

RESEARCH ARTICLE

Ratioing the President: An exploration of public engagement with Obama and Trump on Twitter

Joshua R. Minot^{1,2,3,4*}, Michael V. Arnold^{1,2,3,4}, Thayer Alshaabi^{1,2,3,4}, Christopher M. Danforth^{1,2,3,4}, Peter Sheridan Dodds^{1,2,3,4*}

1 Department of Mathematics and Statistics, University of Vermont, Burlington, VT, United States of America, **2** Vermont Complex Systems Center, University of Vermont, Burlington, VT, United States of America, **3** Computational Story Lab, Vermont Complex Systems Center, University of Vermont, Burlington, VT, United States of America, **4** Vermont Advanced Computing Core, University of Vermont, Burlington, VT, United States of America

* joshua.minot@uvm.edu (JRM); peter.dodds@uvm.edu (PSD)



OPEN ACCESS

Citation: Minot JR, Arnold MV, Alshaabi T, Danforth CM, Dodds PS (2021) Ratioing the President: An exploration of public engagement with Obama and Trump on Twitter. PLoS ONE 16(4): e0248880. <https://doi.org/10.1371/journal.pone.0248880>

Editor: Kazutoshi Sasahara, Tokyo Institute of Technology, JAPAN

Received: August 10, 2020

Accepted: March 7, 2021

Published: April 14, 2021

Copyright: © 2021 Minot et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The data associated with this study can be downloaded from Twitter's public API (<https://developer.twitter.com/en/docs.html>). While Twitter's API terms of service will not allow us to share the individual messages we used, we used Twitter's gardenhose stream to provide a 10% random sample of tweets, from which we selected tweets relating to the accounts of interest. We have also provided the IDs of the replies and retweets necessary to replicate our findings in a public repository (<https://www.kaggle.com/joshminot/ratiometer>). Those looking to replicate

Abstract

The past decade has witnessed a marked increase in the use of social media by politicians, most notably exemplified by the 45th President of the United States (POTUS), Donald Trump. On Twitter, POTUS messages consistently attract high levels of engagement as measured by likes, retweets, and replies. Here, we quantify the balance of these activities, also known as “ratios”, and study their dynamics as a proxy for collective political engagement in response to presidential communications. We find that raw activity counts increase during the period leading up to the 2016 election, accompanied by a regime change in the ratio of retweets-to-replies connected to the transition between campaigning and governing. For the Trump account, we find words related to fake news and the Mueller inquiry are more common in tweets with a high number of replies relative to retweets. Finally, we find that Barack Obama consistently received a higher retweet-to-reply ratio than Donald Trump. These results suggest Trump's Twitter posts are more often controversial and subject to enduring engagement as a given news cycle unfolds.

1 Introduction

The ability for US presidents to communicate directly with the public changed dramatically during the 20th and early 21st centuries. Moving from Franklin Delano Roosevelt's famous fireside chats through the television addresses of Harry Truman and John F. Kennedy, we now find ourselves in the era of social media campaigns and presidencies [1]. Communication technology has especially accelerated the ability for presidents to “go public” [2] and make appeals directly and instantly to the electorate.

With its introduction in the 2000s, social media provided a new platform for direct communication. New mechanisms for information sharing have consequences for contagion: rapid spreading of content through a user-base capable of responding in real-time. While these

these results could either use a 10% historical collection or the Twitter API to query retweets and replies pertaining to specific users.

Funding: P.S.D. and C.M.D. received funding from the Massachusetts Mutual Life Insurance Company and Google Open Source under the Open-Source Complex Ecosystems And Networks (OCEAN) project. Neither funder had any role in the design of the study or presentation of its results.

Competing interests: JRM, TA, CMD, and PSD have been supported by a gift from MassMutual. Our program at-large receives support from MassMutual and Google, although neither organization specifically funded this study. This does not alter our adherence to PLOS ONE policies on sharing data and materials.

platforms democratize sharing for many categories of individuals on social media (e.g., celebrities interacting with fans), US presidential accounts present a unique opportunity to analyze particularly salient signals indicative of broader sociotechnical phenomena in an increasingly prominent aspect of politics.

Advertisers and political campaigns alike are concerned with engagement metrics for social media posts. Twitter itself markets its data as an early warning system for customer satisfaction and reputation management [3]—although the company has recently banned political advertising [4]. The rate with which a given post garners interactions (e.g., clicks, profile views, mouse-hovers, etc.) is a common measure of engagement in the digital realm. Activities in response to social media posts include retweets/shares, likes/favorites, and replies/comments. These activities all reflect specific user actions such as endorsing a post or expressing a divergent opinion. Some have theorized that simple ratios of activities may be used as proxies for of the polarity of the public's response to a given message [5, 6].

The term 'ratioed' is a Merriam-Webster "word we're watching" [7]. In the most general sense, the term refers to a social media post that receives more replies or responses than likes, favorites, or shares. In recent years the ratio has attracted growing interest by tech-journals and internet researchers, although this has largely been relegated to news articles and industry blog posts. [8–10]. On Twitter, the ratio value is generally taken to be defined by the ratio or balance of replies to likes or retweets. Here we will focus on the ratio of retweets to replies, as we show that like volume is often stable, while replies and retweets seem more reflective of specific public responses (Sec. 3).

In an effort to explore the dynamics of ratio space, in 2017 the website FiveThirtyEight Politics presented ternary ratios with activity counts normalized across retweets, comments, and likes [11]. Using this metric, the authors compared US politicians based on the ratio values their tweets receive. This work found noteworthy differences between politicians and political parties. As of late 2017, Trump tweets tended to be met with relatively more replies, while Obama tweets were met with relatively more retweets. Tweets for both presidents had a high value of normalized activities that were likes (formerly known as 'favorites'). Beyond presidents, FiveThirtyEight Politics compared responses to the tweets of Republican and Democratic senators. The senatorial accounts exhibited normalized response values largely reflective of their party's most recent president—Republican senators tended to have high values of normalized reply activities while Democratic senators had more retweets and likes. We have found minimal academic research on the topic of response activity ratios and their temporal dynamics, although response activities have been studied separately.

More broadly speaking, time series of Twitter response activities have been used to investigate the relationship between on- and off-line political activity by analyzing correlations between current events and trends in cumulative activity sums [12]. Fundamental models relating retweet activity to external activity have been proposed at the hour and day scale for the 2012 South Korean presidential election [13]. Kobayashi and Lambiotte find future retweet activity to be log-linear correlated with early tweet levels [14]. Outside of Twitter, analysis of Instagram posts immediately preceding the 2016 US presidential election showed a higher volume of user activity for Trump supporters as well as more intense pro-Trump activity at the sub-day time-scale when compared to pro-Clinton posts [15].

Twitter specifically has been identified as playing an active role in the political arena as evidenced by studies of improved social organization during the Arab Spring [16, 17], disinformation campaigns during the 2016 US presidential election [18–20], and political cohesion on the platform [21, 22]. Twitter has been shown to be a valuable source for data pertaining to domestic protests and social movements in the US [23–25]. The 2016 US presidential election in particular has been extensively analyzed through the lens of Twitter data [26–29]. There is

also increasing academic interest in the unique communication style of President Trump on Twitter [30, 31].

Twitter has been used to study several large-scale socio-technical phenomena such as the stock market [32, 33] and medical research [34, 35]. Messages on the platform have been shown to have a relationship with presidential approval rating polls [36–41]. Wang *et al.* examine how likes on Twitter can identify the preferences of Trump-account followers [42]. Although noisy at times, Twitter data can provide valuable insights on how audiences and populations respond to current events and specific messaging. Information diffusion and cascades on Twitter and other social networks has been the topic of much research [43–47]. Beyond politics on Twitter, Candia *et al.* propose a bi-exponential model to describe collective attention for online videos and manuscripts [48]. Media coverage of major events has been found to decay on roughly a weekly cycle, and scaling superlinearly with the population size of the affected area [49]. Amati *et al.* model retweet actions on dynamic and cumulative activity networks [50] while ten Thij *et al.* identify retweet graph characteristics common among viral content [51].

Some researchers have proposed critical levels of activity required to trigger information cascades on Twitter in light of tweet characteristics [52]. Others have proposed more general cascade requirements on social networks [53]. Jin *et al.* investigate how messages cross language and state communities on the platform [54]. Lee *et al.* predict retweet volume based on prior retweet spacing, user meta-data, tweet content [55]. Others have identified celebrity-users (including Barack Obama) that are central in spreading new content [56].

Regarding the representativeness of social media data, several studies have examined the degree to which Twitter represents the general population and voting public [57–60]. Grinberg *et al.* investigate fake news exposure for Twitter accounts belonging to eligible voters and found disproportionate exposure to a small number of accounts [61].

How likely are Twitter users, and specifically POTUS account followers, to be eligible voters in the US? Pew Research Center has estimated that 26% and 19% of US adults follow the Obama and Trump accounts, respectively [62]. The New York Times estimated however that less than 20% of Trump's followers are voting-age US citizens [63]. These considerations are important to remember when attempting to establish a relationship between Twitter metrics and political outcomes. Nevertheless, the platform has been shown to be a powerful tool for evaluating public sentiment towards politicians, and even detecting adversarial actions in the US political process.

Presidential communications are a timely topic in light of rapidly changing norms surrounding how the executive branch communicates with the public. Moreover, highly influential Twitter users such as US presidents produce signals in the medium that largely transcend the social network dynamics. When a president tweets, the message is nearly instantly amplified and spread by official and unofficial sources on the platform. Indeed, US presidents are often discussed at rates approaching function words on Twitter (i.e., the rate of usage of the name “Trump” may approach usage rates for words such as “the” and “RT” on the platform) [64]. Examining ultrafamous users somewhat removes the more prominent effects of network topology on the response activity and provides an opportunity to view response behavior largely as a function of timing and message content.

The volume of response activities on Twitter is arguably the most tangible metric that is publicly available for evaluating user-base responses to content on the platform. Furthermore, these values are highly visible to users and may influence behavior when users seek to affect the collective response to a communication (e.g., ‘take a stand’ by responding [65]). Establishing characteristic scales of activity volume and temporal dynamics for engagement is an important step in understanding how these values can provide insights on user-base response.

Counts of user activities responding to tweets can be readily viewed through the icons at the bottom of every tweet, which also provide the interface for user engagement. These values provide the end user with an indication of the tweet popularity and, in some cases, controversy. Taken together with the cultural context surrounding the original tweet, the ratio of these values open up the possibility of studying tweets through the lens of “ratiometrics”. This allows for the distillation of response activities into aggregated measures of the user base’s reaction. Moreover, enables the comparison of response activities on Twitter across accounts and across time. From there, we can begin to use the ratio values as a criteria when evaluating the content of tweets.

A simple calculation of the publicly accessible response activity counts is a starting point, but insufficient to fully investigate the full potential of ratiometrics. Here we present a suite of tools—collectively referred to as the “ratiometer”—that help us understand response activity profiles for tweets. Many factors contribute to interpreting the activity counts for a tweet, including typical activity ratios for the user and the age of the tweet. Understanding the typical response that an account receives requires building a historical view of the user’s tweets and their subsequent response activities. Another challenge is determining the typical response volume at a given time step since the tweet was issued.

We show example time series for retweets and replies in response to four Trump tweets authored after the 2016 election in [Fig 1](#). In the case of an ultra-famous user like Trump, we see an immediate response to tweets with early retweets and replies occurring seconds after the original tweet. For Trump’s tweets from after the 2016 election, we generally observe thousands of response activities within the quarter-hour, with response-activity counts nearing their final values within 24-hours of the original tweet. There may be modest growth during the proceeding week, but generally these values have stabilization periods at the day-scale ([S1](#) and [S2](#) Figs). The insets in [Fig 1](#) show the retweet-to-reply ratio time series. These values often fluctuate even while individual ratio time series appear to have a more stable trend. The above highlights some of the challenges in building an understanding for the expected volume and time scale associated with response activities for a specific account.

In the present study, we explore the time series of the volume of user activities in response to presidential tweets. We summarize the ratio values of responses to Obama and Trump tweets over three distinct periods of recent US political history. Our investigation includes ratio values for all three activity counts normalized against the sum of all activities (we refer to these as ternary ratios below). We also present characteristic time scales for response activity volume over the three political periods for both presidents. Finally, we present words that tend to appear more often in ratioed tweets for the @realDonaldTrump account. In section 2.1, we describe our data and methodology. In section 3, we present and discuss the results of our study. Finally, in section 4, we offer concluding remarks and point to areas for future work.

2 Methods

2.1 Data

We construct the dataset analyzed in the present study by filtering Twitter’s Decahose Streaming API [[66](#)] an, in principle, random 10% sample of tweets authored since 2009. The stream contains a variety of activity types including tweets, retweets, quote tweets, and replies. While the Twitter REST API serves up-to-date information for a given user or activity, the Streaming API data offers a snapshot of each activity at the moment of generation. The specific form of this dataset allows us to create a historical timeline of activities; responses to each original activity arrive with timestamped counts for several metrics including replies, likes, and retweets. For the purposes of this study, we define ‘activity’ to refer to any user action recorded

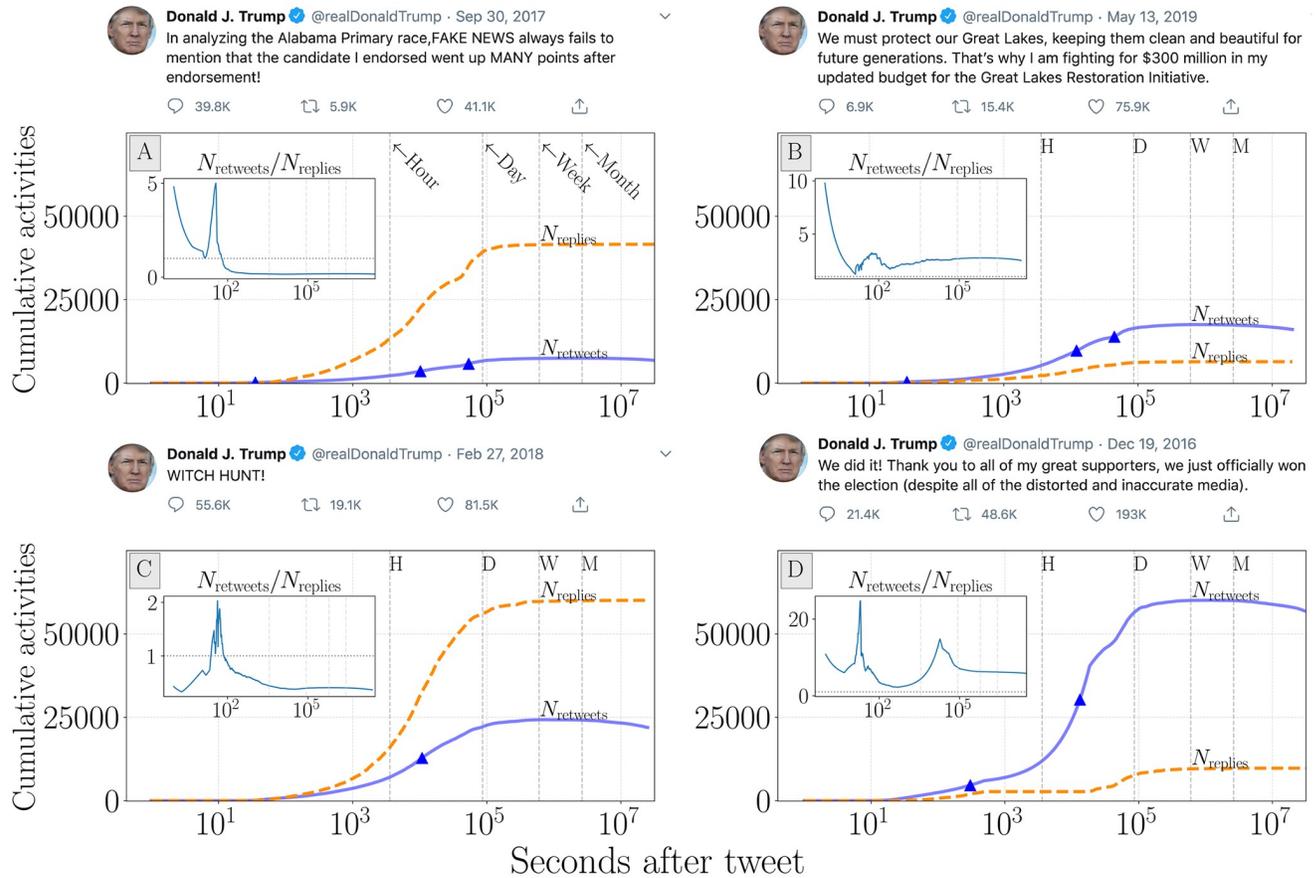


Fig 1. Time series for the cumulative sum of retweet and reply activities for four tweets authored by the Trump account after the 2016 presidential election. The temporal axis is logarithmically spaced to show early-stage growth. Insets represent the cumulative ratio of $N_{retweets}/N_{replies}$ for the same time period. Vertical lines with *H*, *D*, *W*, *M* correspond to hour, day, week, and month intervals. Triangles indicate inflection points where the second derivative of retweet volume transitions from positive to negative (see Sec. 3.2 and Fig 6). The tweets examined here only provide some illustrative examples of how response activity time series behave. Direct links to tweets for panel (A): <https://twitter.com/realDonaldTrump/status/914269704440737792>, panel (B): <https://twitter.com/realDonaldTrump/status/1128051913419837442>, panel (C): <https://twitter.com/realDonaldTrump/status/968468176639004672>, and panel (D): <https://twitter.com/realDonaldTrump/status/810996052241293312>. Tweet screenshots were collected on May 28, 2020.

<https://doi.org/10.1371/journal.pone.0248880.g001>

in the historical sample including original tweets, retweets, retweets with comments, and replies.

Given the sampling rate of the Decahose API, each individual activity has an approximately 10% chance of appearing in our data set. For popular users and tweets (i.e., tweets that garner retweets and replies that number in the thousands or more), we are almost certain to record activities responding to the original activity. Moreover, we are able to observe a number of activities with sufficient temporal resolution to construct historical time series down to the level of a single second.

For this study, we are looking at 2 very popular users on Twitter. Their tweets regularly receive sufficient response-activity volume for our sampling method to be effective. It is worth noting that our method would yield noisier estimates for less popular accounts (and tweets).

2.2 Tracking retweets and replies

Starting with the 10% random sample of Twitter activities, we collect activities responding to presidential tweets by filtering for replies and retweets responding to the @BarackObama and

@realDonaldTrump accounts. We do not include the official US Presidential Twitter account (@POTUS) or the official White House account (@WhiteHouse). The random sample feed serves original tweets, retweets, and replies. Each of these activities are served in a tweet object that contains metadata which includes creation time and follower counts for the author of the activity. In the case of retweets, the tweet object contains a count of retweets and likes at the moment the activity is retweeted. Replies are tagged with user metadata and time of activity. Altogether, the metadata provides the historical data allowing construction of the response activity timeline.

Quote tweets appear in our data stream but we do not count them towards our tally of retweets. This is partially motivated by a desire to maintain backwards compatibility of our analysis with earlier periods before quote tweets were an official feature on Twitter. In a preliminary analysis of 20 Trump tweets with high engagement, we determined that for 16 out of 20 tweets the number of replies and quote tweets both out numbered the number of regular retweets. We mention this as reference point but leave in-depth analysis of quote tweets to future research.

2.3 Measuring the ratio

While retweets counts are observable via tweet metadata, replies must be gathered by maintaining a cumulative summation of reply activity counts (multiplying the final value by 10 to account for our sample rate). We collect counts of retweets and likes from the metadata, while replies are tallied using the cumulative sum method. Validating these estimates for final reply values, we find the error to be less than 1% in terms of actual/predicted reply counts for high activity tweets (actual reply counts were collected for a random sample of tweets via the Twitter web-interface).

In order to calculate the ratio at a common time step, we linearly interpolate the values for replies, retweets, and favorites, resulting in time series that can be sampled at arbitrary intervals for all activity types. We report time steps at the second resolution unless otherwise stated.

There are several challenges associated with the historical feed, stemming from contradicting metadata on activity objects that occurred simultaneously. With timestamps at the 1 second resolution, sometimes tweet activities are reported as occurring at the same time (inferring time stamps from Twitter snowflake IDs did not resolve these conflicts [67]). Activities that reportedly occur at the same time often have conflicting values for retweets and likes (sometimes differing by hundreds of activities). This may be due the count values growing so quickly that the activities within a given second window have differing values and/or it may be an artifact of the slow update times within the platform's database. This latter point is made more notable when taken in combination with Twitter's practice of deleting millions of tweets per week (this would lead to fluctuations in the activity counts for a given tweet).

Regardless of the cause, fluctuations in activity counts result in time series that increase non-monotonically. To make model fitting and data analysis easier, we remove observations that result in non-monotonically increasing behavior before the maximum value of the activity time series. When calculating first and second derivatives we use a Gaussian filter to smooth the time series so that we avoid incorrectly identifying noisy regions as notable transitions from positive to negative derivatives. There is decay in values for retweets while replies only increase—the exact reasons for the decay are not investigated here, but the lack of decay in the replies is due to the cumulative sum method we use for estimating reply values. There is further work to be conducted on the possibility that these jagged portions are a signal of banned account activity, and the dynamics of the jagged regions may indicate activity by accounts that the platform ultimately deems ban-worthy.

Ternary activity values are calculated by dividing each activity count by the sum of all activities at a given time step. The ternary ratio value for activity type τ at time step t is given by

$$\mathcal{R}_\tau(t) = \frac{N_\tau(t)}{N_{\text{retweets}}(t) + N_{\text{likes}}(t) + N_{\text{replies}}(t)}, \tag{1}$$

where $N_\tau(t)$ is the count of the activity at time step t . With the above each observation is comprised of a 3-dimensional vector representing the normalized activity values for a tweet.

The ternary ratio values are represented on a ternary plot (2-dimensional simplex) where the values of activities at each time step sum to 1,

$$\sum_\tau \mathcal{R}_\tau(t) = 1. \tag{2}$$

2.4 Vocabulary of ratioed tweets

We are also interested in how the text content of tweets relates to ratios of response activities. To investigate this we calculate rank turbulence divergence—as defined by Dodds *et al.*—between ratioed and non-ratioed tweets [68].

The rank turbulence divergence between two sets, Ω_1 and Ω_2 , is calculated as follows,

$$\begin{aligned} D_\alpha^R(\Omega_1 || \Omega_2) &= \sum_\tau \delta D_{\alpha,\tau}^R \\ &= \frac{\alpha + 1}{\alpha} \sum_\tau \left| \frac{1}{r_{\tau,1}^\alpha} - \frac{1}{r_{\tau,2}^\alpha} \right|^{1/(\alpha+1)}, \end{aligned} \tag{3}$$

where $r_{\tau,s}$ is the rank of element τ (n -grams in our case) in system s and α is a tunable parameter that affects the impact of starting and ending ranks.

While there are numerous techniques to compare two corpora, we use rank divergence in light of its robustness when working with subsamples of heavy tailed distributions. Working with ranked values also alleviates some challenges that may arise when normalizing n -gram occurrences to arrive at frequency of usage. Further, rank divergence gracefully deals with the problem of n -grams only occurring in one of the two corpora. Finally, ranked values make interpretation more straightforward when presenting results (with the reader need not worry about normalization, etc.). These factors combined with the tunability of rank divergence lead us to find it was best suited for our use case.

Here, we take the content of tweets from the Trump account after his declaration of candidacy on June 16, 2015. We then split this set by ratioed ($N_{\text{retweets}}/N_{\text{replies}} < 1$) and non-ratioed tweets ($N_{\text{retweets}}/N_{\text{replies}} > 1$). Obama’s account did not have a sufficient number of ratioed tweets to conduct this analysis. From here we calculate the rank divergence, between the two sets. We set $\alpha = 1/3$ based on experimentation outlined in the original rank divergence piece by Dodds *et al.* This value of alpha tends to balance the influence of high- and low-ranked items.

3 Results

3.1 Overall time series

Viewing replies, retweets, and likes for individual tweets over time, we see how activities in response to Trump and Obama tweets have changed in the years around the 2016 election. In Fig 2, we show the final count (the activity volume 168 hours after the original tweet is authored) of activities for each tweet in the POTUS data set. Then-candidate Trump’s tweets

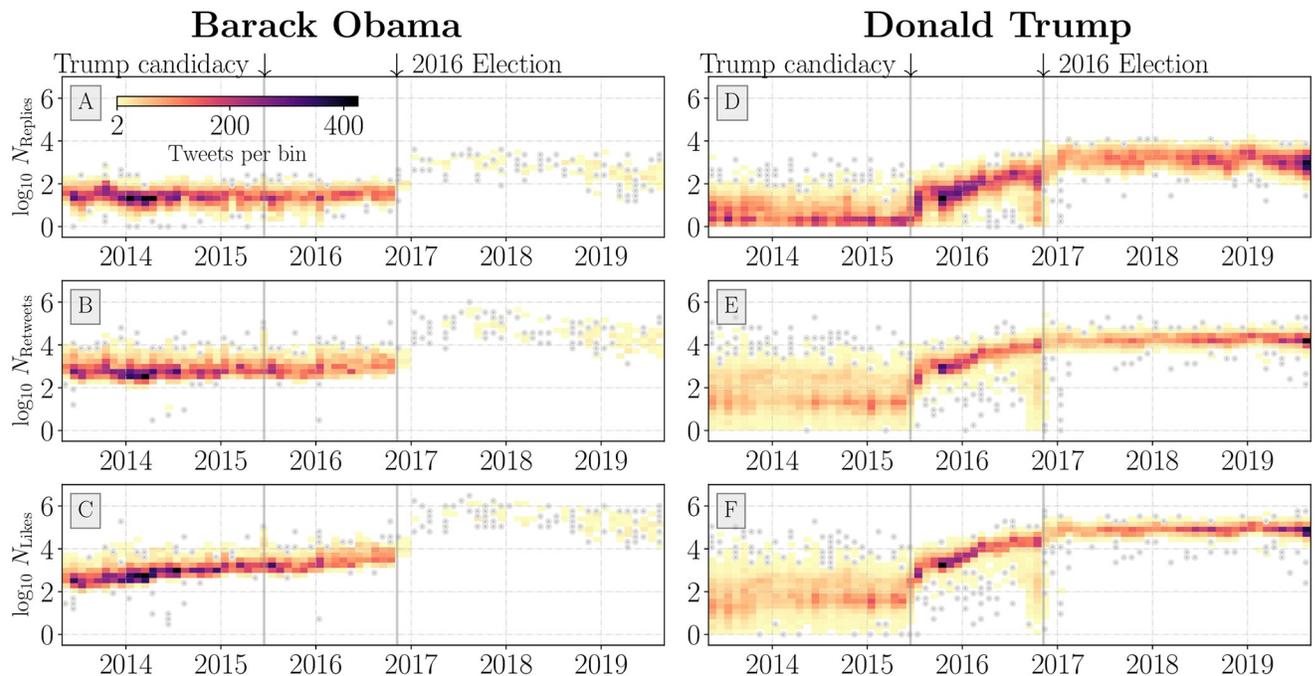


Fig 2. Time series histogram of maximum activity counts for tweets posted from the Barack Obama (A–C) and Donald Trump (D–F) Twitter accounts. Each bin represents a collection of tweets at their original author date along with their respective maximum observed activity count. Bins with less than 2 tweets are shown as grey dots. We include all tweet types (e.g., advertisements, promoted, etc.); see Section 2.1 for information on collection methods. We annotate Trump’s declaration of candidacy and the 2016 US general election with solid vertical grey bars. A marked decrease in Obama account activity is apparent immediately following the 2016 election. The region of outliers in the Trump time series immediately preceding the 2016 election has been determined to be largely reflective of promoted tweets which have abnormal circulation dynamics on the platform [69].

<https://doi.org/10.1371/journal.pone.0248880.g002>

experienced a notable increase in response activity following his June 16, 2015 declaration of candidacy. Prior to this point, Trump’s tweets garnered 2 to 3 orders of magnitude less activities than during the height of his campaign and his subsequent time in office.

Because of the difficulty in sampling replies using our data (see Sec. 2.1 for tweets that receive few replies), many of the tweets from the Trump account from before his candidacy have a high degree of uncertainty in their true value. This uncertainty is introduced by our method for inferring reply counts introduces variance in the overall activity balance.

In Fig 3 we present the normalized activity values for tweets from Obama and Trump from 2013 to 2019. The ternary histograms allow us to compare all activity values, while the time series in Fig 3D and 3H show the retweet-to-reply ratio. We choose the latter as it appears to offer a more descriptive measure of response activities (i.e., counts of likes are generally more stable within sub-regions of the yearly time series). The Obama $N_{retweets}/N_{replies}$ time series (Fig 3D) demonstrates how the Obama account tends to receive more retweets than replies (median ratio value of 3.67 between June 6, 2015 and November 8, 2016).

Once Obama leaves office, the account receives consistently more retweets than replies on the limited number of tweets authored (median ratio value of 7.77 after November 8, 2016).

The Trump $N_{retweets}/N_{replies}$ time series (Fig 3H) shows the tendency of then-candidate Trump’s account to be ratioed less during the campaign season. Soon after transitioning to office the Trump account begins receiving more replies relative to retweets—with a transition period roughly corresponding to the time between the day of the election and inauguration. For the campaign period Trump has a median ratio value of 3.1 while after the election the account garners a median ratio value of 1.32.

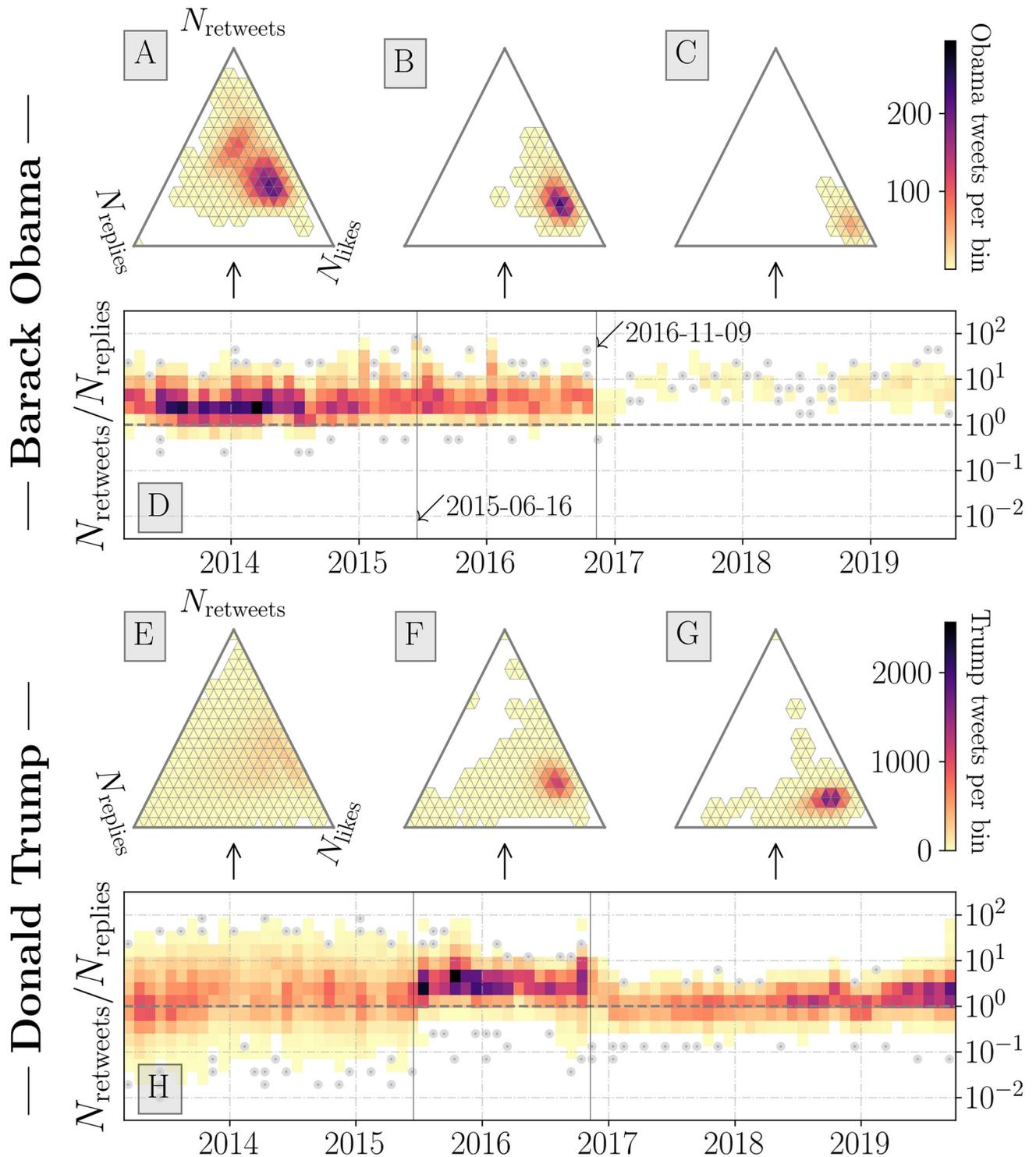


Fig 3. Ternary histograms and $N_{retweets}/N_{replies}$ ratio time series for the @BarackObama (A–D) and @realDonaldTrump (E–H) Twitter accounts. The ternary histograms (A–C and E–H) represent the count of retweet, favorite, and reply activities normalized by the sum of all activities. White regions indicate no observations over the given time period. See Fig 4 for examples of full time series for response activity for example tweets. Heatmap time series (D and H) consist of monthly bins representing the density of tweets with a given ratio value. Single observations (bin counts <2) are represented by grey points. The two dates annotated correspond to the date of Trump’s declaration of candidacy (2015–05–16) and the 2016 general election (2016–11–09). We show the tendency for Trump tweets to have ternary ratio values with a greater reply component—with pre-candidacy tweets having higher variability and pre-election tweets having a higher $N_{retweets}/N_{replies}$ ratio value. Post-election Obama tweets have ternary ratio values with more likes than other periods for both Obama and Trump.

<https://doi.org/10.1371/journal.pone.0248880.g003>

In Fig 3E we show a ternary histogram for the pre-campaign Trump account, where the high variance in reply estimates are evidenced by broad coverage of most regions of the simplex. Error in estimating replies alone does not account for this high variance, with the like and retweet time series from Fig 2 showing higher variance for Trump as well. Another time series artifact is the presence of Trump tweets that were met with low activity counts around the 2016 election. Around the same time, the Obama account activity drops precipitously. After the election, Obama's limited number of tweets are met with consistently high counts of likes, retweets, and replies. The rapid increase in activities in response to Trump tweets, and the corresponding decrease in the overall variance of counts for activities, are important insights visible in Fig 2. These measures are indicative of the meteoric rise of then-candidate Trump, along with his now pre-eminent Twitter presence (his name now appears more frequently than the word "god" on most days [64]). While the two time periods are not directly comparable, by the end of the 2016 election, Trump's account consistently garnered more response activities than President Obama's. After the 2016 election, the Obama account's tweeting frequency is reduced while also experiencing a notable rise in response activities. These results helped inform the selection of distinct periods around the 2016 election. These time periods were both meaningful in a political context—marking Trump's declaration of candidacy and the 2016 election day—as well as in the context of response activity time series.

3.2 Example ternary time series

Using observations of responses to historical tweets, we can construct the time series for all ratios of activities. In Fig 4 we show the ratio time series for three Trump tweets. We selected the tweets based on the value of their final retweet-to-reply ratio—with a tweet from approximately the 90th, 50th, and 10th percentiles of final $N_{\text{retweets}}/N_{\text{replies}}$ ratio values. The response activities to most tweets (with high response activity) tend to experience some early volatility, partially owing to the low number of observations. One hour after the original tweet has been authored, the response ratios tend to fall into a stable region, or at least adopt a stable trend. Within this signal there is also the effect of bots (S2 Appendix) that are likely programmed to respond to Trump account activity within seconds.

We show how ratios tend to stabilize by presenting ternary histograms of activity response ratios over the seconds and days after a tweet is published. S1 Fig shows how the Obama account largely has ratios biased towards retweets and likes throughout the period after a tweet is authored. Ratios for the Trump account (Fig 4) tend to have greater variance in their ratio values in the seconds and hours after a tweet is released. The ratios for Trump's account are also biased towards more replies, relative to the Obama account, throughout the period after issuing a tweet. This is consistent with the final ratio values for each president across three political periods surrounding the 2016 election.

3.3 Distribution of ratios

The final observed ratio value for a tweet offers an aggregate measure of user-base reactions. We examine tweets which were authored at least 168 hours prior to the activity calculation to ensure the responses have largely stabilized. Of course, there is still the possibility that users will respond to the tweet long after the original activity period, but we have found these lagging response activities to minimally affect the normalized ratio value. It is worth noting that this delayed response behavior is common for prominent users such as Obama and Trump, with some response activities taking place months or years after the original activity (often in reference to current events that are addressed in the old tweet: 'there is always a tweet').

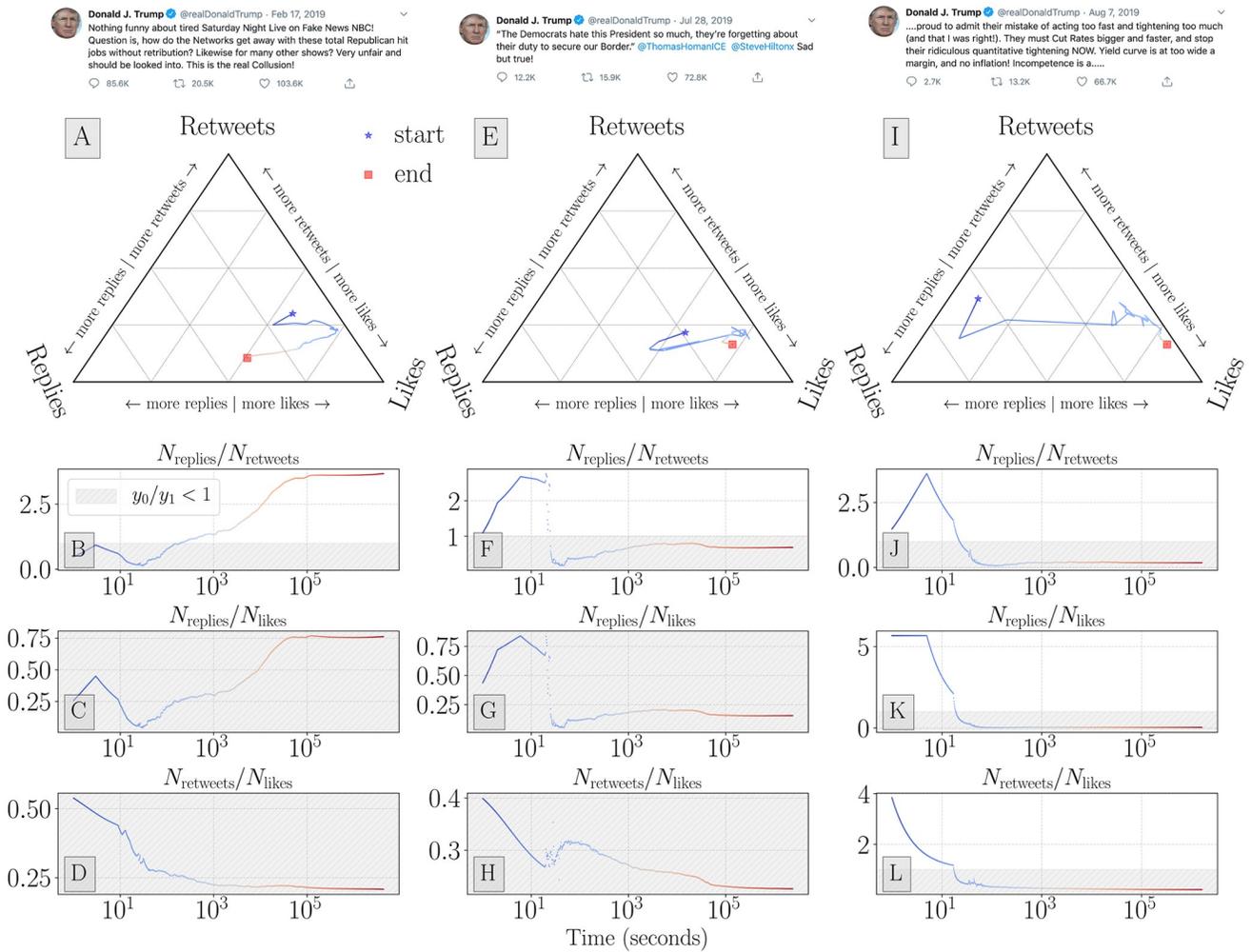


Fig 4. Ternary time series for three popular Trump tweets, selected to represent messages in approximately the upper 90th percentile, 50th percentile, and bottom 10th percentile of $N_{retweets}/N_{replies}$ ratios. The time series represent observations from first activity observation to the final observation of the tweet. These ternary time series contrast the simple 2-dimensional ratio trajectories and illustrate example trajectories through the 3-dimensional ratio space. See S1 and S2 Figs for the distribution of ternary ratio values for Obama and Trump tweets over time. Each of these three tweets were authored on May 25, 2019. Direct links to tweets for panels (A–D) <https://twitter.com/realDonaldTrump/status/1097116612279316480>, panels (E–H) <https://twitter.com/realDonaldTrump/status/1155657137076527104>, and panels (I–L) <https://twitter.com/realDonaldTrump/status/1159083364965654528>. Screenshots were collected on May 28, 2020.

<https://doi.org/10.1371/journal.pone.0248880.g004>

In Fig 5 we show the distribution of Trump and Obama $N_{retweets}/N_{replies}$ ratios, highlighting the tendency for the Trump account to “get ratioed” more often relative to the Obama account. Of 13,639 tweets in Fig 5 from the Trump account, 18% receive more replies than retweets. The Obama account generates a relatively limited number of tweets that receive more replies than retweets (less than 1%, see S4 Fig for Obama’s ratioed tweets). This simple ratio calculated with two of the three activity measures loses some context in terms of overall user-base responses, and we now move to incorporate the like activity volumes.

We subset the tweets for Presidents Obama and Trump based on the time periods introduced in Fig 2. In Fig 3 we show that for both Obama and Trump accounts there is a higher degree of variation in final ternary ratio values for earlier years. This is likely in part due to the reply thread mechanisms and structure changing on Twitter’s platform, as well as the growing popularity of political accounts on Twitter, aside from background changes in how users

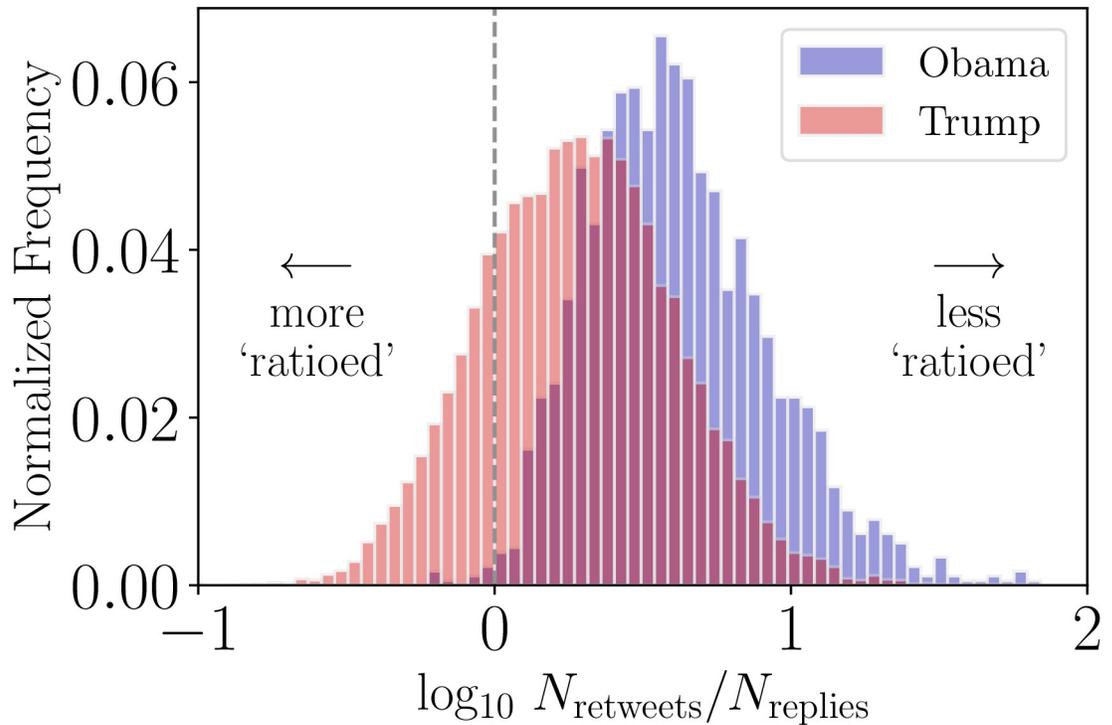


Fig 5. Histogram of final observed ratios for tweets authored by @BarackObama and @realDonaldTrump accounts. Observations left of the dashed vertical line correspond to tweets that are considered ratioed. The shift in the distribution corresponding to Trump’s account can be plainly seen—with the account producing more tweets that are ratioed than compared with Obama’s account. For both accounts, tweets shown here are restricted to those authored after Trump’s declaration of candidacy. Of 16,708 tweets included from the Trump account, 3,015 (18%) have a $N_{retweets}/N_{replies}$ value less than or equal to 0 and 13,693 (82%) have a score greater than 0. Of 1,786 tweets from the Obama account, 7 (<1% have $N_{retweets}/N_{replies}$ values ≤ 0 while 1,779 (>99%) have values > 0 . From the distribution of ratio values we see that ratioed tweets are outliers for the Obama account while the Trump account often gets ratioed and has a lower ratio value on average.

<https://doi.org/10.1371/journal.pone.0248880.g005>

engage on the platform. This shift is especially prominent for the Trump account, which experienced a marked growth in response activity after Trump’s declaration of candidacy. The higher variance of final ratio values is accentuated when the sample of replies is lower, owing to our method of estimating reply activities. Comparing Obama and Trump ternary histograms (Fig 3) we observe more Trump activity ratios in the reply region of the simplex.

For the Obama account, Fig 3A–3D, we see notably higher values for normalized retweet and like activity compared with the Trump account. The retweet-to-reply ratio is consistently higher in the time series presented in Fig 3D. Once leaving office, the Obama account demonstrated a further tendency to receive a higher proportion of likes and retweets.

3.4 Characteristic time scale

To empirically investigate the characteristic time scale of response activities, we report the time index t for the local maxima of the instantaneous retweet count for each tweet (the local maxima of the first derivative of cumulative retweets $N_{retweets}$, or $\text{argrelmax } \frac{d}{dt} N_{retweets}$). More specifically, we use the second derivative test to find points where the second derivative of cumulative retweet counts drops below 0,

$$\frac{d^2}{dt^2} N_{retweets}(t_{i-1}) > 0 \text{ and } \frac{d^2}{dt^2} N_{retweets}(t_i) < 0. \tag{4}$$

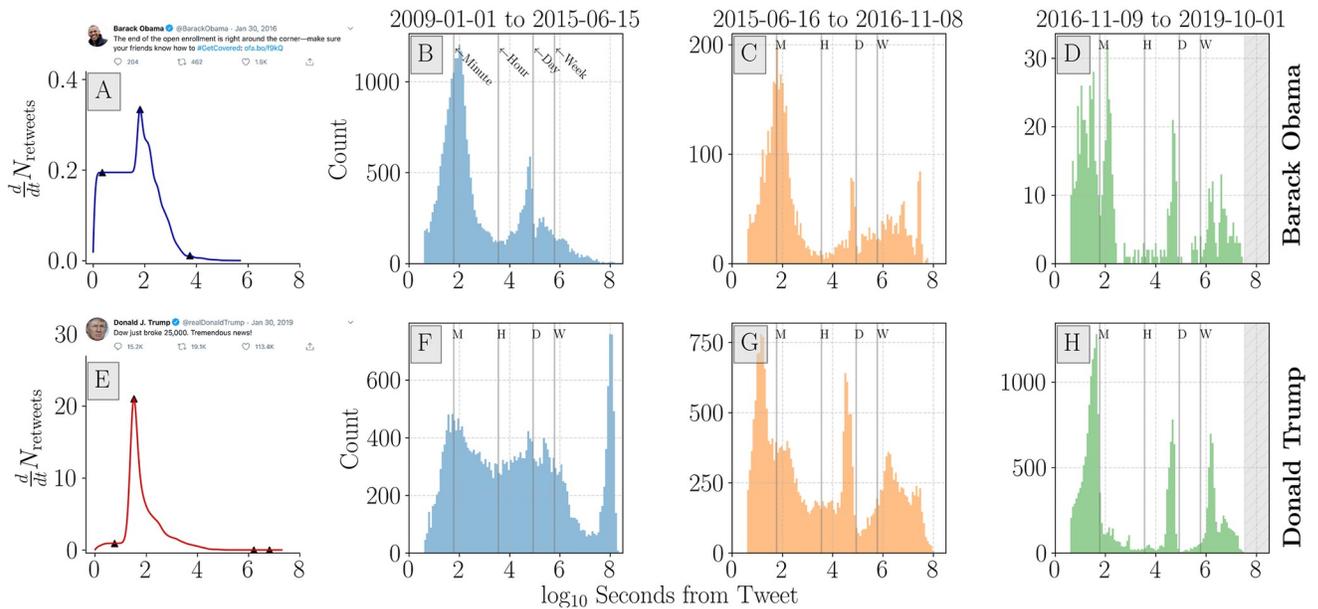


Fig 6. Distribution of inflection points, the local maxima of instantaneous retweet volume, $\text{argrelmax} \frac{d}{dt} N_{\text{retweets}}$, where $\frac{d^2}{dt^2} N_{\text{retweets}}(t_{i-1}) > 0$ and $\frac{d^2}{dt^2} N_{\text{retweets}}(t_i) < 0$. A and E: Example cumulative retweet time series and inflection points (solid triangles) for Obama and Trump tweets. Histograms show the distribution of inflection points across all tweets binned by time periods before (B and F), during (C and G), and after (D and H) the 2016 US presidential election campaign. The January 1st 2009 to June 15th 2015 period for Trump (F) contains inflection point counts that are largely reflective of low initial activity and high(er) late activity (months or years later) leading to unusually high values for seconds to first inflection point ($>10^8$ seconds). For Obama’s and Trump’s time in office, tweets experience inflection points around 1-minute and 1-day after the tweet is authored—indicating characteristic time-scales of activity waning. Histogram bin widths are in effect logarithmically spaced over the displayed time span. This has the effect of providing a higher temporal resolution for shorter timer periods. Direct links for Obama tweet (A): <https://twitter.com/BarackObama/status/693571153336496128> and Trump tweet (E): <https://twitter.com/realDonaldTrump/status/1090729920760893441>. Screenshots were collected on May 28, 2020.

<https://doi.org/10.1371/journal.pone.0248880.g006>

These points are viewed as being indicative of the start of a decreasing trend for the volume of new response activity, specifically marking the moment when the rate at which new activities are generated begins decreasing. Going forward we will refer to these points as inflection points (examples of these points can be seen with triangular markers in Fig 1).

A similar measure could examine the point where the second derivative rises above 0. This would be a valid measure, but we prefer our measure for a couple of reasons. First, we found it better characterizes the time scale of the initial burst of activity that most POTUS tweets receive (i.e., how long the first wave of response activities is sustained). Second, it provides a closing bound on a period of increased response activity (i.e., the point indicates when activity starts to fade). We were more interested in characterizing when attention, as indicated by response activity, is no longer maintained for the POTUS tweets. We include results for “inverse-inflection” points in S3 Fig as a point of reference.

When viewing time since an original tweet in Fig 6, we see a higher density of inflection points around the 1 minute and 1 day time intervals. This suggests that many periods of intense activity take place within the first day of a tweet being authored. We also observe inflection points in cases where more sporadic engagement occurs later in the tweet response-activity timeline. This leads to another spike in activity around 1 week after the original tweet. Over 90% of inflection points are observed before 10^5 seconds (~ 1.15 days) for both accounts and for all time periods as evidenced by complementary cumulative distribution functions (see Fig 7).

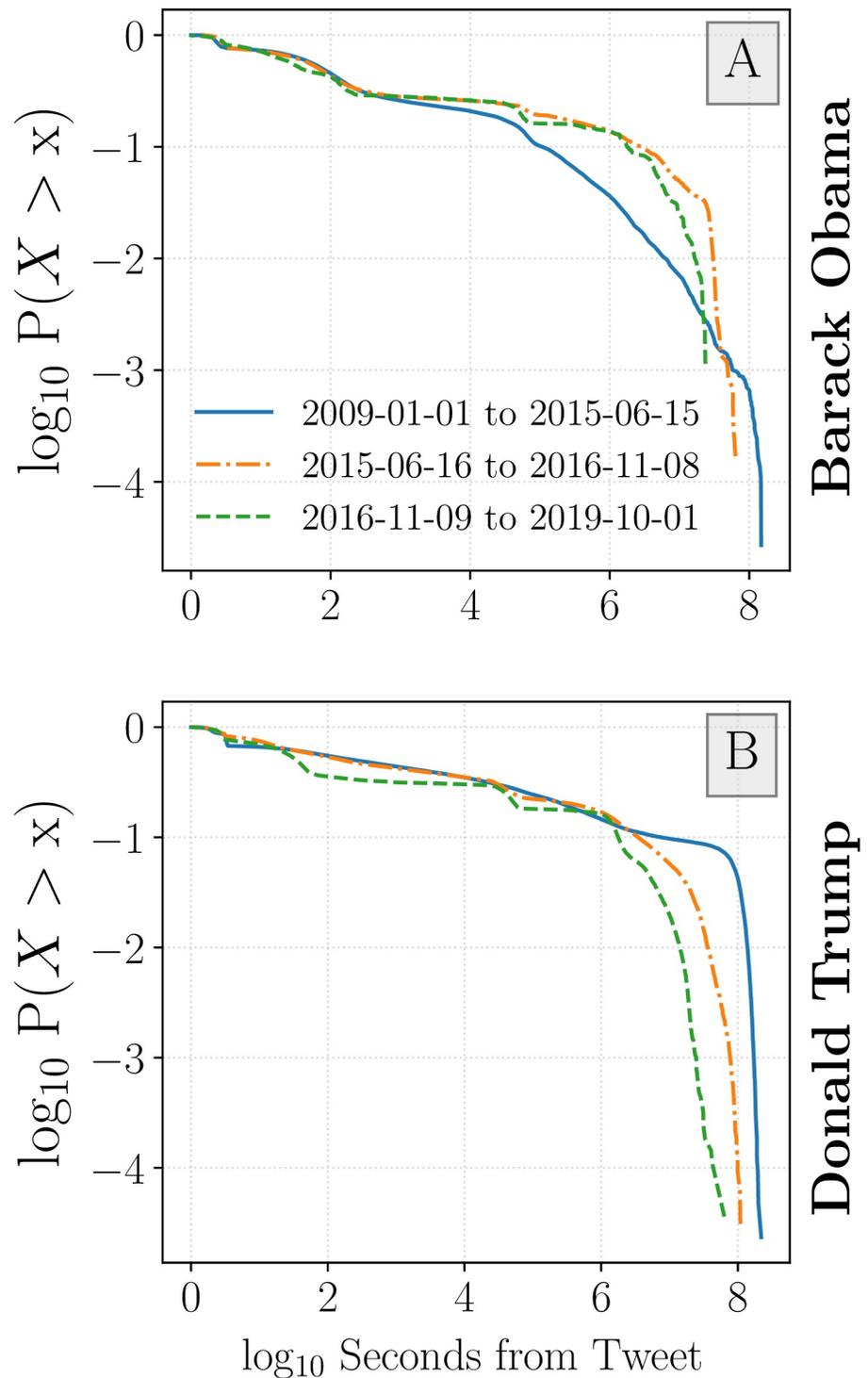


Fig 7. Complementary cumulative distribution function for inflection point timing. Roughly 90% of Obama inflection points (A) took place before 10^5 seconds (~ 1 day) for the period before the 2016 election cycle. For the period during and after the 2016 campaign season, both Obama and Trump tweets (B) receive 10% of inflection points after roughly 10^6 seconds (~ 10 days) from the initial tweet.

<https://doi.org/10.1371/journal.pone.0248880.g007>

There is a notable difference between the governing periods for Obama and Trump, Fig 6A and 6G, respectively. Trump tweet response time lines demonstrate a tendency to generate inflection points before the 1-minute mark—indicative of rapid response to his account’s tweets. Additionally, we find more time series inflection points after 1-week has passed for the Trump time series than compared with Obama. This suggests that users may return to Trump tweets later to comment on the content (perhaps as new political developments occur). This is further illustrated in Fig 7B, where we show that 10% of Trump inflection points take place after 10^6 seconds (~ 10 days). This is true for the pre-campaign, campaigning, and governing periods. By contrast, 90% of Obama inflection points take place before 10^5 seconds (~ 1 day).

3.5 Ratio content

We can investigate how tweet content relates to response activities by comparing the distributions of words that appear in ratioed and non-ratioed tweets. Taking the rank of the frequency of occurrence for each group, we then compare the rank-turbulence divergence for the two groups [68]. Per the allotaxonograph in Fig 8, we are able to tune the impact of starting and ending rank magnitude on the overall score for the word. The allotaxonograph is an

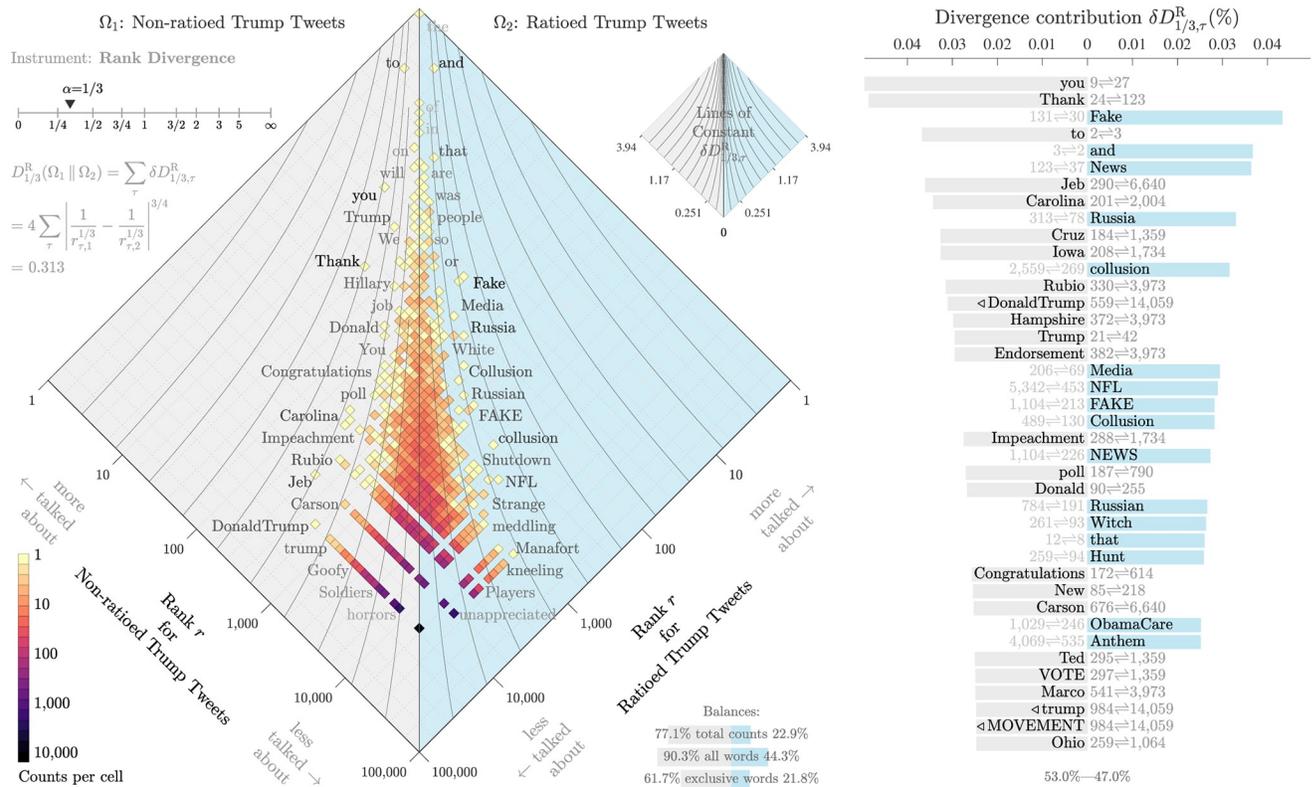


Fig 8. Rank divergence allotaxonograph [68] for 1-grams from Trump account tweets authored after Trump’s declaration of his candidacy on June 16, 2015. “Ratioed” tweets are those where the proportion of retweets over replies is less than 1 ($N_{retweets}/N_{replies} < 1$), non-ratioed tweets accumulated a ratio greater than 1 ($N_{retweets}/N_{replies} > 1$). There is a notable class imbalance. The majority of tweets are non-ratioed, with a median value of ~ 2 . To generate the above figure we examined 3,313 ratioed tweets and 15,274 non-ratioed tweets. The 1-grams “Fake”, “News”, “Russia” and “NFL” are ranked higher in the ratioed corpus. For the non-ratioed corpus, the 1-grams “Jeb”, “Carolina”, and “Ted” are ranked higher. This illustrates the tendency of ratioed tweets from this period to contain more politically contentious 1-grams related to Trump scandals. Non-ratioed tweets from this period more often contain campaign related messages. In the main horizontal bar chart, the numbers next to the terms represent their rank in each corpus, while terms that appear in only one corpus are indicated with a rotated triangle. The smaller three vertical bars describe the balances between the ratioed and non-ratioed corpora: 77.1% of total counts occur in the non-ratioed corpus; we observe 90.3% of all 1-grams in the non-ratioed corpus; and 61.7% of 1-grams in the non-ratioed corpus are unique to that corpus.

<https://doi.org/10.1371/journal.pone.0248880.g008>

instrument that allows us to compare the Zipf distributions of two corpora. The figure features a rank-rank histogram and a vertical bar plot describing each 1-grams' contribution to the overall rank divergence value.

In Fig 8, we show that ratioed tweets contain more words related to fake news, the Mueller inquiry into Russian interference, and Colin Kaepernick and kneeling during the national anthem in the National Football League (NFL). For non-ratioed tweets, we see words from campaign-related communications (e.g., “Jeb”, “Iowa”, and “VOTE”) tend to appear more often than in ratioed tweets. The imbalance of word counts in the ratioed and non-ratioed tweets is roughly proportional to the tweet imbalance with 22.9% of words occurring in the ratioed tweets. Of all unique words, nearly twice as many occur once or more in the non-ratioed tweets compared to the ratioed tweets.

4 Discussion

We have considered how both the volume and kind of user activities in response to US presidents on Twitter varies over multiple time scales. We found the Trump account to have greater variability in the normalized ratio of activities—including a tendency to receive more replies relative to likes and retweets than when compared to the Obama account. Obama's tendency to receive more retweets and likes was only amplified after his departure from public office. We found that in ratioed tweets authored by Trump, words pertaining to fake news, the Mueller inquiry, and the NFL are more common than in non-ratioed tweets. We also showed more general results for response activity profiles, with responses to the two Twitter presidents often stabilizing around the 1-day scale. The Trump account also experiences more activity fluctuations after 1-week than compared to the Obama account, perhaps owing to users more often returning to Trump's comments as political actions unfold.

How the public responds to messages over the course of a politician's career is a fundamental dynamic in politics. As candidates rise to prominence and are elected to office, the reach and importance of their communication changes. Further, the rapid proliferation of digital technology provides a backdrop of constant evolution in political communications. Response activity by users on social media provides a valuable set of features for gauging how social networks respond to political messaging. Beyond simple counts, response time series dynamics provide insight into the characteristics of political engagement. This includes determining when activity counts stabilize and perhaps establishing how polarized push-and-pull dynamics play out in response to messages. We speculate these techniques may also be useful for detecting non-human user behavior in response to tweets. When combined with more conventional natural language processing techniques, these “ratiometrics” hold promise in improving our understanding of how specific messaging (e.g., word choice) affects user responses.

This study was not able to control for background changes in the Twitter interface—changes to the manner in which retweets are served could have effects on the response profiles. Further, there are some types of posts—namely “promoted” or “ad” posts—that have different distribution characteristics (i.e. may be viewed by a more targeted, limited audience). We were not able to fully filter these posts and they were included in the analysis.

Twitter is a constantly changing system, and this is especially true for the nature of political discourse on the platform. While we make efforts here to highlight some notable exogenous events, we cannot control for background flux of user-engagement with politics. Simple changes that may be occurring include shifts in the distribution of users from across the political spectrum (i.e., more left- or right-leaning user activity) or more subtle changes in content (e.g., more spam and less in-depth conversation). There is also the potential for modifications to Twitter's timeline algorithms to affect how users engage with content through response

activities. Related to this point, we are unsure of algorithmic factors that affect the circulation of inflammatory content on the platform—it is conceivable that increased views of controversial activities could lead to a snowball effect with heated debates taking place between users in the reply section of a tweet.

With social media playing such a prominent roll in today's political process, it is our goal that the methods presented here can contribute to a broader suite of instruments for analyzing political communication in the digital realm. This is the first piece of an effort to systematically evaluate the communications of US presidents on Twitter. Building from our understanding of ratiometrics, the "POTUSometer" is envisioned as sitting alongside instruments such as the Hedonometer [70] (which provides a measurement of collective sentiment on Twitter). The POTUSometer would tap into the 'wisdom of the crowd' (or at least the reality of collective responses) in order to evaluate and better understand how presidents speak and are listened to on social media platforms.

Future research could explore how the sentiment changes with response activities (commented retweets and replies). Similarly, topic modelling could be used to explore which subjects are discussed throughout the response activity time line. A more sophisticated null model for the ternary ratio values (and ratio time series) would be worthwhile to enable anomaly detection. There would be further work in the area of researching attention reinforcement along with analyses of time series stability. With recent advancements in language modelling, text regression on tweet content in order to predict ratio scores may be worthy of investigation. Finally, replicating this work across a broad base of users beyond Obama and Trump would better inform how certain social network attributes effect the above results.

Supporting information

S1 Appendix. Ternary histogram snapshots.

(PDF)

S2 Appendix. Brief note on bots.

(PDF)

S1 Fig. Snapshots of ternary activity ratio values for tweets in response to the Obama account. Observations are recorded at logarithmically spaced intervals after the release of the original tweet. Included here are 3,015 tweets from the period of time after Trump's declaration of candidacy on June 16, 2015. Whereas Fig 3 shows a final ratio value for tweets over distinct political periods, here we show how ratios unfold over the life of each tweet. This is serves as a snapshot of the ternary ratio time series presented in Fig 4. Obama's tweets tend to receive a greater volume of likes and retweets relative to replies, and this ratio is often maintained throughout the response timeline.

(TIF)

S2 Fig. Snapshots of ternary activity ratios for tweets in response to the Trump account. Observations are recorded at logarithmically spaced intervals after the release of the original tweet. Included here are 16,708 tweets from the period of time after Trump's declaration of candidacy on June 16, 2015. Whereas Fig 3 shows a final ratio value for tweets over distinct political periods, here we show how ratios unfold over the life of each tweet. This is serves as a snapshot of the ternary ratio time series presented in Fig 4. We can see the greater variation in ternary ratio values compared to the snapshots for Obama (Fig 4). There is also a greater tendency for tweets to have higher reply counts when compared to the Obama snapshots.

(TIF)

S3 Fig. Distribution of inverse-inflection points, the local minima of instantaneous retweet volume, $\text{argrelmin} \frac{d}{dt} N_{\text{retweet}}$, where $\frac{d^2}{dt^2} N_{\text{retweet}}(t_{i-1}) < 0$ and $\frac{d^2}{dt^2} N_{\text{retweet}}(t_i) > 0$. Here we switch the criteria from Fig 6 in order to show how our choice affects the results for the characteristic time scale. For the results in this figure we show points where the second derivative of the retweets activity time series goes from negative to positive. **A** and **E**: Example cumulative retweet time series and inverse-inflection points (solid triangles) for Obama and Trump tweets. Histograms show the distribution of inverse-inflection points across all tweets binned by time periods before (**B** and **F**), during (**C** and **G**), and after (**D** and **H**) the 2016 US presidential election campaign. The January 1st 2009 to June 15th 2015 period for Trump (**F**) contains inflection point counts that are largely reflective of low initial activity and high(er) late activity (months or years later) leading to unusually high values for seconds to first inflection point ($>10^8$ seconds). For Obama's and Trump's time in office, tweets experience inflection points around 1-minute and 1-day after the tweet is authored—indicating characteristic time-scales of activity waning. Direct links for Obama tweet (**A**): <https://twitter.com/BarackObama/status/693571153336496128> and Trump tweet (**E**): <https://twitter.com/realDonaldTrump/status/1090729920760893441>. Screenshots were collected on May 28, 2020. (TIF)

S4 Fig. Obama's ratioed tweets from the study period. Here we show the instances where we observed Obama's tweets garnering more replies than retweets. For the tweets in panels (B, D and G) the response activities later changed to reflect fewer replies than retweets. This behavior is consistent with what we have seen in other cases, with activity shifts either owing to authors deleting the activity or responding-accounts being deleted. Here it seems Obama's tweets are a combination of controversial (the Iran nuclear deal) and engagement-seeking (soliciting entries in a contest to meet the Obama). Direct links to tweets for panel (**A**): <https://twitter.com/BarackObama/status/733055491664797696>, panel (**B**): <https://twitter.com/BarackObama/status/639553864006430721>, panel (**C**): <https://twitter.com/BarackObama/status/794926969829920768>, panel (**D**): <https://twitter.com/BarackObama/status/666396723191808000>, panel (**E**): <https://twitter.com/BarackObama/status/652914197454561281>, panel (**F**): <https://twitter.com/BarackObama/status/732691885139984384>, and panel (**G**): <https://twitter.com/BarackObama/status/715194440487309312>. Tweet screenshots were collected on November 15, 2020. (TIF)

Acknowledgments

We greatly appreciate the many helpful conversations with members of the Computational Story Lab. We are also grateful for the computational resources provided by the Vermont Advanced Computing Core. We thank David Rushing Dewhurst and Jane Adams for feedback on analytical approaches and data visualization.

Author Contributions

Conceptualization: Joshua R. Minot, Peter Sheridan Dodds.

Data curation: Joshua R. Minot, Michael V. Arnold, Thayer Alshaabi.

Formal analysis: Joshua R. Minot, Peter Sheridan Dodds.

Investigation: Joshua R. Minot, Peter Sheridan Dodds.

Methodology: Joshua R. Minot, Peter Sheridan Dodds.

Supervision: Christopher M. Danforth, Peter Sheridan Dodds.

Validation: Joshua R. Minot, Christopher M. Danforth, Peter Sheridan Dodds.

Visualization: Joshua R. Minot, Michael V. Arnold, Thayer Alshaabi, Peter Sheridan Dodds.

Writing – original draft: Joshua R. Minot.

Writing – review & editing: Joshua R. Minot, Christopher M. Danforth, Peter Sheridan Dodds.

References

1. Frantzych SE. Presidents and the Media: The Communicator in Chief. London: Routledge; 2018.
2. Kernell S. Going Public: New Strategies of Presidential Leadership. 3rd ed. Washington, D.C: CQ Press; 1997.
3. Twitter. Target—Twitter Developers; 2019. Available from: <https://developer.twitter.com/en/case-studies/target>.
4. Twitter. Political Content; 2019. Available from: <https://business.twitter.com/en/help/ads-policies/prohibited-content-policies/political-content.html>.
5. Isaac M. The ratio establishes itself on Twitter. The New York Times. 2018;.
6. O’Neil L. How to know if you’ve sent a horrible tweet. Esquire. 2017;.
7. Merriam-Webster. Words we’re watching: ‘ratioed’;. <https://www.merriam-webster.com/words-at-play/words-were-watching-ratio-ratioed-ratioing>.
8. Data for Progress. The Ratio Richter Scale; 2019. Available from: <https://www.dataforprogress.org/the-ratio-richter-scale>.
9. Mounk Y. Only Cowards Don’t Get Ratioed; 2019. Available from: <https://www.theatlantic.com/ideas/archive/2019/10/ratio-twitter-good/600334/>.
10. Victoria G. Is the Twitter ratio getting out of hand?; 2019. Available from: <https://www.pulsarplatform.com/blog/2019/is-the-twitter-ratio-getting-out-of-hand/>.
11. Roeder O, Mehta D, Wezerek G. The Worst Tweeter In Politics Isn’t Trump. FiveThirtyEight. 2017;.
12. Alashri S, Kandala SS, Bajaj V, Ravi R, Smith KL, Desouza KC. An analysis of sentiments on Facebook during the 2016 U.S. presidential election. In: 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM); 2016. p. 795–802.
13. Ko J, Kwon HW, Kim HS, Lee K, Choi MY. Model for Twitter dynamics: Public attention and time series of tweeting. Physica A: Statistical Mechanics and its Applications. 2014; 404:142–149. <https://doi.org/10.1016/j.physa.2014.02.034>
14. Kobayashi R, Lambiotte R. TiDeH: Time-dependent Hawkes process for predicting retweet dynamics. In: Tenth International AAI Conference on Web and Social Media; 2016. Available from: <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM16/paper/view/13026>.
15. Schmidbauer H, Rösch A, Stieler F. The 2016 US presidential election and media on Instagram: Who was in the lead? Computers in Human Behavior. 2018; 81:148–160. <https://doi.org/10.1016/j.chb.2017.11.021>
16. Howard PN, Duffy A, Freelon D, Hussain MM, Mari W, Maziad M. Opening closed regimes: What was the role of social media during the Arab Spring? Rochester, NY: Social Science Research Network; 2011. ID 2595096. Available from: <https://papers.ssrn.com/abstract=2595096>.
17. Tufekci Z, Wilson C. Social media and the decision to participate in political protest: Observations from Tahrir Square. Journal of Communication. 2012; 62(2):363–379. <https://doi.org/10.1111/j.1460-2466.2012.01629.x>
18. Ruck DJ, Rice NM, Borycz J, Bentley RA. Internet Research Agency Twitter activity predicted 2016 U. S. election polls. First Monday. 2019; 24(7). <https://doi.org/10.5210/fm.v24i7.10107>
19. Garrett RK. Social media’s contribution to political misperceptions in U.S. Presidential elections. PLOS ONE. 2019; 14(3). <https://doi.org/10.1371/journal.pone.0213500> PMID: 30917154
20. Guess A, Nagler J, Tucker J. Less than you think: Prevalence and predictors of fake news dissemination on Facebook. Science Advances. 2019; 5(1). <https://doi.org/10.1126/sciadv.aau4586> PMID: 30662946

21. Cherepnalkoski D, Karpf A, Mozetič I, Grčar M. Cohesion and coalition formation in the European Parliament: Roll-call votes and Twitter activities. *PLOS ONE*. 2016; 11(11). <https://doi.org/10.1371/journal.pone.0166586> PMID: 27835683
22. Barberá P. Birds of the same feather tweet together: Bayesian ideal point estimation using Twitter data. *Political Analysis*. 2015; 23(1):76–91. <https://doi.org/10.1093/pan/mpu011>
23. Jackson SJ, Foucault Welles B. Hijacking #MNYNYPD: Social Media Dissent and Networked Counterpublics. *Journal of Communication*. 2015; 65(6):932–952. <https://doi.org/10.1111/jcom.12185>
24. Jackson SJ, Welles BF. #Ferguson is everywhere: Initiators in emerging counterpublic networks. *Information, Communication & Society*. 2016; 19(3):397–418. <https://doi.org/10.1080/1369118X.2015.1106571>
25. Theocharis Y, Lowe W, Deth JWv, García-Albacete G. Using Twitter to mobilize protest action: Online mobilization patterns and action repertoires in the Occupy Wall Street, Indignados, and Aganaktismenoi movements. *Information, Communication & Society*. 2015; 18(2):202–220. <https://doi.org/10.1080/1369118X.2014.948035>
26. Bovet A, Makse HA. Influence of fake news in Twitter during the 2016 US presidential election. *Nature Communications*. 2019; 10(1):1–14. <https://doi.org/10.1038/s41467-018-07761-2>
27. Lee J, Lim Y. Gendered campaign tweets: The cases of Hillary Clinton and Donald Trump. *Public Relations Review*. 2016; 42(5):849–855. <https://doi.org/10.1016/j.pubrev.2016.07.004>
28. Darwish K, Magdy W, Zanoua T. Trump vs. Hillary: What went viral during the 2016 US presidential election. In: Ciampaglia GL, Mashhadi A, Yasseri T, editors. *Social Informatics*. Lecture Notes in Computer Science. Cham: Springer International Publishing; 2017. p. 143–161.
29. Quelch J, Teixeira T. *The Twitter Election*. HBS Working Knowledge. 2016;.
30. Clarke I, Grieve J. Stylistic variation on the Donald Trump Twitter account: A linguistic analysis of tweets posted between 2009 and 2018. *PLOS ONE*. 2019; 14(9). <https://doi.org/10.1371/journal.pone.0222062>
31. Ahmadian S, Azarshahi S, Paulhus DL. Explaining Donald Trump via communication style: Grandiosity, informality, and dynamism. *Personality and Individual Differences*. 2017; 107:49–53. <https://doi.org/10.1016/j.paid.2016.11.018>
32. Bollen J, Mao H, Zeng X. Twitter mood predicts the stock market. *Journal of Computational Science*. 2011; 2(1):1–8. <https://doi.org/10.1016/j.jocs.2010.12.007>
33. Bloomberg. Bloomberg launches a Twitter feed optimized for trading | Bloomberg L.P. 2018;.
34. Hamed AA, Wu X, Erickson R, Fandy T. Twitter K-H networks in action: Advancing biomedical literature for drug search. *Journal of Biomedical Informatics*. 2015; 56:157–168. <https://doi.org/10.1016/j.jbi.2015.05.015>
35. Wetsman N. How Twitter is changing medical research. *Nature Medicine*. 2019;. <https://doi.org/10.1038/s41591-019-0697-7>
36. Cody EM, Reagan AJ, Dodds PS, Danforth CM. Public opinion polling with Twitter. *arXiv:160802024 [physics]*. 2016;.
37. Bovet A, Morone F, Makse HA. Validation of Twitter opinion trends with national polling aggregates: Hillary Clinton vs Donald Trump. *Scientific Reports*. 2018; 8(1):1–16. <https://doi.org/10.1038/s41598-018-26951-y>
38. Yaqub U, Chun SA, Atluri V, Vaidya J. Analysis of political discourse on Twitter in the context of the 2016 US presidential elections. *Government Information Quarterly*. 2017; 34(4):613–626. <https://doi.org/10.1016/j.giq.2017.11.001>
39. DiGrazia J, McKelvey K, Bollen J, Rojas F. More tweets, more votes: Social media as a quantitative indicator of political behavior. *PLOS ONE*. 2013; 8(11). <https://doi.org/10.1371/journal.pone.0079449> PMID: 24312181
40. Pasek J, McClain CA, Newport F, Marken S. Who's tweeting about the President? What big survey data can tell us about digital traces? *Social Science Computer Review*. 2019; p. 0894439318822007. <https://doi.org/10.1177/0894439318822007>
41. O'Connor B, Balasubramanian R, Routledge BR, Smith NA. From tweets to polls: Linking text sentiment to public opinion time series. In: *Fourth International AAAI Conference on Weblogs and Social Media*; 2010. Available from: <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM10/paper/view/1536>.
42. Wang Y, Luo J, Niemi R, Li Y, Hu T. Catching fire via “Likes”: Inferring topic preferences of Trump followers on Twitter. In: *Tenth International AAAI Conference on Web and Social Media*; 2016. Available from: <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM16/paper/view/13054>.
43. Hui C, Tyshchuk Y, Wallace WA, Magdon-Ismail M, Goldberg M. Information cascades in social media in response to a crisis: A preliminary model and a case study. In: *Proceedings of the 21st International*

- Conference on World Wide Web. WWW'12 Companion. Lyon, France: Association for Computing Machinery; 2012. p. 653–656. Available from: <https://doi.org/10.1145/2187980.2188173>.
44. Baños RA, Borge-Holthoefer J, Moreno Y. The role of hidden influentials in the diffusion of online information cascades. *EPJ Data Science*. 2013; 2(1):6. <https://doi.org/10.1140/epjds18>
 45. Bhattacharya D, Ram S. Sharing news articles using 140 characters: A diffusion analysis on Twitter. In: 2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining; 2012. p. 966–971.
 46. Goel A, Munagala K, Sharma A, Zhang H. A note on modeling retweet cascades on Twitter. In: Gleich DF, Komjáthy J, Litvak N, editors. *Algorithms and Models for the Web Graph*. Lecture Notes in Computer Science. Cham: Springer International Publishing; 2015. p. 119–131.
 47. Crane R, Sornette D. Robust dynamic classes revealed by measuring the response function of a social system. *Proceedings of the National Academy of Sciences*. 2008; 105(41):15649–15653. <https://doi.org/10.1073/pnas.0803685105>
 48. Candia C, Jara-Figueroa C, Rodriguez-Sickert C, Barabási AL, Hidalgo CA. The universal decay of collective memory and attention. *Nature Human Behaviour*. 2019; 3(1):82–91. <https://doi.org/10.1038/s41562-018-0474-5>
 49. Prieto Curiel R, Cabrera Arnau C, Torres Pinedo M, González Ramírez H, Bishop SR. Temporal and spatial analysis of the media spotlight. *Computers, Environment and Urban Systems*. 2019; 75:254–263. <https://doi.org/10.1016/j.compenvurbsys.2019.02.004>
 50. Amati G, Angelini S, Capri F, Gambosi G, Rossi G, Vocca P. Modelling the temporal evolution of the retweet graph. *IADIS International Journal on Computer Science & Information Systems*. 2016; 11(2).
 51. ten Thij M, Ouboter T, Worm D, Litvak N, van den Berg H, Bhulai S. Modelling of trends in Twitter using retweet graph dynamics. In: Bonato A, Graham FC, Prałat P, editors. *Algorithms and Models for the Web Graph*. Lecture Notes in Computer Science. Cham: Springer International Publishing; 2014. p. 132–147.
 52. Pramanik S, Wang Q, Danisch M, Guillaume JL, Mitra B. Modeling cascade formation in Twitter amidst mentions and retweets. *Social Network Analysis and Mining*. 2017; 7(1):41. <https://doi.org/10.1007/s13278-017-0462-1>
 53. Kempe D, Kleinberg J, Tardos E. Maximizing the spread of influence through a social network. In: *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining. KDD'03*. Washington, D.C.: Association for Computing Machinery; 2003. p. 137–146. Available from: <https://doi.org/10.1145/956750.956769>.
 54. Jin H, Toyoda M, Yoshinaga N. Can cross-lingual information cascades be predicted on Twitter? In: Ciampaglia GL, Mashhadi A, Yasseri T, editors. *Social Informatics*. Lecture Notes in Computer Science. Cham: Springer International Publishing; 2017. p. 457–472.
 55. Lee K, Mahmud J, Chen J, Zhou M, Nichols J. Who Will Retweet This? Detecting Strangers from Twitter to Retweet Information. *ACM Transactions on Intelligent Systems and Technology (TIST)*. 2015; 6(3):31:1–31:25. <https://doi.org/10.1145/2700466>
 56. Wu S, Hofman JM, Mason WA, Watts DJ. Who says what to whom on Twitter. In: *Proceedings of the 20th international conference on World wide web. WWW'11*. Hyderabad, India: Association for Computing Machinery; 2011. p. 705–714. Available from: <https://doi.org/10.1145/1963405.1963504>.
 57. Barberá P, Rivero G. Understanding the political representativeness of Twitter users. *Social Science Computer Review*. 2014;. <https://doi.org/10.1177/0894439314558836>
 58. Mellon J, Prosser C. Twitter and Facebook are not representative of the general population: Political attitudes and demographics of British social media users. *Research & Politics*. 2017; 4(3):2053168017720008. <https://doi.org/10.1177/2053168017720008>
 59. Wojcik S, Hughes A. How Twitter users compare to the general public; 2019. Available from: <https://www.pewresearch.org/internet/2019/04/24/sizing-up-twitter-users/>.
 60. Murthy D, Gross A, Pensavalle A. Urban social media demographics: An exploration of Twitter use in major American cities. *Journal of Computer-Mediated Communication*. 2016; 21(1):33–49. <https://doi.org/10.1111/jcc4.12144>
 61. Grinberg N, Joseph K, Friedland L, Swire-Thompson B, Lazer D. Fake news on Twitter during the 2016 U.S. presidential election. *Science*. 2019; 363(6425):374–378. <https://doi.org/10.1126/science.aau2706>
 62. Wojcik S, Hughes A, Remy E. 19% of U.S. adults on Twitter follow Trump;. Available from: <https://www.pewresearch.org/fact-tank/2019/07/15/about-one-in-five-adult-twitter-users-in-the-u-s-follow-trump/>.
 63. Shear MD, Haberman M, Confessore N, Yourish K, Buchanan L, Collins K. How Trump Reshaped the Presidency in Over 11,000 Tweets. *The New York Times*. 2019;.

64. Dodds PS, Minot JR, Arnold MV, Alshaabi T, Adams JL, Dewhurst DR, et al. Fame and Ultrafame: Measuring and comparing daily levels of 'being talked about' for United States' presidents, their rivals, God, countries, and K-pop. arXiv:191000149 [physics]. 2019;.
65. Jackson SJ, Bailey M, Welles BF, Lauren G. #HashtagActivism: Networks of Race and Gender Justice. Cambridge: The MIT Press; 2020.
66. Twitter. Decahose stream; 2019. <https://developer.twitter.com/en/docs/tweets/sample-realtime/overview/decahose>.
67. Twitter. Twitter IDs (snowflake); 2019. <https://developer.twitter.com/en/docs/basics/twitter-ids>.
68. Dodds PS, Minot JR, Arnold MV, Alshaabi T, Adams JL, Dewhurst DR, et al. Allotaxonomy and rank-turbulence divergence: A universal instrument for comparing complex systems. arXiv:200209770 [physics]. 2020;.
69. Twitter. What are Promoted Tweets?; 2019. Available from: <https://business.twitter.com/en/help/overview/what-are-promoted-tweets.html>.
70. Dodds PS, Harris KD, Kloumann IM, Bliss CA, Danforth CM. Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter. PloS one. 2011; 6(12). <https://doi.org/10.1371/journal.pone.0026752>