign. *New Media & Society* 14(5): 729–747.

McKinney MS, Houston B and Hawthorne J (2014) Social watching a 2012 republican presidential primary debate. *American Behavioral Scientist* 58(4): 556–573.

McPherson M, Smith-Lovin L and Cook JM (2001) Birds of a feather: homophily in social networks. *Annual Review of Sociology* 27: 415–444.

Marwick AE and boyd d (2011) I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New Media & Society* 13(1): 114–133.

Nagarajan M, Purohit H and Sheth A (2010) A qualitative examination of topical tweet and retweet practices. In: Proceedings of the fourth international AAAI conference on weblogs and social media. Available at: https://www.aaai.org/ocs/index.php/ICWSM/ICWSM10/paper/viewFile/1484/1880

Nyhan B and Reifler J (2010) When corrections fail: the persistence of political misperceptions. *Political Behavior* 32(2): 303–330.

Nyhan B and Reifler J (2012) Misinformation and fact-checking: research findings from social science. New America Foundation Media Policy Initiative Research Paper. Available at: http://www.dartmouth.edu/~nyhan/Misinformation_and_Fact-checking.pdf

Papacharissi Z (2009) The virtual sphere 2.0: the internet, the public sphere and beyond. In: Chadwick A and Howard P (eds) *Handbook of Internet Politics*. London: Routledge, pp. 230–245.

Park CS (2013) Does Twitter motivate involvement in politics? Tweeting, opinion leadership, and political engagement. *Computers in Human Behavior* 29(4): 1641–1648.

Rainie L and Smith A (2012) Social networking sites and politics. Available at: http://www.pewinternet.org/files/old-media/Files/Reports/2012/PIP_SocialMediaAndPoliticalEngagement_PDF.pdf

Riffe D, Lacy S and Fico FG (1998) *Analyzing Media Messages: Using Quantitative Content Analysis in Research*. London: Routledge.

Rojecki A and Meraz S (2014) Rumors and factitious informational blends: the role of the web in speculative politics. *New Media & Society*. Epub ahead of print 19 May. DOI: 10.1177/1461444814535724.

Rosvall M and Bergstrom CT (2008) Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences USA* 105(4): 1118–1123.

Shibutani T (1966) *Improvised News: A Sociological Study of Rumor*. Indianapolis, IN: Bobbs-Merrill.

Smith MA, Rainie L, Shneiderman B, et al. (2014) Mapping Twitter topic networks. Available at: http://www.pewinternet.org/2014/02/20/part-2-conversational-archetypes-six-conversation-and-group-network-structures-in-twitter/

Sunstein C (2009) *On Rumors*. New York: Farrar, Straus and Giroux.

Tonkin E, Pfeiffer HD and Tourte G (2012) Twitter, information sharing and the London riots? *Bulletin of the American Society for Information Science and Technology* 38(2): 49–57.

Wasserman S and Faust K (1994) *Social Network Analysis: Methods and Application*. Cambridge: Cambridge University Press.

Weeks BE (2015) Emotions, partisanship, and misperceptions: how anger and anxiety moderate the effect of partisan bias on susceptibility to political misinformation. *Journal of Communication* 65: 699–719.

Weeks BE and Garrett RK (2014) Electoral consequences of political rumors: motivated reasoning, candidate rumors, and vote choice during the 2008 U.S. presidential election. *International Journal of Public Opinion Research* 26(4): 401–422.

Williams H, McMurray JR, Kurz T, et al. (2015) Network analysis reveals open forums and echo chambers in social media discussions of climate change. *Global Enviornmental Change* 32: 126–138.

World Economic Forum (2014) The rapid spread of misinformation online. Available at: http://reports.weforum.org/outlook-14/top-ten-trends-category-page/10-the-rapid-spread-of-misinformation-online/

Author biographies

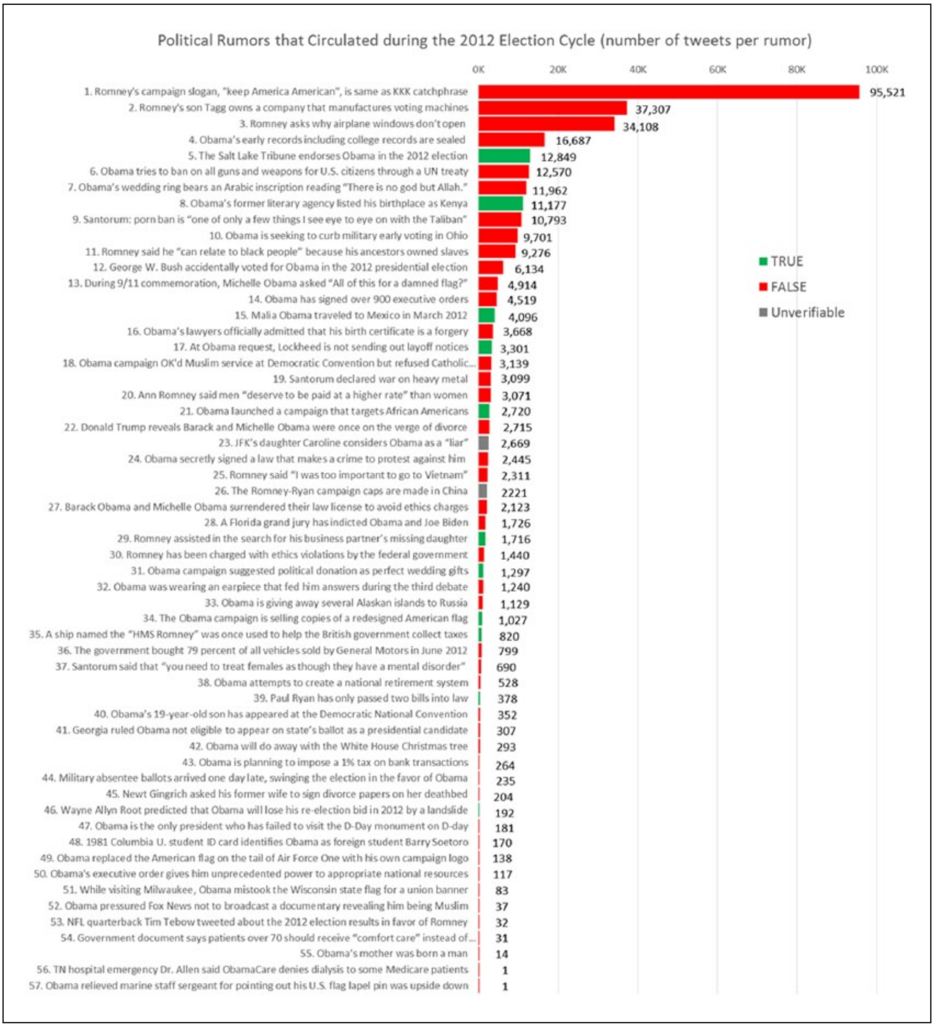
Jieun Shin is a Ph.D candidate at the Annenberg School for Communication and Journalism at the University of Southern California. Her research interests focus on diffusion of information and social influence on the web.

Lian Jian is an Assistant Professor at the Annenberg School for Communication and Journalism at the University of Southern California. Her research interests include the economics of information, crowd-sourcing, crowd-funding, and collective intelligence.

Kevin Driscoll is a Postdoctoral researcher at Microsoft Research. His research concerns the popular and political cultures of networked personal computing.

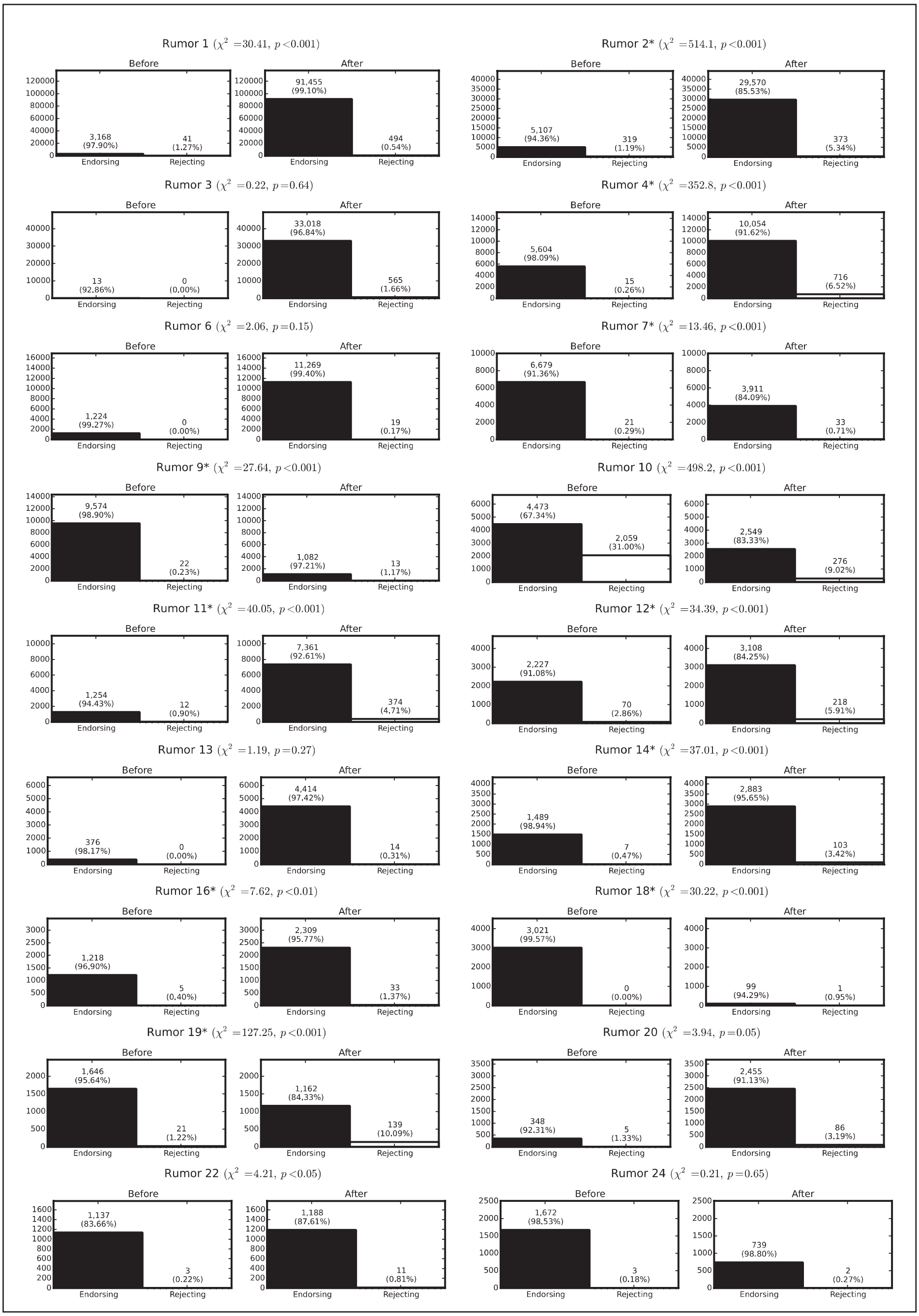
François Bar is an Associate Professor of Communication and Spatial Sciences at the University of Southern California. His research and teaching explore the social and economic impacts of information technologies with a specific focus on telecommunication policy, user-driven innovation and technology appropriation.

Appendix 1



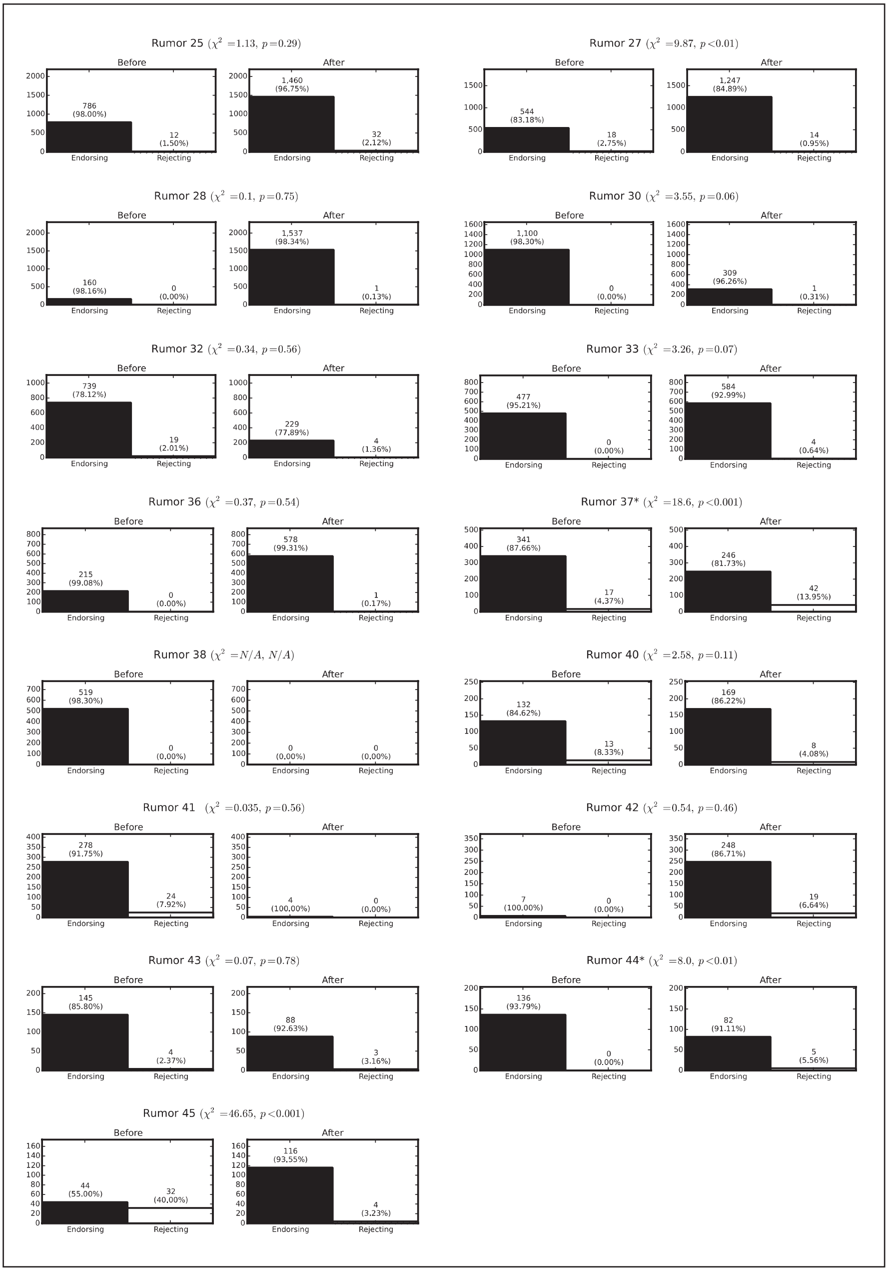
Descriptions of 57 rumors.

Appendix 2



(Continued)

Appendix 2. (Continued)



Frequencies of endorsing and rejecting tweets before and after debunking.
Details about each rumor are in Appendix 1. Values in parentheses indicate percentage of endorsing or rejecting tweets before and after debunking. Asterisk next to Rumor ID indicates a significant increase in the proportion of rejecting tweets and a significant decrease in the proportion of endorsing tweets after debunking.