zettacap

Analyzing the 2016 US Presidential Election
Social Media Influence Index

August 2016

Summary

Social Media interactions are very similar to traditional polls and surveys in many ways, except that you can get much larger sample sizes, in real time, and with very limited data manipulation. With this in mind, we explore three datasets – our Social Media Influence Index (SMI), candidates' tweet 'likes', and Wikipedia page views. These three datasets show what people are doing on-line, how they are interacting on social media, and what topics interest them. For the modern world, these should provide excellent insights into what people really think in an unfiltered and unbiased way about the candidates and election.

The SMI continues to show Trump dominance. After a brief decline at the beginning of August, the relative Trump versus Clinton SMI has bounced back to its July peak implying that polls will soon follow. Given historical trends, we expect polls to pull even and show Trump with a slight lead very soon. The SMI made very accurate election calls even before primary voting started and appears to be the most accurate election forecaster of this cycle. It currently continues to call for a Trump victory.

Tweet 'likes' favor Trump as well with his average tweet over the past month receiving almost four times the number of 'likes' as those of Clinton. This is an extremely strong indicator for interest in the candidate and an indirect measure of enthusiasm of potential voters. Individual tweet analysis is also important in helping to identify key topics that excite 'the base' of each candidate. We analyze these Top Topics for each candidate and provide advice for where each side might want to focus. In the end, it seems like it will be a battle of which candidate is more unfit for the presidency.

Wikipedia page views again shows that on-line activity favors Trump. There are consistently more people visiting Trump's page, his campaign's page, his slogan's page, and the Republican page. Clearly, Trump holds the advantage on Wikipedia as well.

Trump's Continued SMI Dominance

From the start of the electoral cycle in 2015, Trump has dominated the social media landscape. As reported in previous updates, our Social Media Influence index predicted Trump's rise within the Republican race well before pundits or polls took him seriously. Over the last month, his lead has continued over Clinton, as can be seen in the following comparative chart. This is especially impressive considering the fact that in polls, Clinton's lead has skyrocketed over the same period.

Chart 1: SMI for Trump minus Clinton
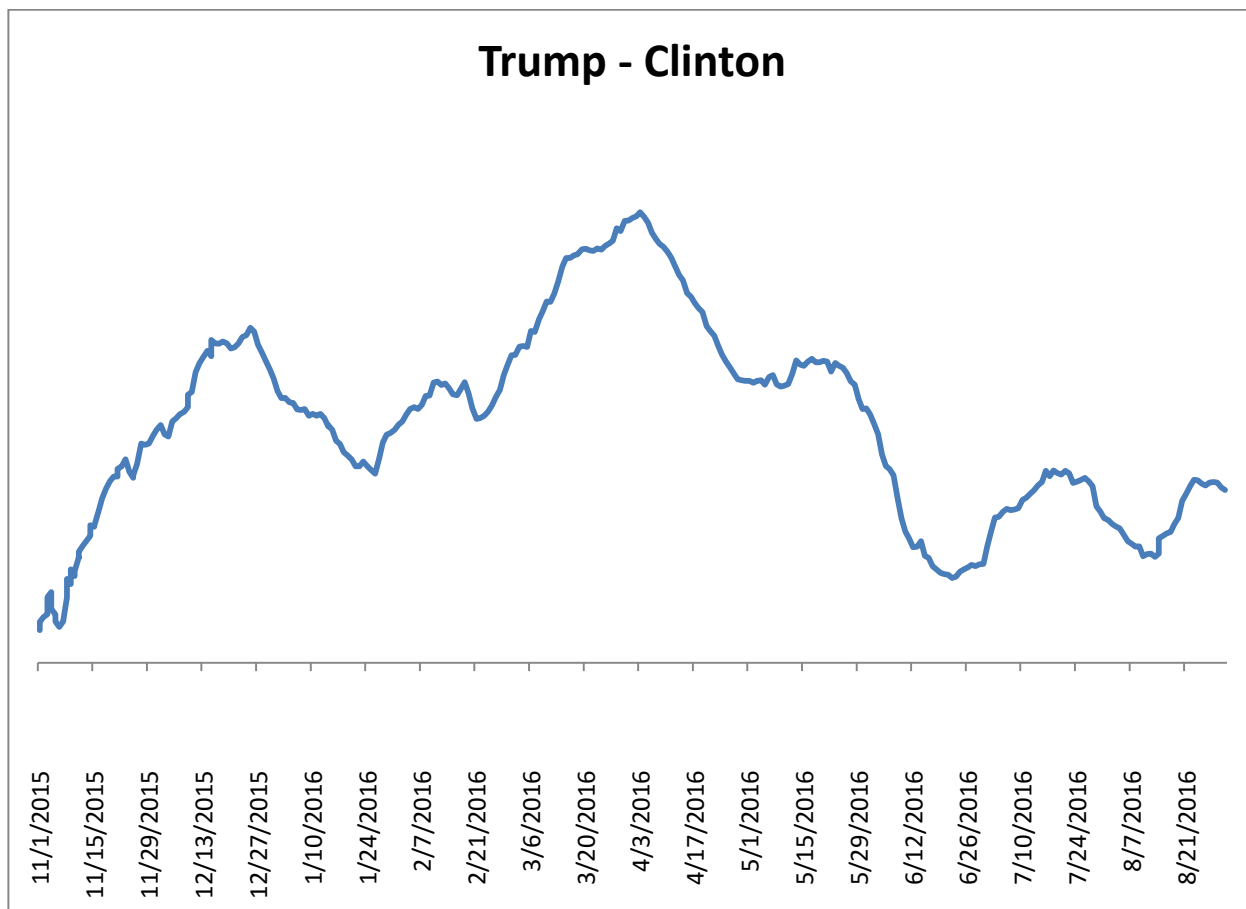
Trump - Clinton

It should be noted that the baseline, or the horizontal line, is a zero net SMI rating or when the candidates' ratings are equal. Above this line, Trump's SMI is higher than Clinton's, which forecasts a Trump victory. Below this line, Clinton's is better. As we can see, Trump has been consistently on top in terms of SMI.

One of the most important elements of this chart is the continued dominance of Trump in the face of declining polls. Right after the DNC at the end of July, Clinton's poll numbers improved dramatically and many in the media all but declared the race over. In terms of SMI, however, Trump's decline was rather minor and quick. His SMI reversed upwards by mid-August and returned to mid-July highs.

This is counter intuitive. Polls showed that Trump's candidacy had failed with the media reporting that it would be almost impossible for a candidate to come back after being behind by so much so close to the election. In reality, however, online activity showed the opposite – namely that Trump continued to have more influence than Clinton. Though not a topic for this report, it clearly seems like either the polls are correct or social media and on-line activity are correct. If social media turns out to be correct, which we firmly believe it will, it will rewrite much of what we know about data and election analysis.

We have stated previously that we expect the relative SMI of Clinton versus Trump to start to become more accurate after the national conventions. In other words, from this point forward the SMI should be increasingly accurate for forecasting true poll levels and then the election. As it stands currently, it looks like the SMI is leading poll results. Given that the SMI has spiked higher, in Trump's favor, we expect polls to follow suit. Keeping all things constant, polls should show similar results to when the SMI peaked in July. In other words we expect the polls to quickly pull even and then show Trump with a minor lead. Afterwards, we will have to see how the SMI develops, but certainly it is currently signaling the end of Clinton's post DNC poll dominance.

As previously stated, we believe that social media activity and influence is extremely powerful for forecasting election results. Our analysis is the only quantitative analysis that forecast Trump dominance in the Republican primaries and Clinton eventual winning but only after a very difficult race with Sanders in the Democratic primaries, before voting started. In fact our social media influence model predicted such results from August 2015. In an election cycle that has seen the best and more experienced pundits and analysts publically apologize for making terrible calls our track record is phenomenal.

Our disclaimer is that our data and analysis methods are new with short track records. Social media and on-line activity analysis has simply not been around that long to establish such a long track record. So, take the SMI and all other social media and on-line activity analysis with a grain of salt. Regardless, to-date, there is not one model or pundit who has come close. And, the SMI continues to make the out-of-consensus call of a Trump victory by a not very close margin.

4

Social Media Analysis vs Polls
There has been a sharp increase in the focus on polls. Because this election cycle has been more volatile perhaps polls might provide some necessary direction or grounding for observers. It is not entirely clear as to the reason, but certainly when a new poll can make a major headline, there is a lot of interest there.

It seems rather odd then that there has not been more interest shown in social media analysis of the race as in many ways they are the new polls, albeit indirect.

Social media, among other things, shows interactions on-line. It depicts what is on peoples' minds, what they are interested in, and where they are spending their time. If someone for instance is searching for topics related to Hillary Clinton, interacting with the Democrat's social media pages, and 'liking' tweets made by Clinton, we can be fairly sure her interest and enthusiasm level for Clinton is higher than someone who spends considerably less time and energy doing so. We can say the same thing for Trump or any candidate.

Now, if we multiply this analysis by the population of Internet, we can see what is really going on in a race.

Additionally, social media often makes up for shortcomings of polls. Polls are great, there is no question that they are extremely useful. However, like everything, they have weakness and social media analysis tends to make up

for some of these weaknesses. First, polls rely on relatively small sample sizes which must be created in a particular way or it could mess up the results. Social media analysis in contrast is capable of using the entire dataset available, so there is really no need to worry about proper sampling. Second, polls are not real time as there is always a lag. Social media is immediate. Third, polls are open to various forms of bias – polling company bias could inadvertently massage data during the analysis process for a variety of reasons and/or respondents can change their answer as they might not want their true feelings to be known. Social media analysis simply portrays how people are interacting in a free and open environment.

In a normal case, polls (or surveys) and social media analysis should support one another. Or, each should confirm the conclusion of the other. One of the striking elements of the 2016 race is that polls and social media analysis have diverged when analyzing the nominees. Prior to and during the primaries, social media analysis did a great job in forecasting results even better than and before polls. However, polls tended to be more or less in-line with social media analysis for the larger races.

Now, in the general election, there is a significant divergence. Clinton tends to consistently lead in polls whereas Trump tends to consistently lead in social media. And, in neither case does it appear to be a close race.

We firmly believe in social media and on-line activity as leading indicators of people's intentions. We further believe that as this data is unfiltered and not significantly open to bias, that it could be more valuable than polls in emotional elections where increasing bias could come into effect.

Furthermore, our social media and on-line activity analysis is the only standard that we are aware of that actually correctly called the nominations before voting started. We quantitatively forecast that Trump would take the Republican nomination and that Clinton would win the Democratic nomination, but only after a very tight race with Sanders. Such performance was far superior to any polls or pundits prior to the beginning of voting.

5

We do not disregard polls. We simply have seen our analysis work extremely well in an election cycle that has thrown off many previously strong models and analysts.   We expect the results of polls and our analysis to tighten as election-day comes closer. Regardless, know that our analysis is refuted by polls, pundits, and betting markets who are clearly picking Clinton as the 2016 victor.

Tweet Popularity as an Indirect Poll
Twitter has emerged as the go-to platform for communication during the 2016 race. Each of the main candidates is very active on twitter and has millions of followers. Their streams are lively and cover almost every topic related to the race ranging from more serious policy type statements to questionable political accusations. In other words, twitter is a great unfiltered environment for researching what people really think.

Before getting into the analysis let's cut to the conclusion – Trump leads by a significant degree. Keeping all other things constant, the greater interest that people show in Trump's tweets implies a higher level of enthusiasm for his message and candidacy. Data from previous elections show that the candidate with higher levels of enthusiasm almost always wins. Judging from this data, and other social media based information, we continue to forecast a Trump victory in the general election.

Each tweet has four icons at the bottom of it. At the left, there is a curved arrow for 'reply', to its right there are two curved arrows for 're-tweet', to its right is a heart for 'like' and finally there is an icon for 'more'. In most

cases, the 'like' heart is the most used. It is a very fast and simple way for people to show their approval of a tweet.

We assume for political candidates, more 'likes' on average imply more interest and enthusiasm for that candidate and for specific tweets more 'likes' imply higher levels of agreement with that particular message. It is very simple and straightforward. In this way we can not only judge which candidates are getting better traction but which topics as well.

Trump's Tweet Popularity

Taking Trump's twitter stream from July 27 to August 27, we measure each tweet's level of popularity by the number of 'likes'. We then interpret the entire dataset to see what the distribution of the 'likes' says about the candidate. Finally, we look for the main topics where people have shown the most interest.

The following chart show Trump's tweet 'like' distribution.

Chart 2: Distribution of Tweet 'Likes' for Trump, July 27 to August 27, 2016
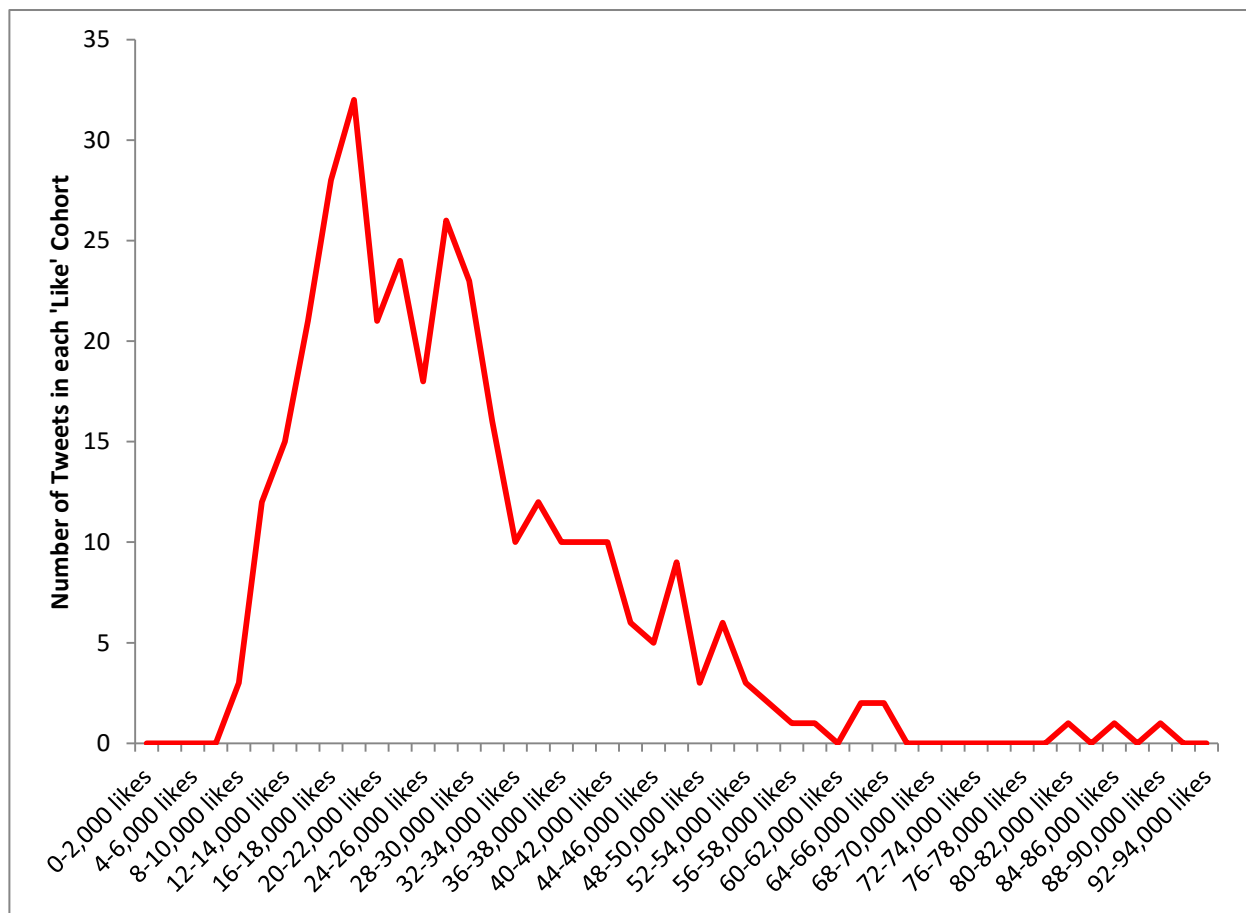
Source: Twitter

In general, the distribution of these 'likes' closely resembles what we would expect. Its skewness is 1.4, meaning that the distribution is further off to the right than normal. However, in analyzing social media 'likes' we want this type of skewness as it allows for easier Top Topic identification. In other words, those outliers on the right of the distribution contain the topics that are more 'liked' and therefore of more interest. Such a marginal right skew makes this far easier and accurate. Its kurtosis is 2.9 and extremely close to the accepted norm of 3. Kurtosis tells us how high the distribution is or how clumped up the data is around a very narrow point. This data appears very normal in terms of skewness and kurtosis for a social media dataset, making it adequate for analysis. Put another way, if there were significant manipulation, which some people fear (not in Trump's case but whenever looking at a dataset taken from publically available data), it would have likely been flagged here.

Next, let's drill down on the Top Topics of the tweets as measured by 'likes'. We roughly rank them by the peak 'likes' for each theme. Analyzing themes is difficult as some themes are fairly sparse (like "Trump is a normal person") whereas other repeat often and in different forms (like "Anti-Media"). In other words, these themes and their rankings include more qualitative analysis.

Table 1: Top Topics of Trump's Tweets from July 27 to August 27, 2016 as measured by 'Likes'

TOP TOPICS	'Likes' in thousands
Anti-Media	
Trump vs Media Bias	88
Khan argument, but Hillary voted for war, where's media?	59
Secret Service Did NOT speak with Trump	46
Baby, get out of room, not true	42
Crowds at Trump rallies, not covered	41
Obama not doing job	
Obama worst president	80
Louisiana, where is Obama?	42
Hillary Crooked	
Deleted Emails, maybe Russia has them	64
Iran Scandal	49
Emails	46
What Bernie Sanders thinks of Hillary Clinton	42
Debates during NFL games	41
Law and Order Candidate	
Shooting deaths of Police officers	64
Thanking Police	47
Trump is a normal person	
Trump with KFC, eating chicken	63
Hillary Health / Incompetent	
Where's Hillary? Sleeping	62
Isis, Hillary allowed	40
Hillary short-circuits	40

7

Borders	
Hillary wants to let refugees in	53
Hillary wants borderless country	45
Strong border and wall	32
Presidential Trump	
Trump is against Insiders	47
Work together to overcome	46
Visiting Louisiana	45
Minority Outreach	
African-Americans deserve better	34
African-American and Hispanics: "What do you have to lose?"	30

Source: Twitter

The most popular tweet from Trump was his statement that he is not just running against Hillary but the biased media. If this is accurate or not is for another day, but in terms of political analysis it is important to understand that the Republicans are very enthusiastic about the topic of media bias. Though not shown here Trump often tweets about specific media outlets and even media personalities. These tweets do not get as much traction. He would be well to lay off specific media attacks and increase the attacks on general media bias with examples. He often complains about the media not showing the size of the rallies – tweets showing their size and the media ignoring them would likely get a positive response. If Trump hones this message, this topic could gain even more traction. But again, random shots at media outlets do not seem to gain the necessary traction – it needs to be general media bias with examples.

Obama not doing his job or not doing it well is another popular topic in the Trump tweet stream. This was especially the case when it came to Louisiana and Obama not visiting. However, this is very punctual as a topic. Plus, as we will see in the Clinton stream, she benefits greatly from her association with Obama. So, a status quo of Trump making minor attacks on Obama will likely be counteracted by Clinton making positive remarks.

The' Crooked Hillary' series of tweets that purport to show that she is not honest are effective. It is most interesting to note that the top tweet was one where Trump used off-the-cuff humor (paraphrasing, maybe Russia can help access the lost emails). This was also the case in a few other highly 'liked' tweets (such as when people asked where Hillary was and Trump responded 'sleeping'). It seems that Trump gets a lot of traction through slightly questionable and certainly not traditionally political humor. A dangerous strategy, but if he could increase his use of humor to nail some of the more serious topics, his tweet popularity could increase significantly.

Trump being the law and order candidate also has generated traction on twitter. There has not been that much follow-through here in that the absolute number of tweets is not that high for this topic. However, the few that he has made has resulted in traction. In contrast to the media topic where it seems like the better strategy is to remain general and not use specific people, for the law and order theme the reverse is likely better. In other words, the tweets that show Trump shaking hands with police or discussing a very specific police topic seem to get the most traction.

The theme of Trump being a normal person is real, it is not a joke. Surprisingly, a simple tweet of a travel agenda that included Trump with a bucket of KFC was one of his most popular tweets. There was no politics here or

attacks on his opponent. And yet it received a lot of traction. It seems like Trump should really increase any tweets or activity that shows that he is a real person, doing normal things. Because Clinton's strategy as we will see later includes attempting to paint him as a money-grubbing charlatan with a dubious character who is surrounded by racists, it would make sense as a counter-measure to show Trump doing normal things in an attempt to defuse some of these attacks. Certainly the unusual popularity of the tweet showing him doing something as normal as eating KFC shows that people want to see more of this.

Clinton's Tweet Popularity

Clinton's twitter stream is somewhat surprisingly more active than that of Trump in terms of absolute number of tweets. However, the number of 'likes' associated with her tweets is far below those of Trump. In fact, it is not even close. This follows a trend highlighted in other social media streams and analysis, namely that when on-line people tend to favor Trump. And, in polls people tend to favor Clinton.

The following chart show Clinton's tweet 'like' distribution.

Chart 3: Distribution of Tweet 'Likes' for Clinton, July 27 to August 27, 2016
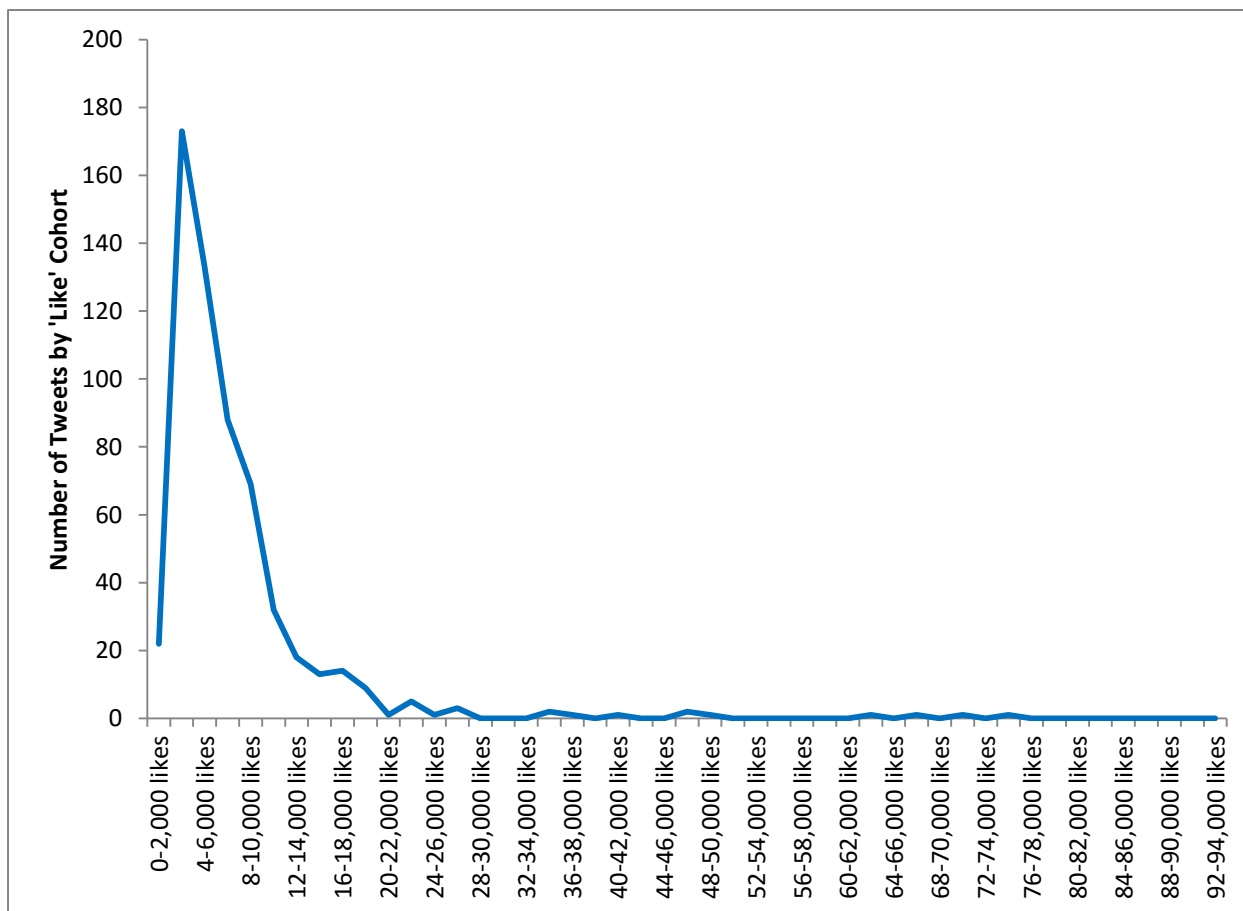
Source: Twitter

The distribution of Clinton's 'Likes' is slightly unusual. Its skewness is 4.8, as there are clumps of 'likes' going far off onto the right side of the distribution. Its kurtosis is 31.1, which is extremely high. Essentially, the distribution is very concentrated around a single cohort (4 – 6,000 likes). In other words, when Clinton tweets, it is highly likely the number of 'likes' for that tweet will fall within this relatively tight range or very close to it. This certainly makes outlier analysis easier, but it does not bode well for the health of Clinton's twitter stream. It does not appear that her twitter stream is receiving the type of interaction with followers that defines a healthy social media site. If the peak cohort were extremely high number of 'likes', we might rethink this statement, but it is actually fairly low considering her raw number of followers on twitter. In short, the fact that she is not generating many likes and the fact that those likes are extremely concentrated around a relatively low cohort is not good news for Clinton. She is just not generating very much excitement on twitter which does not provide a very indication for getting people to the polls.

In terms of Top Topics, Clinton's are fairly easy to identify.

Table 2: Top Topics of Clinton's Tweets from July 27 to August 27, 2016 as measured by 'Likes'

TOP TOPICS	'Likes' in thousands
Personal Connections w/ Obama	
Hugging Obama at DNC	71
Happy Birthday To Obama	27
Obama not founder of ISIS	22
Defending Obama	18
Obama and Hillary - friendship	17
Image of Hillary, Bill, Obama	17
Obama and Hillary - successor (pass baton)	14
Personal Connections w/ Family	
Hillary and Bill hug at DNC, 'next president – and her husband'	66
Hillary watching Chelsea at DNC	62
Happy Birthday To Bill Clinton	48
Happy National Dog Day, image Bill with dogs	17
Flashback Image with Bill	13
Trump does NOT have temperament / character	
Anyone provoked by tweet should not have nuclear codes	40
Should not condone violence	34
Some Republicans will vote for Hillary due to poor character	27
Captain Khan and family represent best of America	27
Not role model for kids	23
Changed Campaign manager often, bad judgment	16
Trump says knows more than Generals	10
Women, Women's Rights, Pro-Women	
First woman candidate of major party	24
19th Amendment	22
Hispanic Women from US competing	21

First woman from US won gold, from 1900	20
Women winning at 2016 Olympics	20
Muslim Women from US competing	20
Girl for her birthday wants to meet Hillary	18
Let's break glass ceiling	15
Women winning at past Olympics	14
Woman Card	13
Meet the women of the senate	11
Human's rights are women's rights	7
Trump and Taxes	
Buffet says Trump should release taxes	18
Many think Trump is hiding something in taxes	13
Trump could be paying zero taxes	11
Tax returns, Trump should release	10
Tax cut for rich like Trump	7
Trump campaign is racist, sexist, mean, hateful	
Breitbart headlines are mean	15
KKK likes Trump	15
Breitbart hire at Trump campaign	13
Trump's bigoted rhetoric	11
Running Campaign of hate and fear	10

Source: Twitter

The first thing that likely jumps out is the importance of Obama. Clinton's relationship with Obama is undisputedly one of her largest assets. By mentioning his name she tends to get an outsized positive response on twitter. She should stick very close to him and tweet often about him. Any joint pictures or statements of Clinton and Obama would likely be very positive for her tweet popularity and social media interactions.

A very close second place is her family. Note when the election began, most thought she would derive much benefit from Bill Clinton, but he has essentially turned out to be an almost liability as an individual. His social media influence has decreased during the campaign and prior attempts by Hillary of trying to include him in direct campaigning have not gone well at all.

However, people seem to really enjoy the idea of Hillary having a sweet family. Perhaps they recall some of the glory days of the 1990s when they see pictures of her family together, it is not clear. Certainly, Hillary has one of the most recognizable families in the country. One of Hillary's highest rated tweets is her watching her daughter give a speech at the DNC. More tweets showing her loving relationship with her daughter is likely a good strategy to help to humanize Hillary. Also, images or tweets that show Hillary and Bill together or Bill as a supportive spouse would likely be very good as well – note, not Bill as a politician but as a spouse.

It should be noted that even non-relevant tweets, like a Happy Birthday Bill tweet got many 'likes' as did a tweet of an image of Bill with their dogs. Even a tweet of an image of them having breakfast together from the 1990s was more popular than the vast majority of her political tweets. Such non-political feel-good tweets get many more 'likes' than almost all those criticizing Trump. Think about that for a little while. People enjoy seeing Hillary happy and like seeing Bill as the tamed spouse supporting his wife's career. This plays very much into their

narrative of strong women, it is shocking they have not tried to exploit this earlier. Having analyzed what her followers 'like', an image of Bill carrying Hillary's briefcase or her coffee could likely become the most 'liked' tweet of her campaign. This is what her followers want, not Bill defending some decision he made in the 1990s or accepting some pseudo-official title as an economic czar.

The next Top Topic is that Trump apparently does not have the temperament or character to be president. This topic has received a significant amount of traction. It is fairly loaded however and not entirely clear if it will be effective. Recall that many of Trump's Republican competitors tried the exact same tactic. At that time too, it got traction, but not enough. Perhaps it will resonate better with Democrats and Independents. Ammunition which the Democrats could exploit further is quotes concerning his temperament made by other Republicans. Additionally, nuclear codes in the hands of someone with a rash temper could also be relatively 'new' material against Trump and has so far been effective. A major plus with this theme is that Trump has little defense if these attacks stick to quotes from well-known Republicans and other individuals or entities seen as neutral. The temperament theme will likely be one of Clinton's most lasting attacks and will likely take her through the debates and into the election.

The theme of women, women's rights, women in leadership, and anything pro-woman is a very strong theme for the Clinton campaign and for her twitter stream. The gender breakdown of her followers is not clear, but certainly tweets that fall in this category receive many more 'likes' than most any of the tweets highlighting attacks on Trump. People enjoy the idea of electing the first woman president and tend to enjoy reading tweets that are uplifting for women. Seemingly generic tweets, such as one mentioning the first American woman to win a gold medal in the 1900 Olympics received more 'likes' than any tweets concerning Trump's taxes or his apparently 'racist' and 'hateful' campaign. The same can be said about tweets highlighting an American Hispanic women and an American woman in a hijab competing in the Olympics. Hillary tweets that are overtly pro-woman tend to receive outsized attention and more 'likes'. If she is looking for social media attention, she likely will want to amp up her pro-woman talk even further.

It would be easy to think that there could be an exaggeration here concerning a pro-woman stance. If you look at Clinton's actual twitter stream however you will see its heavy slant towards women. As for example, she tweeted exclusively about women during the Olympics. Of the 17 Olympic-related tweets, all 17 highlight exclusively women. Tweets include highlighting the:

- first US woman to win four gold medals,
- first US woman to win gold in shotput,
- first US woman to compete in Olympics in a hijab,
- first US-born Hispanic woman to compete on Olympics gymnastics team since 1984,
- first US women's gymnastics team to win the gold, from 1996 Olympics,
- first US woman to win gold in long jump, from the 1990 Olympics,
- first US woman to win three gold medals in track and field in a single Olympics, from 1960 Olympics,
- first US woman to win a gold medal, from the 1900 Olympics,

The list simply looks like someone got a list of women's firsts in Olympic history and then compiled them into tweets. These are frankly fairly generic tweets. But, they received on average many more 'likes' than Clinton's political tweets. The underlying theme here of course is not only women doing important things, but being the FIRST to do them. The message is not subtle at all as Hillary will be if elected the first female US president.

The remainder of the Olympic tweets includes mostly congratulatory ones for US women competing. There is a single tweet that mentions Michael Phelps but it is questionable to categorize it as a true Olympics tweet as it is used to attack Trump whereas none of the Olympic tweets mentioned do so. In short, Clinton stayed on message when commenting on the Olympics – focus on women's achievements and especially women's firsts. This is one of her core messages, and according to our analysis of which tweets have received the most 'likes', this strategy has been working.

In comparison, the other Top Topics that target Trump do not appear to be gaining that much traction. One focuses on his taxes, essentially that he should release them but also created to plant concern about why he does not want to release them. This has gained only marginal traction. The most popular tweet from Hillary on this topic actually highlighted Warren Buffet saying that Trump should release returns. This is a familiar pattern in that Hillary's allies are very well known and popular and them stepping up to defend her or deliver her message often receives more 'likes' than Hillary giving the same message. She really should increase her use of these famous surrogates.

Another topic focuses on Trump's campaign being apparently racist, sexist, mean and hateful. This has received some traction but surprisingly little for the amount of tweets under this broad theme and the intensity of the accusation. In US politics, there are very few things worse than being accused of as being a racist, and yet this theme has not received as many 'likes' as expected. This seems like more of a talking point and not something that will result in much further social media traction.

Comparing Clinton and Trump Twitter Streams

As with many other social media metrics, Trump dominates in the number of 'likes' his tweets receive. The comparison is fairly stark. As we will see, a tweet from Trump that is on the lower end of number of 'likes' would be a fairly popular one for Clinton. If it were not for the polls showing Clinton as having an advantage, data such as the number of 'likes' per tweet certainly would provide an unquestionable upper hand to Trump.

Chart 4: Clinton and Trump Tweet Popularity by Number of 'Likes' per Tweet
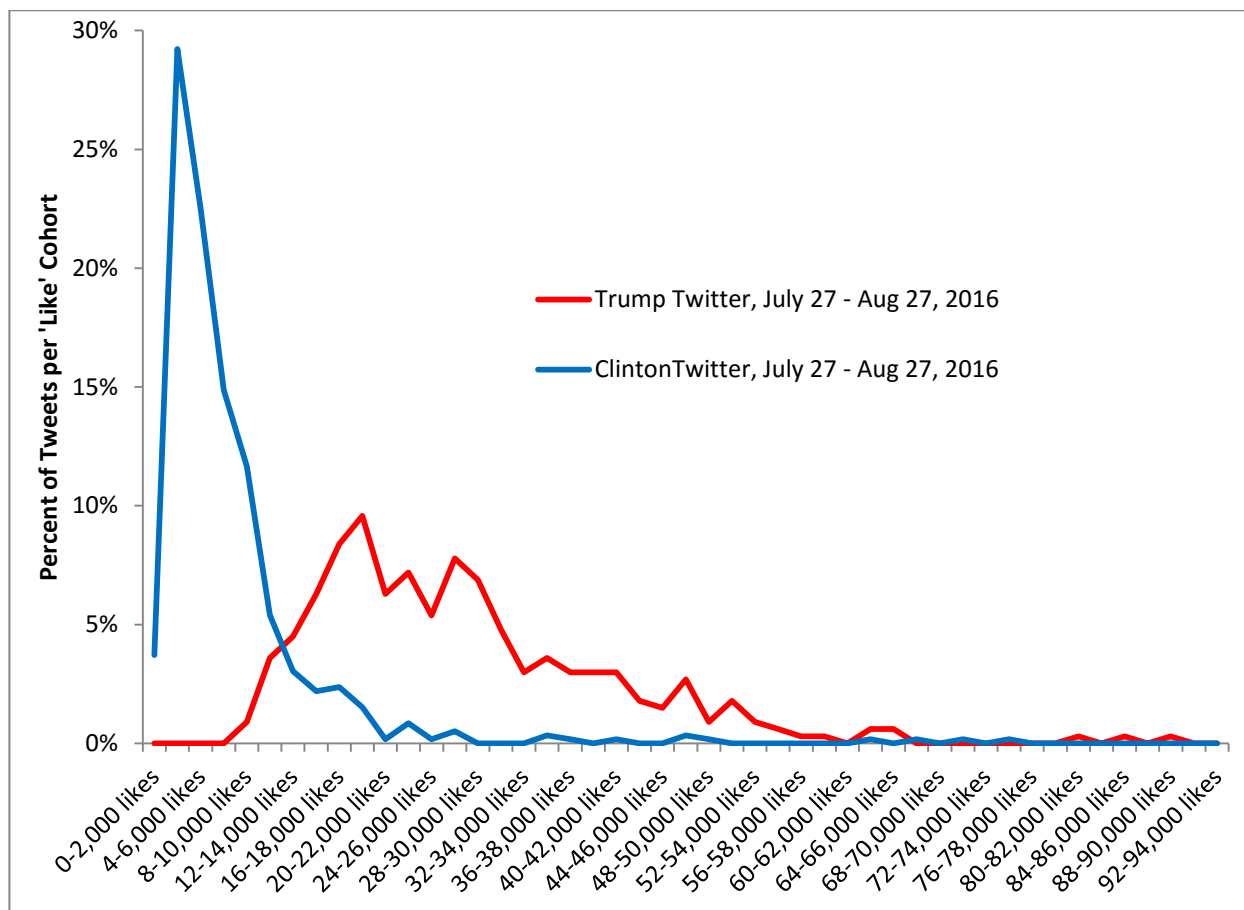
Source: Twitter

The average number of 'likes' for a Clinton tweet during this period was 7 thousand versus 27 thousand for Trump. On average, **Trump's tweets are almost 4 times as popular**. This is truly an amazing statistic.

It should also be noted that Clinton held the advantage during this time period which began on July 27th so it actually includes the last two days of the DNC. These days featured speeches from Chelsea Clinton, Obama, and Hillary Clinton. In other words, there were plenty of tweets about the DNC and these speeches offered many photo opportunities, soundbites, and videos that would almost by definition be popular among Democrats. In fact, if you look at Clinton's highest rated tweets in terms of number of 'likes' in the Top Topics table you will see that the top three come from the DNC.

Additionally, during the sample period, Trump's polls numbers collapsed. After the DNC, almost every poll showed Clinton's lead swell. Many in the media all but threw in the towel for Trump. Again, this time period should have provided Clinton an advantage. But these numbers show that when it comes to on-line activity Clinton tends to have more of a problem – she just does not generate the same excitement and support, as measured by tweet popularity.

If you believe that social media activity is a reasonable indication of interest and if you believe that social media activity can be used to forecast the success of a variety of things like new product releases or even stock performance, then why wouldn't it be able to equally forecast political elections? Again, if it were not for the polls showing something completely different, this would not even be a question – we would just assume Trump would be the expected victor.

Analysis of Top Topics, Clinton versus Trump

Translating social media insights into strategy is not always easy. Basically it is the challenge of taking quantitative data and making it practical and useful. Comparing the twitter streams Top Topics offers this challenge.

Clinton's primary strength as measured by tweet popularity is her relationships. Trump will need to counteract this.

She would if elected be Obama's successor. It is in both of their best interests to get Clinton elected. Obama knows this and is doing everything possible, beyond what previous lame duck presidents have in order to get their successors elected. Clinton should get even closer to Obama and tweet as many photos with Obama, defenses of Obama's policies, and anything else to try to bring back the excitement that he generated during 2008 and 2012.

Trump will likely counteract this by tweeting footage from the 2008 primaries showing Clinton and Obama as adversaries. Anything that can show there was once a wedge between them, even old news articles or video clips, will likely be dug up and tweeted by Trump.

Her relationship with her family was also shown to be a big draw on twitter.   Clinton benefits enormously by being seen as having a solid family life. She likely benefits more than other candidates as the nation knows her husband and daughter so well. They are familiar faces to almost everyone. As previously mentioned, Bill Clinton appears to be morphing into the beta-male version of himself who will play the supportive spouse role. Hillary's supporters, the majority of which are women, will love such a role for Bill. Judging by how he has been acting and how the number of 'likes' of tweets showing Bill being more neutral and subdued, this strategy should produce traction for her. Also, pulling Chelsea closer especially during public appearances would work out well for Hillary – anything to show her as a supportive mother and Chelsea as a supportive daughter seems to activate her base and help her on social media.

Trump was very aggressive against Bill Clinton earlier this year when the Clinton campaign allowed Bill to take a more active role. Trump neutralized him very quickly by bringing up his previous scandals and Bill has not bounced back. His role appears to be more supportive and neutral. Trump will have to further neutralize Bill's positive effect on the campaign as the supportive spouse. He will likely remind the nation of the strained marriage during the Bill's sex scandal and the problems that caused for Hillary. The fact that Hillary's assistant Huma Abedin recently split with her husband over a cheating scandal offers a lot of ammunition for Trump just a few months from election-day. Trump will be successful if he can draw parallels between these cheating scandals and place doubt in people's minds of who made the correct decision – Hillary for staying or Huma for leaving. If this becomes a national discussion topic surely it will chip away at the Clinton's relationship advantage.

Taking the relationship topic one step further, we should note that many of her most popular tweets included other notable people.   So, her most popular tweet about Trump and taxes highlighted Warren Buffet. One of her most popular tweets regarding the fact that if elected she would be the first woman president in the US quotes Meryl Streep. Similarly tweets that contain direct comments by or video clips of Biden, O'Maley, Warren,

15

Michelle Obama, and a list of non-political celebrities tend to receive significantly more 'likes' than those from Clinton on similar topics.

It is no secret that Clinton is not a very popular and is not a well liked candidate. It is also no secret that she has many more political and celebrity surrogates than does Trump. She should adapt to this situation and radically increase the messages given by these surrogates as it is a way to increase enthusiasm and interest in her campaign and to get her message out. This might be a bit of a bitter pill to swallow, but using tweet 'likes' clearly shows that people literally 'like' getting the same information / statements about or concerning Clinton from her surrogates and not from her.

Trump does not have much of a defense for this. He can show how Hillary and Obama fought in 2008, which could drive a perceived wedge between them. He could also depict the Clinton family as not so happy and Bill as a philandering spouse neutralizing that advantage as well. But if an army of Clinton surrogates appear in a semi-organized fashion, giving out her message and not having her message depend on her delivery, there is not much Trump can do about that. He cannot attack Buffet, Warren, Michelle Obama and half of Hollywood at once.

Clinton's next Top Topic was depicting Trump's temperament and character as unsuited for the presidency is likely her best attack and this should increase substantially. There is plenty ammunition here and the media appear to like to report on this topic too so there could be follow-through. Also, Trump cannot do much to directly refute these claims as most are based on imaginary fears of something that may or may not occur in the future. For instance, implying that Trump having the nuclear codes puts you at risk – this tweet was one of her most popular. This theme has gotten a lot of traction with 'likes' and appears to be popular with her base. And, Trump cannot do that much to defend against it. The biggest drawback is that Bush and Cruz tried very similar type attacks during the primaries and they were not able to get enough traction with them to defeat Trump. The bet is that painting Trump as having the wrong temperament will impact Independents and Democrats more than it did Republicans.

The 'Woman Card' seems to be Clinton's go to topic that is constantly on in the background. Even when it comes to something as apparently gender-neutral as the Olympics, her campaign exclusively focuses on women. And, as previously stated, these Olympic tweets were some of her most popular, though having nothing to do with the election! Clinton's base seems to really enjoy the focus on women. Other popular tweets included a little girl wanting to hug Clinton for her birthday, a little girl being held by her mother who wanted to be at the DNC when Clinton accepted the nomination, and another little girl asking Clinton if she will get paid the same as a man if she becomes president. Given the popularity and generally non-abrasive tone of these tweets, Clinton should continue to and maybe even increase her use of pro-women themes. Women's firsts could also become a leading topic as well, which would tie in nicely to Clinton being the first female US president if elected.

Trump will likely counter attack by showing how many other countries have had women presidents, so the US election should be more about the candidates and less about electing the first woman. A more risky strategy would be to accuse the Clinton campaign of sexism, which is the reverse of what you normally would hear but it could be just as effective. Certainly this is risky for Trump to attack, and perhaps why the Clinton campaign feels so secure with it.

Surprisingly, specific attacks on Trump, outside of him not apparently having the correct temperament, do not appear to be going so well. They seem spotty at best – a little traction here or there but nothing too reliable. Trump making his products abroad, a too-close relationship with Putin, and his apparently racist and sexist

16

campaign do not really get consistently high 'likes'. This could be one of the main problems for Clinton's campaign in that if her other Top Topics fail, she is left with very little.

Having analyzed the popular tweets on both sides, having broken them down into Top Topics, and having studied how Trump could neutralize some of Clinton's advantages, it seems like Clinton could be left at the very end of the campaign with the battle cry that Trump is simply not fit to be president due to his temperament and character. It seems like Trump could neutralize many of the advantages she gets from her relationships and from her pro-woman stance. Many might vote for Clinton due to her relationships and/or to vote for the first woman president, but it does not seem like they would be enough to swing the election if Trump works to neutralize them further.

As for areas where Clinton can focus, one is to copy some of the general moral outrage from Sanders. His tweets are laced with moral outrage over all types of topics in which the government should do more. Clinton has had some success with this already but it has not become a real focus yet. Her 'epipen' tweets have generated many 'likes'. These tweets have the moral outrage element and the proposed cure of increased government action. Likewise, her tweets that include LGBT issues could be increased. Her tweet of LGBT kids getting bullied was actually one of her most popular, but she needs to create a theme around it.

Trump's most popular topic is how unfair and biased the media has become.

This is a huge 'like' getter for him. Assuming the media is biased, this will only get him back to neutral as his coverage by the media will be against him. This topic will continue to be niche Republican sticking point until / unless it is shown that the media truly is biased. Up until this point, there has been smoke without fire, or at least from the general public's perspective it does not seem like a proven case yet. As such, Trump is playing to his base and not to the general populace. If he can effectively depict bias, it could be a real winning topic for him in the general election.

Clinton can actually play this topic as well. She can start to complain intensely about how the media is hounding her about her unauthorized server, deleted emails, the Clinton Foundation, and Wikileaks and that she needs to prepare for the debates or the election and that this is unfair. It would be very interesting to see how this would play out. Certainly it could help to neutralize some of the Republican gusto for complaining about media bias.

The anti-Obama theme is a strong one for Trump. However, Obama still has very strong support within the Democratic base and Clinton will likely take the opposite position as Trump on whatever he says about Obama. This topic though very popular with Trump's base, will likely be neutralized in the general election.

Hillary is unfit for the presidency could be his strongest and most popular argument. This combines Clinton's scandals, her questionable behavior, and her health. An interesting side note here is that Trump's most popular tweets on these have come when he has used humor. For instance, when people were asking where Hillary was, Trump responded 'sleeping'. And, when discussing Hillary's deleted emails, Trump proposed that maybe Russia has them. Now whether you think this humor is funny or not is another point – more importantly Trump's humor amplified his message.

The themes of Law and Order, of Borders, and of Minority Outreach could come together – maybe it can all morph into Trump as the Safe and Secure candidate. Part of this is that he seems to be trying to siphon off even a small amount of voters that traditionally vote Democrat. As African-Americans have historically voted in excess of 90% for Democrats, even a small decrease could mean a lot. This strategy is unlikely to swing many votes, but

really Trump does not need that many. A few percentage points could actually make the difference in a tight race and falling from such high Democratic voting rates seems feasible. Additionally, Clinton appears to be having a tough time coming up with a counter attack to this. <u>Trump's Law and Order, Borders and Minority Outreach themed tweets are getting far more 'likes' than Clinton's tweets about Trump's campaign being racist. It seems like these have some room to move and Trump will likely increase them going forward</u>.

The last main themes are Trump is a normal person and Trump doing presidential things. Basically, tweets showing Trump doing real things have done very well. In other words, tweets not just showing him speaking to thousands of people or flying on his own plane. An example tweet might show him eating something normal people eat or meeting some normal everyday people in crisis – not as their boss but as a normal guy who cares. <u>These types of tweets have become very popular. A very large added benefit is that this theme counteracts many of the Clinton tweets that paint him as out-of-touch, money grubbing, and of questionable moral character</u>. Again, an added benefit is that Clinton counter attacks will be limited. Who can mock someone for having a hotdog at the World Series?

Summarizing, by looking at the popularity of tweets, we can determine which themes have emerged as the most popular for the candidates. There will likely be new themes that come up from now until the election, but not that many. In fact, we are fairly close now and these themes are close to maturity. Many of these themes have opposing forces or strategies that can neutralize them. For instance, Obama's support of Clinton will likely be neutralized by Trump's criticism of him. Trump's complaints of a biased media could be neutralized by Clinton complaining about the same. Clinton's pro-woman stance could be neutralized by accusations of sexism. These will likely not be completely neutralized in either direction, but neutralized enough so that they will not make oversized impacts on the general election.

Taking out topics on each side that will likely cancel each other out or become so diluted so as to not make a significant difference, what we will likely be left with, as based on tweet popularity, topic analysis, and qualitative analysis, includes the following not so positive picture:

Clinton Campaign	Trump Campaign
Trump is _unfit_ for the presidency, due to his temperament and character. You should be very scared at the thought of a Trump presidency. Even lots of Republicans agree. Plus, here is a list of famous people saying the exact same thing. Lots of famous people do not support Trump. We are _morally outraged_ that anyone would even vote for Trump. We are morally outraged about other stuff too, like _____. You must vote for Clinton, it's our last chance.	Clinton is _unfit_ for the presidency, due to her moral character, scandals, and health. It's surprising she can even be in the race. Trump will bring safety, security, and jobs back. Trump is a normal guy just like the rest of us, see all the normal stuff he does. He is not scary. We laugh and _make fun_ of how Hillary disappears for long stretches and the media does not cover her scandals, inspired by Trump's humor. You must vote for Trump, it's our last chance.

This should turn very nasty. You might think that it is already nasty, but looking at the previous outline shows that <u>the debates will mostly be baseless and with little policy discussions. If the table is correct it will be a battle of **which candidate is more unfit for the presidency**. Clinton will use the Democrat's chosen weapon of moral outrage and Trump will use his own brand of slightly questionable humor. Each side hates the weapon of the other.</u> The moral outrage will inflame the Republicans and the questionable humor will set off Democrats.

In a final ironic twist, it will be Trump who will have new policy issues to discuss and not Clinton, who is known for her command of issues and her wonkiness. Even in the boiled down table of what the campaigns will have left at the end, it is Trump who has new policies. The Safe and Secure theme, which would combine border security, keeping jobs at home, law and order, and some sort of special deal for urban minorities, offers many new policies.

Clinton, if this analysis is correct, will have to take the mantel of attacking these policies and defending the status quo or the Obama-economy and policies.

One last point is that trade is expected to explode as a theme on social media at some point before the election. Perhaps it will be one of the main topics at one of the debates. Expect for this to happen. On this topic, Trump will likely hold the upper hand as he established it as one of the main talking points of the Republican primaries and Clinton will have to defend NAFTA which was passed during her husband's presidency. Clinton again on the defensive does not bode well for her.

Wikipedia Page Views
Let's admit it, we all rely on Wikipedia. When you want to know about something or even you want a refresher on a topic, you go to its Wikipedia page. The information is extremely useful and almost always up-to-date. It should come as no surprise then that trends in Wikipedia page views are extremely useful in analyzing levels of interest in a topic.

In a sense Wikipedia page views are similar to Google Search Trends. However, with Google you can mostly determine what people are asking about. With Wikipedia Page Views you know exactly where they are going. So someone searching on Google for "Hillary Clinton" might simply be interested in the election or about some news item concerning her. With Wikipedia, we can determine if they ended up learning more about the party (like a view of the "Democratic Party" page), the candidate ("Hillary Clinton" page), her campaign ("Clinton Presidential Campaign 2016" page) or even her slogan ("I'm with Her" page). We can do the same for the Republican candidate and compare the results for each. <u>The core idea is the more views to a specific page means more people are interested in that topic.</u>

Measuring interest in a candidate by looking at page views is fairly straightforward and leverages the power of tracking on-line activity. This method basically tracks people's daily actions around a topic. There are no sample sizes, sophisticated poll methodology, or margins of error here. This data includes the hits to Wikipedia pages and effectively measures interest level on a national scale without any filters, adjustments, or black boxes. <u>Instead of random samples of the populace amounting to somewhere around 1,000 people in a poll, we can look at +100,000 daily hits on Trump's Wikipedia page.</u>

The weakness with Wikipedia page views is that there is an element of discovery in addition to pure interest. For instance, an argument can be made that there could be more searches for a less known candidate, in this case Trump as Clinton has been known in the political sphere since at least the early 1990s. This is a strong argument, but as we will see the strength of the Republicans and Trump on Wikipedia appears to transcend this argument due to its consistency.

First, let's look at how the candidates compare head-on, or comparing the page views for Donald Trump and Hillary Clinton. This shows the page views for Donald Trump divided by those for Hillary Clinton. A level of 100% is parity, anything over that shows strength for Trump.

![zettacap]

Chart 5: Wikipedia Page Views for Donald Trump / Hillary Clinton, 1 month moving average


(Trump / Clinton), 1 month moving average
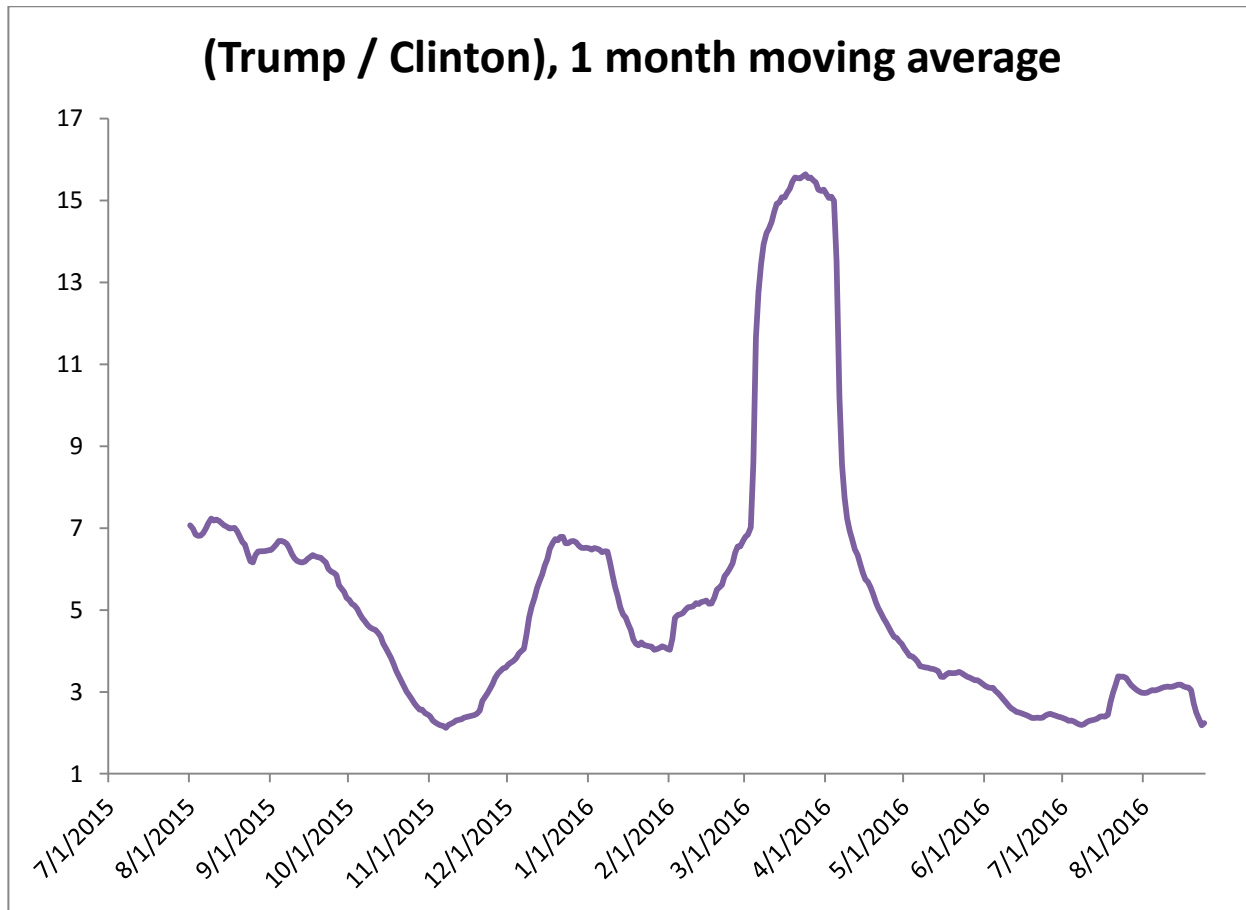
Source: Wikipedia

When it comes to comparing the page views for the candidates, there really is no comparison. Page views for Trump far exceed those for Clinton. On this metric, it clearly seems like there is much more interest around Trump's candidacy. Even making a discount for the fact that Clinton is more well known in the political realm does not seem like it could make up for such a disparity. For example, over the last week of the chart, the average number of daily page views for Trump was approximately 75,000 versus 23,000 for Clinton. There simply seems to be more general interest around Trump.

Second, let's look at page views for their campaigns. So, we will compare page views for Donald Trump Presidential Campaign 2016 / Hillary Clinton Presidential Campaign 2016. Again, we see that interest in Trump is much higher.

Chart 6: Wikipedia Page Views for Donald Trump Presidential Campaign 2016 / Hillary Clinton Presidential Campaign 2016, 1 month moving average

(Trump Presidential Campaign / Clinton Presidential Campaign), 1 month moving average
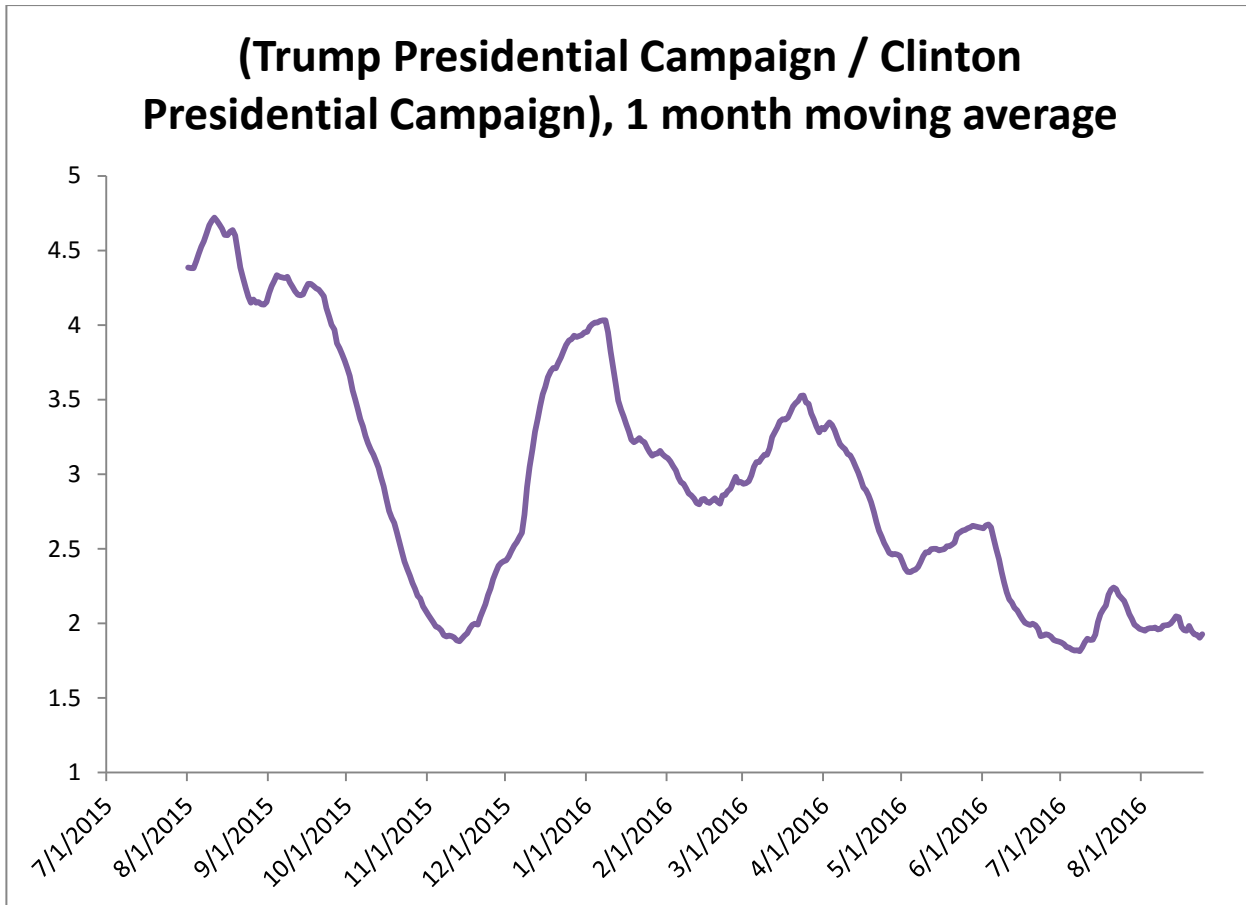
Source: Wikipedia

Similar to the page views of the candidates as individuals, the page views for Trump's campaign well exceed those for Clinton's. Even though it has fallen more recently, it has fallen to a level of around 200%. In other words roughly twice as many people are visiting the Trump Campaign Wikipedia page than those visiting the equivalent for Clinton.

Third, let's look at their most well-known slogans. For Trump it is "Make America Great Again" and for Clinton it is "I'm with Her".

Chart 7: Wikipedia Page Views for "Make America Great Again" / "I'm with Her", 1 month moving average
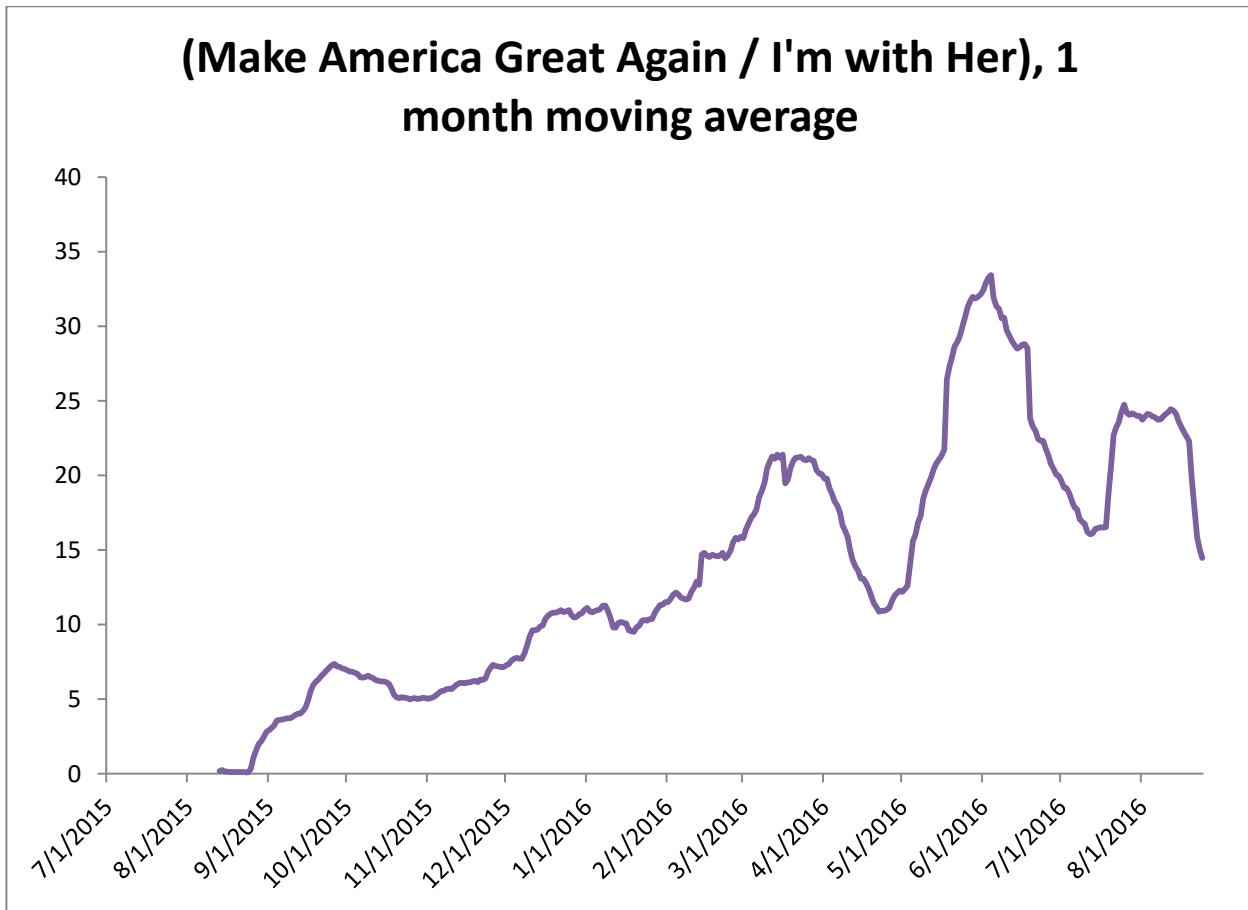
Source: Wikipedia

The comparison is not even close. On average during the last week of the chart, there was over 10 times the number of page views for Trump's campaign slogan than that for Clinton's. Even making allowances and discounts, it would be difficult to argue against the fact that there seems to be more interest in the Trump campaign.

Fourth, let's look at the page views for the Republican Party versus the Democratic Party.

Chart 8: Wikipedia Page Views for Republican Party / Democratic Party, 1 month moving average

(Republican Party/ Democratic Party), 1 month moving average
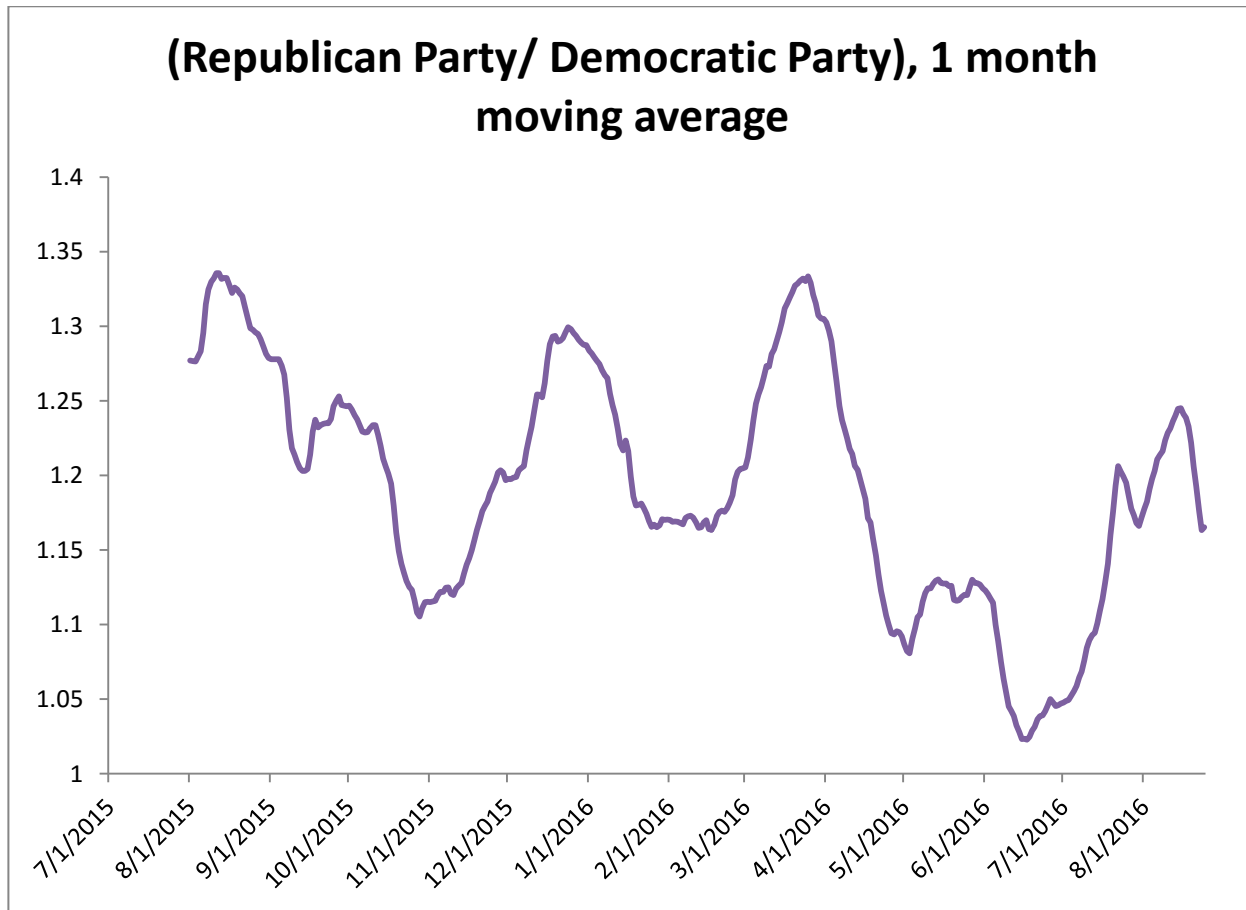
Source: Wikipedia

The results of this comparison are much tighter. However, the page views for the Republican Party are consistently higher than those for the Democratic Party. The difference is very small in comparison to the other trends but the consistency is impressive. Also, recall that the there are many more registered Democrats than Republicans, so the Republican Party page consistently receiving more page views is still more impressive. Again, more people are showing interest in the Republican Party this election season.

Summarizing, using Wikipedia Page Views as a metric to measure general interest, there are many more people interested in Trump, his candidacy, and the Republican Party this election season. This bodes well for the Republicans going into the general election. Though it is possible that these trends could invert, it does not seem likely so close to the general election as these trends have consistently shown Trump/Republican strength. Also, it is possible, though not likely, that these trends could be produced by a greater curiosity surrounding Trump as his candidacy is highly unusual. If these trends could be explained away due to curiosity, there would likely be more spikes around gaffes and questionable situations, but the pro-Republican trends have dominated fairly consistently on a variety of topics throughout the race.