Campaign 2016—The First Words and Their Audiences

George Robert Boynton

Department of Political Science, University of Iowa, Iowa City, USA
Email: Bob-boynton@uiowa.edu

Abstract

When do citizens begin to attend to election campaigns for the presidency? The timing is changing as the infrastructure of communication changes. What could be said in the days when television was dominant is no longer a full answer given the emergence of new media as a vehicle for political communication. This report addresses the question of when by examining the messages posted to Twitter after Jeb Bush said he would consider running for the nomination of the Republican party. What was the reaction in those first two weeks of the campaign?

Keywords

Republican Candidate, Jeb Bush, December 2014, Twitter, Tweets, Audience

1. Introduction

On December 16, 2014 Jeb Bush announced that he was seriously considering seeking the nomination of the Republican party for the presidency. The campaign began. Other candidates started their own planning considering how their plans might have changed with this announcement. The news media had something to write about. Survey researchers started asking which candidate you like best all over the country. And the public? A few began to go public with their thoughts, and the prime candidate for public communication about politics is Twitter.

The question has been: when do citizens begin to attend to the election in large numbers? An early answer was Labor Day. However, adding the caucuses and primaries in the late winter and spring of the year of the election moved attention earlier than Labor Day (Balz, 2012). One way to address the question has been to look at audiences for news programs as they begin to feature news about the election. When television dominated public communication about politics, that was an adequate way to make an assessment. However, attention to television news has been declining quite substantially. Pew Research reported a twenty percent drop in viewing

cable news between 2009 and 2014 (Holcomb, 2015). At the same time the use of Twitter and Facebook as sources of news was increasing to 63% of users, and 59% of Twitter users said they followed news about developing events on Twitter (Barthel et al., 2015). To answer the question about citizen attention the use of social media as a source of news must be investigated. This note is a report on the volume of messages about the candidates and the audiences for those messages at what might be thought of as the beginning of the campaign.

There have been two foci of research about Twitter and elections. At least as early as 2009 scholars started attempting sentiment analysis of twitter messages to predict the outcome of an election (Tumashan, et al., 2010). There are some indications of success and some of failure. The results of this research stream are summarized in "Limits of electoral predictions using social media data" (Gayo-Avello et al., 2011). Notwithstanding, the research continues. The second stream is research on the use candidates make of social media in their campaigning. This has been, in part, inspired by the Obama campaigns of 2008 and 2012 because of the effectiveness of their campaign activity. Chadwick (2014) has been influential in setting the focus for this research. ShantoIyengar's Media Politics has a chapter "New Media, New Forms of Campaigning" (Iyengar, 2016). New Media in American Political Campaigns; Controlling the Message is a collection of research reports about new media and campaigning (Farrar-Myers & Vaughn, 2015). Only a modest amount of research has been conducted examining the communication activity of users of Twitter in elections. Andreas Jungherr is one scholar who has done research by looking at the political communication of Twitter users (Jungherr, 2014). FlippoMenczer, leading the Center for Complex Networks and Systems Research at Indiana University, has been investigating communication activity related to politics for some time. And that is also the focus of the research reported here. As far as we can tell no one has attempted to systematically assess the audience for twitter messages about electoral politics.

After a brief description of the procedures for collecting the Twitter messages used in the analysis the paper goes through the two steps needed to begin to answer the question about public interest in the 2016 election from its very first days. The first step is examining the number of messages posted to Twitter about the primary candidates in that two-week period. The second step is estimating the audience for those messages.

2. Methods

The Twitter messages were collected using the computer program Desktop Archivist running on Windows computers; it queries the search API every five minutes. The Twitter API responds with up to 1900 tweets to each request. In the unusual case of messages being posted more frequently than 1900 in five minutes it is a time based sample. However, the program can acquire as many as 547,000 tweets in a single day, but Twitter samples the tweets posted when responding to the search request. It does not provide the full stream of messages so one is working with a sample. The search terms used were: Jeb Bush, Chris Christie, Hillary Clinton, Ted Cruz, Mike Huckabee, Rand Paul, Marco Rubio, Paul Ryan, and Scott Walker. Hillary Clinton, the leading Democratic candidate, was included as a point of reference. Both first and last names by which they are known were included in the search terms. While this is necessary to differentiate Bush from the many other individuals named Bush, including his father and brother, it does mean some references to the candidates are lost when both names were not included in the tweets. These were the primary candidates being communicated about at the time. Romney had not announced his intentions, and other candidates would become active later. In addition to the message written and posted by users the program acquires 180 items of information about each tweet and the individuals who posted them. Information about the frequency of messages posted by accounts is used in the next section and the number of followers for each account is used in the following section.

3. The Magnitude of the Twitter Message Stream

Twitter regularly divulges information about the number of users and the number of messages being posted by those users as part of their quarterly financial report. As of the fourth quarter of 2014 they reported 288 million users who are active at least once a month and who posted 500 million messages a day. Twenty-three percent of the 288 million accounts are users in the United States. Five hundred million is the total stream of messages. That does not provide much information about political messages.

The number of messages captured during the two weeks, December 16 through December 31, is displayed in Figure 1.

The announcement came and data collection started later in the day on December 16 when just over 50,000

tweets were captured. The 17th saw a spike of over 10,000 messages posted to Twitter. The number declined for the next few days and settled at around 20,000 a day with a day off for Christmas. Spikes of tweets later in the campaign would exceed the attention to Mr. Bush's announcement, but it was a spike starting what would become the extended campaign for the Republican nomination.

Table 1 disaggregates the stream by candidate.

Four hundred and seventy thousand tweets were collected in the two weeks. That is a subset of the total since Twitter does not supply the total stream through the search API. The number of tweets mentioning Bush, 182,554, exceeds the number for any other candidate. Hillary Clinton was second in terms of total number of messages mentioning her with 88,780 tweets. The Republican candidate with the second largest number of mentions was Marco Rubio. That was primarily the result of his criticism of the Obama administration opening relations with Cuba. He was vociferous in his opposition and that inspired a large number of mentions (Grossman & Shiskin, 2014). The number of messages mentioning Bush and the large number of tweets mentioning Rubio illustrate the potential volatility of communication using Twitter. Though it is not shown in the table mentions of Rubio fell quite dramatically after these two weeks. At the bottom of the distribution were Kasich, Huckabee, Jindal and Ryan. Ryan was to be the first drop out of the campaign when he announced that he would not run later in January. The other three were not deterred by the small number of tweets mentioning them.
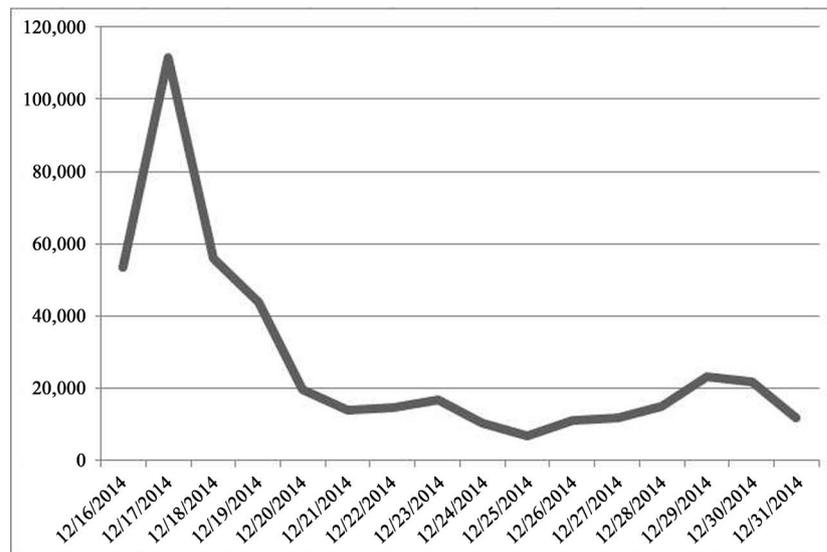
Figure 1. Twitter messages per day for all candidates 12/16/2014 through 12/31/2014.

Table 1. Tweets and tweeters December 2014.

Candidate	Total tweets	Unique Users	Avg Tweets per User
Bush	182,554	75,364	2.41
Christie	17,713	12,530	1.41
Clinton	88,780	47,892	1.85
Cruz	41,511	18,790	2.21
Huckabee	2616	1997	1.31
Jindal	4122	2704	1.52
Kasich	1557	1187	1.31
Paul	45,671	21,717	2.10
Rubio	68,933	30,405	2.27
Ryan	7656	5929	1.29
Walker	10,501	6154	1.71
Totals	471,614	224,669	2

The four hundred and seventy thousand tweets were posted by 224,669 unique Twitter accounts. Individuals posting a message about Bush were the most loquacious averaging 2.41 tweets apiece. Ryan, Kasich, and Huckabee were at the bottom with only about 1.3 tweets apiece.

The number of Twitter messages per day is a good reflection of attention being paid to a subject. One might plausibly argue that the campaign for the 2016 election started among political insiders as soon as the 2012 election ended. However, this is the first burst of attention from a broader public that we know about.

4. Audiences

This volume of tweets early in the campaign is an important sign of the place of Twitter in the evolving communication structure of politics. However, posting messages is only the first step. The number of people who have access to the messages through their accounts is a measure of audience, and audience is critical to the importance of Twitter as communication about politics.

There are two institutional arrangements that structure the flow of communication in Twitter. One is search. A user can request the twitter messages containing a search term. Twitter has the information about tweets found using search, but they do not share that information. The second structuring arrangement is following. The followers of a user are sent messages posted by the user. The number of followers of every person posting a tweet is available as part of the meta data for the tweet. That is the information on followers reported here.

The total number of followers associated with each tweet is an indication of the total views in the system. If a user posts more than one message to Twitter the followers see as many as the user posted and the followers for that user are counted as often as the user tweets. But that does not give the audience, i.e., the number of followers of the unique individuals posting a message. **Table 2** starts by deleting the duplicates, for the users who posted more than once, so their followers would only be counted one time.

The users who posted a message mentioning Bush had, according to Twitter, 469,172,141 followers. However, the size of fake followers for an account like Bush's has been documented and roughly 30% are fake (Boynton et al., 2013). The first column reports the number of followers of users mentioning each candidate. The second column adjusts for the known 30% of fake followers. Even after that adjustment the number of followers is too large. How does one get such large numbers? In this case it was the news media.

The news media from CNN to the Washington Post have millions of followers, and fifty of the user accounts posting messages had more than one million followers. If a candidate is mentioned by even half of these user accounts in a two week period the reach of the communication is very great.

That is not the whole story, however. The distribution of followers is so skewed that columns 4 through 10 is a better way of getting a grasp on the distribution than any single summary number. The accounts listed in **Table 3** are all in the 10th column with more than 10,000 followers. For the totals 25% of the users have fewer than

Table 2. The number of followers of users who posted a message mentioning a candidate.

	Number Followers	Adjusted Followers	<50	50 - 100	100 - 500	500 - 1000	1000 - 5000	5000 - 10,000	10,000+
Bush	469,172,141	332,788,846	9998	6992	23,953	16,111	13451	2259	2626
Christie	97,257,066	68,839,943	1365	1084	3813	1979	3029	564	564
Clinton	305,851,993	216,872,829	6504	4671	15,361	6672	10,305	1954	2424
Cruz	100,107,215	71,104,754	2088	1412	5001	2706	5453	1061	1069
Huckabee	6,581,611	4,712,569	369	212	595	253	407	81	80
Jindal	13,416,624	9,540,047	267	172	695	406	820	158	187
Kasich	6,024,928	4,286,724	122	82	319	193	320	71	80
Paul	166,400,582	117,717,677	2447	1836	6179	3197	5602	1168	1286
Rubio	240,122,816	169,947,051	2954	2438	9363	4826	7790	1373	1661
Ryan	19,543,880	14,058,388	706	559	2158	897	1198	179	232
Walker	22,073,836	15,796,145	657	391	1680	908	1814	372	332
Totals	1,446,552,692	1,025,664,973	27,108	17,479	27,477	38,148	50,189	9240	10,541

Table 3. Followers of new media.

Twitter User	Number Followers
CNN	15,123,388
The New York Times	14,353,808
The Economist	6,756,640
TIME.com	6,592,851
Reuters Top News	5,649,816
Wall Street Journal	5,479,243
Huffington Post	4,963,526
Fox News	4,647,524
VEJA	4,496,025
The Associated Press	4,456,569
G1	4,154,670
Forbes	4,135,798
Rolling Stone	4,111,264
Washington Post	3,794,809

Following is very high among users of Twitter for political communication, and that means the "message" goes out to an extensive audience.

However, there is a problem with these counts that has not been taken into account. The problem is overlapping following or network density. If a user followed four of the users posting a tweet they would be counted four times. We computed overlapping followers of the major republicans considering running in March. On average individuals following one candidate also followed three others. That means the total number of followers would be divided by four to get a reasonable estimate of audience. A more adequate procedure requires computing the overlap in the network. But the network has an exceedingly large number of individual users. Computing the network connections in a network of this size is not feasible. There is a sampling procedure suggested by Bruns and Stieglitz (12/6/2012) that can be used to estimate the density of the network. That procedure will be used in future research.

5. Conclusion

This is a very preliminary step in assessing the audience of Twitter messages about politics. Four hundred and seventy thousand tweets is a very large sample by almost any standard. We know the audience receiving those tweets is large. We also know there are several problems with the counts of the audience. If Twitter communication is going to become as important as communication on television, it will be in part because it reaches an equal audience. That makes it important to begin to estimate the size of the audience, which is what is done in this report.

References

Balz, D. (2012) Lead up to Labor Day May Determine Winner of Presidential Race. *Washington Post*. http://www.washingtonpost.com/politics/lead-up-to-labor-day-may-determine-winner-of-presidential-race/2012/04/14/gIQAPJiPHT_story.html

Barthel, M., Shearer, E., Gottfried, J., & Mitchell, A. (2015). The Evolving Role of News on Twitter and Facebook, Pew Research Center. http://www.journalism.org/2015/07/14/the-evolving-role-of-news-on-twitter-and-facebook/

Boynton, G.R., Bates, A., Bettis, E., Bopes, M., Brandt, F.D., Hahn, J., Hart, T., Headley, C., Kopish, J., Maharry, R., Matson, J., Mohoff, K., Mraz, R., Palmer, M., Pena, L., Phillips, B., Rhodes, A., Rosman, H., Sievers, C., Tate, D., Tyrrell, S., Villarrea, J., Wiese, P., & Wignall, A. (2013). The Reach of Politics via Twitter—Can That Be Real? *Open Journal of Political Science, 3,* 91-97. http://dx.doi.org/10.4236/ojps.2013.33013

Bruns, A., & Stieglitz, S. (2012). Quantitative Approaches to Comparing Communication Patterns on Twitter. *Journal of*

Technology in Human Services, 30, 160-185. http://dx.doi.org/10.1080/15228835.2012.744249

Chadwick, A. (2014). *The Hybrid Media System: Politics and Power*. Oxford: Oxford University Press.

Farrar-Myers, V., & Vaughn, J. (2015). *Controlling the Message: New Media in American Political Campaigns*. New York: University Press.

Gayo-Avello, D., Panagiotis, M., & Mustafaraj, E. (2011). Limits of Electoral Predictions Using Social Media Data. *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*, Barcelona, 17-21 July 2011, 4 p.

Grossman, A., & Shiskin, P. (2014). Marco Rubio Criticizes "Obama-Paul" Policy on Cuba. *The Wall Street Journal*. http://blogs.wsj.com/washwire/2014/12/21/marco-rubio-criticizes-obama-paul-policy-on-cuba/

Holcomb, J. (2015) Cable News: Fact Sheet, Pew Research Center. http://www.journalism.org/2015/04/29/cable-news-fact-sheet/#chart

Iyengar, S. (2016). *Media Politics*. New York: W. W. Norton and Company.

Jungherr, A. (2014). The Logic of Political Coverage on Twitter: Temporal Dynamics and Content. *Journal of Communication, 64,* 239-259. http://dx.doi.org/10.1111/jcom.12087

Tumashan, A., Sprenger, T., Sandner, P., & Welper, I. (2010). Election Forecasts with Twitter: How 140 Characters Reflect the Political Landscape. *Proceedings of the Fourth International Conference on Weblogs and Social Media*, AAAI Press, Menlo Park, California.