#WeCouldHaveSeenThemComing
Trump, Sanders, and the Undervalued Metrics of Twitter Followers and Engagement

A PEORIA Project Report
May 25, 2016

By

Michael D. Cohen, Ph.D.  michaeldcohen@gwu.edu
Michael B. Cornfield, Ph.D.  corn@gwu.edu
Lara M. Brown, Ph.D.  larambrown@gwu.edu
Introduction: Surprise, Surprise

The Indiana primaries held on May 3 marked the *de facto* end of the 2016 races for the Democratic and Republican nominations for president of the United States. Former Secretary of State Hillary Clinton and businessman Donald Trump will face off from here. Clinton has secured enough delegates at the ballot box and among super-delegates that Senator Bernie Sanders can no longer defeat her before the convention. Trump’s remaining rivals on May 3 dropped out after his victory.

While most political observers anticipated that Clinton would get the nod of her party, the rise of Trump and Sanders have been surprising. Trump flirted with running for president in the past but never before put his name on a ballot for public office. Sanders, as obscure as Trump was famous in 2015, began his race for the Democratic nomination as a self-described “Democratic Socialist.” He had not won his Senate seat on the Democratic ticket, although he has caucused with the party in the upper chamber.

Many pundits and prognosticators have misjudged the seriousness of the Trump and Sanders bids. With good reasons: these men were party outsiders attempting entry at the top. In summer 2015, GOP eminence Karl Rove called Trump a “sideshow” and noted Clinton’s formidable advantages in campaign money and organization. The sideshow quickly became the main show; as we documented in an earlier PEORIA report [hyperlink], ever since his June 16, 2015 candidacy announcement, coverage of Trump has consistently dominated network news coverage, boosted cable news ratings, and ruled online political news. Meanwhile, by April 21, 2016 Sanders’ campaign committee had actually outraised Clinton by over $2 million despite refusing PAC money. Nate Silver, the statistical virtuoso who wrote *The Signal and the Noise*, admitted on May 18 to falling into a punditry trap, particularly regarding Trump.

Early on Trump was widely viewed as an attention-seeker running mainly to promote his name-branded venues and products. He employed no pollsters and built no “ground game” or field organization to speak of. Yet at the start of 2016 Trump held a 20% lead in national polling, the same as Al Gore did in 2000 and behind only George W. Bush that same year in the last four presidential cycles. A lead of this magnitude was a strong indicator of likely nomination; the only candidate who lost with a polling lead of comparable size to qualify under the “December rule” was Hillary Clinton in 2008. (She held a similar 19% lead in January 2016.) In acknowledging this pattern in mid-January 2016, Larry J. Sabato, Kyle Kondik, and Geoffrey Skelley nevertheless surmised that
Republican party leaders were likely to coalesce around one of the remaining viable establishment candidates: Jeb Bush, Chris Christie, John Kasich, or Marco Rubio.

As for Sanders, his large crowds, small donor millions, and steady climb in the polls garnered him lots of faint and qualified praise from the commentariat. News stories and opinion pieces about how he was unelectable have accompanied him throughout his candidacy. His issue positions are radical (“revolutionary” in his rhetoric) and well to the left of median voters even among Democrats. Clinton has enjoyed overwhelming leads in endorsements from party elites and commitments from super-delegates. Debates, a potential equalizer or reverser of electoral fortune, were scheduled for off hours on off days to minimize the chance of a decline in the front-runner’s lead. Yet it has dwindled markedly in 2016.

And so the central two-faceted puzzle of the effectively concluded primary season is: how did Trump win and Sanders get so close? Based on our PEORIA Project research of social media analytics with our partner Crimson Hexagon, we find it likely that Twitter is part of the answer. At the least, Twitter provides important and heretofore neglected (at least in public) intelligence on the amount of attention candidates are getting from an important segment of the public. Twitter also supplies candidates with direct means to engage and build political support.

Trump and Sanders have made substantial use of Twitter. In an April 2016 interview with CNN’s Anderson Cooper, Trump confirmed that he does most of his own tweeting or dictates tweets to staffers. His constant and unmistakably personal presence on a platform that favors authentic, short statements with an edge could have been a campaign asset for the first-time candidate. A similar advantage may have obtained for Sanders, whose clear and individuated tweets by Hector Sigala, a 27-year old staffer, has contrasted with the more conventional campaign voice of Clinton campaign tweets.

To confirm the intelligence and engagement utilities of Twitter, it’s necessary to examine data about usages. In this report we present strongly suggestive patterns in Twitter follower growth, total engagement, and retweets for the leading presidential candidates from the day they announced their candidacies in 2015 until May 3, 2016. We have yet to solve the puzzle of the presidential cycle, but we believe we have identified a key.

**Concepts and Method**

The enterprise social media analytics company Crimson Hexagon receives and archives a pipeline of all tweets on Twitter. Its tools track the number of followers each candidate has day-by-day, allowing us to analyze follower growth over the course of the campaign.
“Total engagement,” as defined by Crimson Hexagon, adds together the retweets (repeats of a tweet by another account), replies (tweets addressed to a particular account via the @ sign), and mentions of a tweeter. These items vary in their helpfulness to a campaign. Account replies may be hostile. Mentions may be hostile or digressive. Retweets, however, are informational campaign contributions. They increase the reach of a campaign message—and arguably its credibility as well—through the validating vehicle of the persons who deem it worthy of passing along to their own social networks.

A cautionary note: our analysis of the data does not correct for the presence of bots. Bots are automated followers generating tweets designed to inflate frequency counts and the popularity associated with those numbers. There is no question that bots have been in use in the 2016 presidential campaign discourse. As Patrick Ruffini, co-founder of the digital agency Echelon Insights, noted in early April, there are hundreds of Twitter accounts promoting Trump that are likely fake.

However, we do not think this seriously warps our findings. We present large-scale and long-term developments in comparative context, and available evidence does not indicate that any one campaign has promulgated bots to a decisive advantage over others. Philip Howard of the University of Washington noted on May 16 that “According to the site TwitterAudit, one in four of Trump’s followers is fake, and similar ratios run through the accounts of the other presidential hopefuls.” At the same time, the artificial boosting of social media metrics cannot be dismissed outright and we would not be surprised to find evidence of coordination with campaigns, even if unofficially.
Findings

1. A Twitter following is a campaign asset from day one.

<table>
<thead>
<tr>
<th>NAME</th>
<th>ANNOUNCEMENT DATE</th>
<th>TWITTER FOLLOWERS THAT DAY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ted Cruz</td>
<td>3/23/15</td>
<td>371,742</td>
</tr>
<tr>
<td>Hillary Clinton</td>
<td>4/12/15</td>
<td>3,208,701</td>
</tr>
<tr>
<td>Marco Rubio</td>
<td>4/13/15</td>
<td>706,910</td>
</tr>
<tr>
<td>Ben Carson</td>
<td>5/4/15</td>
<td>330,441</td>
</tr>
<tr>
<td>Bernie Sanders</td>
<td>5/26/15</td>
<td>41,732</td>
</tr>
<tr>
<td>Jeb Bush</td>
<td>6/15/15</td>
<td>198,340</td>
</tr>
<tr>
<td>Donald Trump</td>
<td>6/16/15</td>
<td>2,997,242</td>
</tr>
<tr>
<td>John Kasich</td>
<td>7/21/15</td>
<td>75,445</td>
</tr>
</tbody>
</table>

As the table above documents, Donald Trump and Hillary Clinton launched their presidential candidacies with big leads over their rivals in the number of Twitter followers. Their celebrity status and embrace of the medium enabled these leads; both were famous long before Twitter was unveiled in public in March of 2006, and both found tweeting useful in previous capacities.

So what? Isn’t agreeing to follow someone at most a casual act of slacktivism?

We know from Pew Research that 23% of the 85% of US adults online in 2015 used Twitter, up from 16% in 2012 although somewhat plateaued since 2014. Nearly half of Twitter users, 49%, did so weekly or daily. Pew also reported in 2014 that “the 16% of registered voters who follow political figures on social media participate in various traditional campaign activities at high rates.” Our own Battleground Poll in April 2016 found that 52% of registered voters learned about the presidential race through social media.

Twitter lacks the individual reach of texting, the bulk reach of email, and the networked reach of Facebook. But the Twitterverse in which followers, engagers, and retweeters communicate consists of opinion leaders and activists. We know, too, that Twitter content is visible in other media; an American Press Institute survey found that half of non-Twitter users have seen tweets on television, in online news, in messages from social networks, and even the good old newspaper.
To follow, first, is to show interest in a particular campaign. Second, following is interest made demonstrable to campaigns and the public: in the aggregate as in the above table, and to an extent in personally identifiable and group-sorted forms as well. Third, following opens the follower up to direct communications from the followed for as long and often as the campaign chooses to send tweets, unless and until the follower severs the connection. This is as though a person had a television always on and always available with ads surfacing during the programming already being watched. However, that does not mean always attended to; tweets disappear from immediate view rather quickly. Fourth, following is a step toward further engagement with a campaign and on behalf of a campaign, as we will discuss ahead.

In sum, an advantage in the number of followers can translate into stronger name identification, attention-hogging potential (contingent on messaging talent and appeal, a strong qualifier), and potential support (depending on reader impulses and campaign solicitation prowess, also strong qualifiers). Trump has converted on this potential, although we have a ways to go in our research to explain precisely how and to what extent he has done it. Sanders has, too, although in his case he seems to have inherited followings from existing progressive social networks such as those connected with MoveOn.org, Democracy for America, ActBlue, and Senator Elizabeth Warren.

Our main point here is simply that campaigns should know where they are relative to their competition in Twitter followers, on day one and everyday thereafter. Indeed, given the free media boost accorded candidates at every level on their announcement day, it pays to assess beforehand where a campaign stands in followers relative to existing and likely competition.

2. Growth rates in Twitter followings provide actionable evidence of campaign strength and weakness.

Follower growth rates inform campaigners and observers about how well the political potential inherent to having a Twitter following is being converted into political capital. Crucially, the metric does so in a way that lends itself to helping a campaign take steps to improve its growth rate. The instructiveness applies to campaigns with small as well as large numbers of followers. One can’t just go to work and boost a poll number or television rating. But as with fundraising, one can do numerous things on a daily basis to increase a follower growth rate. Twitter metrics are convenient, inexpensive, continual, and thus more susceptible to alteration than most communication channel metrics.

The rate seems to matter, as can be seen in the chart of normalized growth rates for the period under review. In the top tier, Trump overtook Clinton in mid-October, 2015. This an early signal that his campaign was building while hers was sustaining. Trump’s
Twitter following grew by 178% from announcement through Indiana, almost twice Clinton’s rate of 89%. These data points, if better known, might have altered elite views of Trump’s prospects at a time when more could have been done to stop him -- or, alternately, when it might have been more politically rewarding to back him.

Sanders came out of nowhere, growing his Twitter followers 1,526% from his announcement through the Indiana primary, close to three times the rate Ben Carson’s grew (552%), about five times the rate John Kasich grew (343%) and close to six times the percent increase of Ted Cruz’s following (260%) over the same period. All four grew much more than the rest of the packs. Sanders sustained his growth rate as the other three did not. However, it was not a rate significantly greater than Clinton’s, meaning that, with followers as with delegates, he failed to close the gap. Yet he did in polls, which suggests that the mainsprings for his political success lie beyond his Twitter performance as a messenger and more with respect to engagement.
Digital strategists for these campaigns could take follower growth data, plot it against key dates, calculate follower growth rates as goals, and strive to attain them by:

- Reaching out to endorsers with known followings non-congruent with the campaign’s existing following.
- Monitoring trending topics and crafting tweets that match them.
- Improving the campaign’s twitter voice (a move that probably involves reducing the number of people authorized to speak in it to as close to 1 as possible).
We do not know much yet about the Twitter operations of presidential campaigns besides Trump’s. It will be fascinating to learn which campaigns monitored these metrics. It will also be important to learn more about the use of bots to artificially stimulate follower numbers.

3. **Engagement Metrics, Especially Re-Tweets, Provide a Second Dimension of Campaign Intelligence Apart from Follower Metrics**

Crimson Hexagon defines “total engagement” as the sum of retweets, mentions, and replies on the platform. As you can see, Trump leads all candidates in total engagement (the blue line at the top of the first graph below) but Sanders had a 2:1 edge over Clinton (seen clearer in the second graph), which may partly explain how he closed the polling gap despite fewer Twitter followers.

Which component of total engagement was the most important? Replies are responses by Twitter users to something a candidate says. A tweet that is an @reply could reinforce the candidate’s message or clash with it. Even a supportive message addressed to the candidate could change the context of the campaign message. Sentiment analysis aspires to distinguish among these possibilities, but for now it is outside the scope of our inquiry. A @mention is simply the candidate’s Twitter account being called out by another Twitter user. Again, this can vary in what is said and how positive, negative, or neutral the new tweet is. Still, these users are definitely engaged
with a candidate’s messages and, therefore, worthy of study. The charts below show how Twitter users are engaged with each candidate throughout the period from their announcements to the Indiana primaries.

We believe the *key* Twitter metric to watch is a *component* of total engagement:
retweets, which is the cleanest possible echo of a campaign tweet. By repeating the campaign’s line word for word the effect of the tweet is to amplify the message and, in most cases, express enthusiasm for it. Yes, many journalists write in their bios that retweets ≠ endorsements; all the same, the effect of retweeting is to extend the message to a new audience. It’s free media.

Bernie Sanders consistently outperformed Hillary Clinton on retweets. Although Sanders entered the race five weeks after Clinton, he ended up with more retweets through the Indiana primary (6,797,859 vs. 4,020,452). As most campaigns come down to turnout, it’s clear that Sanders partisans were more engaged than Clinton’s. There’s no direct correlation established here. But at some point, with future data from other cases, political analysts will be able to attach a multiplier to retweets to account for their impact. In other words, our results suggest it is worth testing the hypothesis that enthusiastic Sanders retweeters were more likely to vote than Clinton’s supporters. We cannot do that absent voter file data.
4. The potential power of Twitter lives on past the end of campaigns.

At the end of the period we studied, May 3, we had two presumptive nominees in Clinton and Trump, and a remaining candidate in Sanders. As Cruz and Kasich dropped out on May 3rd, and 4th, respectively, we do not want to get too far ahead of the data. This leaves us with Bush, Carson, and Rubio, who all suspended their candidacies with enough time before May 3 to study the trends in Twitter followers and, surprisingly, continued growth.

As you can see from the chart below, all three candidates’ Twitter accounts continued to grow in the absence of active candidacies. The notion that Twitter followers stay – and even grow – may leave opportunities for future campaigns. Rubio’s account had consistent growth but did level off after his announcement to suspend but it is still growing. Carson’s continues to grow as his support of Trump and active involvement as a surrogate continues to build his following, which will likely overtake Rubio’s over the next month. Despite all of the money raised and spent as well as his relatively early exit, Bush’s account continues to grow steadily. But the bottom line is this: leaving the race does not result in fewer followers; it results in more. The irony of an exit boost is not that the departed made a mistake; it’s that the exit is a moment to help another day.
However, we should also note that followers disengage after campaign suspensions as shown in the graph below, leaving plenty of rebooting work if these candidates return. It is also likely, if we had the data, that we’d find the same pattern for Clinton in 2008.

In sum, returning candidates should be looking at Twitter followers as an asset to reengage. Their future opponents should see them coming. This is particularly important with respect to Sanders, the purported head of a sweeping and radical force. Dean institutionalized his following before social media really took off; what will the Sanders people (some the same) do? What will he and his twitter voice ask them and arrange for them to do?
5. **A hashtag is not equivalent to a following or engagement activities.**

Late opposition to Trump's nomination within the Republican Party coalesced in reaction to his victory in South Carolina on February 20. A day later, Marco Rubio joined the hashtag #NeverTrump, which started trending worldwide. By February 28th, CNN called it a movement, which ironically coincides with the moment that it declined being one on Twitter. As the chart below shows, the frequency of posts peaked on the 27th at 533,618, and was halved 282,906 the following day. #NeverTrump never again regained its early promise. There were two more mini-peaks in the movement, one on March 3 where there was a debate in Michigan and May 3, the last gasp of Trump resistance – the Indiana primary, sealing his victory.
So what happened? The central problem with the #NeverTrump effort was that it did not have a viable candidate. In the federal system, veto points and powers distributed across institutions make “no” strategies effective. But in electoral politics you can’t beat someone with no one. #NeverTrump was not equivalent to #Go[fill in the blank].

The GOP opposition to Trump could not agree to follow or even talk up an alternative. Ted Cruz lost South Carolina, undermining his southern strategy. Jeb Bush dropped out after losing the Palmetto State despite an all-in effort. By that time Rubio had already become an enduring robotic caricature of himself from his faceoff with Chris Christie, who quickly left the race to awkwardly endorse Trump. A relatively moderate governor and long-time elected official, John Kasich never fired up the base of the party who did not want an establishment candidate. We can see the implications of this situation in the relative flatness of these candidates’ Twitter metrics.

Furthermore, the #NeverTrump resistance was never a coherent Republican base of support. As this BuzzFeed story demonstrates, the hashtag was used by a diverse group of Twitter users, not just Republicans. It became not a campaign for defeating Trump but a way to uniformly express outrage at his candidacy from his comments about Mexico, those with special needs, his questionable business deals, or his personality. This speaks to an important difference between hashtags, on the one hand, and followings and engagement metrics, on the other: Hashtags express ideas and sentiments that do not necessarily correspond to a campaign effort.

We find that hashtag-tracking is important but in a different way. It clarifies key moments in a campaign. The hashtag extended beyond the Twitter platform. Undoubtedly the emergency of the #NeverTrump hashtag was a sign of angst against
the controversial outsider, which included Republicans. But the reason why it went viral was because it extended beyond the GOP. The fact that it was resurrected on May 3, at the de facto end of the nomination, shows that this was a vessel for strong feelings, not a sustainable movement.

Conclusion

Twitter metrics have been studied in business and academic for quite some time. In this report, we argue that Twitter data metrics can enable political analysts on the job to sketch out campaign stories that diverge from the conventional wisdom of poll- and pundit-driven daily narratives. This can occur because Twitter (along with Facebook, Instagram, Snapchat, and other social media) provides a new type of evidence: low-level expressions of voter preference prompted by a variety of cues other than the questions of pollsters. This is not a knock on polling as practiced responsibly; polling remains an essential research tool. Indeed, we think polling will be soon be part of a standardized multi-method approach to interpreting campaign events and collecting feedback.

Social media metrics are valuable correctives. By monitoring them since March 2015, we did see Sanders coming and Trump lasting through them. And we have here only scratched the surface of Twitter analytics as adaptable to political intelligence and campaign activity.

We will be looking in future weeks at tweet content, word cluster patterns, and polling data in a multi-method probe of the messages the candidates purveyed and generated. We are keen to learn, in particular, how topics come to dominate and frame election discourse; Twitter organizes topical searches and discussions through the hashtag. We are equally intent on looking for telltale signs of movements versus organized campaigns as evident on Twitter: when a comer is spotted, can the analyst discern how much is driven by a big campaign team and how much by a flock of tweeters with homes in the grassroots. We will explore connections between Trump and the Tea Party, and Sanders and MoveOn.org, Democracy For America, and Senator Elizabeth Warren. A third focus of our upcoming inquiry will be the dynamics of social and mainstream media. And a fourth will explore differences in metrics broken out by party.

There is more than slacktivism on display in social media. Even when that is mostly what accumulates, the noise obscures signals being sent by competitors. Attention is a resource as vital to campaigns as money, message, and candidate quality. It too must be managed efficiently, and to do that smartly it must be studied.