

The role of new media in determining how politicians shape their presentation to the public has been debated from the advent of television to the creation of the Internet. While one role that media traditionally serves is as a watchdog for politicians (Sutter 2006; Setala 2006), advancements in media, particularly social media, enable politicians to directly participate in building their own images (Gurevitch 2009). Theoretically, advances in new media technologies such as Twitter should allow politicians to shape their public images in order to win favor with constituents by catering to constituents' needs. Historically, however, politicians have failed to recognize the power of new media.

As shown by studies regarding politicians' use of personal websites, a politician's ability to directly influence constituents' perceptions has gone underused, despite the increased ease of doing so. Ferber, Foltz, and Pugliese (2003, 2005) found that many politicians used websites as an extension of traditional public relations mechanisms with limited interactivity. Variation amongst websites confirmed that political information such as policy stances could easily be provided to users; however, most politicians failed to take advantage of this. Contini et al. (2004) build on this concept through findings in which personal information, constituent services, and communication and feedback were featured more prominently than issue information on personal websites. Taken together, these studies suggest that politicians use media sources that they control to focus on their non-issue images. Based on this pattern, one would expect that few politicians realize the image shaping power of new media forms such as Twitter, as Twitter is too new of a technology. For the politicians who do realize the potential of new media forms like Twitter, one would expect a continuation of self-promotion on a personal or constituent service basis as opposed to self-promotion as policy maker who clearly and concisely state positions on issues.

Given the relatively new prominence of Twitter as a means of communication, it is not surprising that little extensive research exists pertaining to politicians' use of Twitter as opposed to personal websites or television. Studies that have been done (Glassman, Straus, and Shogan 2010)—while serving as a good basis for further study—do not adequately address possible motivations behind politicians' Twitter usage. Looking face-value at what a politician tweets about, while offering insight into how politicians use Twitter, does not address why politicians tweet about the topics that they do. Our research aims to answer the question of why some politicians tweet about policy as opposed to constituency services, for example. We will examine how factors such as constituency demographics, politician demographics, and the number of followers each politician has influence what types of messages are tweeted. While we will discuss traditional focuses regarding politicians using Twitter such as the frequency of tweets, we will move beyond simply providing data into looking for possible explanations. Given the high percentage of politicians who use Twitter at the time of this study, conclusions drawn from this study should be robust.

Background

In order to determine how politicians use Twitter, it is first important to understand what Twitter is and how it functions. Twitter allows users to send tweets, or messages, comprised of 140 characters or less to all of their "followers." Each Twitter user creates a username that appears along with each tweet on their "followers'" feeds. All tweets made by a user or sent to another user can be traced back to the original "tweeter" on his or her newsfeed. Most Twitter feeds are public and can be looked up by username on search engines such as Google. Users are able to "retweet" tweets made by others so that the "retweeted" tweets also become attached to their own accounts. While "retweeting," users may either add commentary or simply share

someone else's opinion, story, etc. Finally, Twitter users are able to link together tweets that share a common theme through hashtags(#). Hashtags that appear frequently on Twitter are considered "trending."

Given the relative ease of tweeting, it is not surprising that most politicians have added Twitter to their political arsenals. When Glassman, Straus, and Shogan completed their congressional report for Twitter use from August 2009 to September 2009, only 39 Senators and 166 Representatives held Twitter accounts (2010, pg. 6). As on January 2013, 40 governors, 388 representatives, and 85 senators have Twitter accounts. In approximately four years, it has become the norm for politicians to have a Twitter account. While some politicians have yet to join Twitter, most recognize the popularity of Twitter in both a media and constituent context. News programs routinely devote time to what is "trending" on Twitter. Even when not being directly spoken about, tweets frequently appear across the bottom of a television screen during news programs. Besides news watchers, many people recognize Twitter-related words and symbols, as these symbols have seemingly inundated American society. Most major events are now promoted with a hashtag label.

As Twitter has grown and continues to grow in popularity, it is logical that politicians would attempt to use Twitter as a way to stay in contact with constituents. According to the Pew Internet and American Life Project (Smith and Brenner 2012), 80% of all adults in the United States use of the Internet. Of this 80%, 15% also use Twitter. While this number may appear small and thus of little importance to politicians, it is important to note that from 2010 to 2012, the number of Internet users using Twitter increased from 8% to 15%. In 2 years, Twitter has doubled in growth and, consequently, has come to include more and more of the U.S. adult population as users. Few if any politicians would want to be late in acknowledging the growth

potential for Twitter as it continues to become more and more influential. Even if the percentage of Twitter users was to stagnate, 15% of the U.S. adult population equates to millions of individuals, many of which have the ability to vote in elections.

Table 1. Percentage of Internet Users Who Use Twitter by Demographics

Total adult internet users	15
Men	14
Women	15
<i>Age</i>	
18-29	26**
30-49	14
50-64	9
65+	4
<i>Race/ethnicity</i>	
White, Non-Hispanic	12
Black, Non-Hispanic	28**
Hispanic	14
<i>Annual household income</i>	
< \$30,000	19
\$30,000-\$49,999	12
\$50,000-\$74,999	14
\$75,000+	17
<i>Education Level</i>	
No HS diploma	22
HS Grad	12
Some College	14
College +	17
<i>Geographic location</i>	
Urban	19**
Suburban	14**
Rural	8

Source: Pew Research Center's Internet & American Life Project Winter 2012 Tracking Survey

** = Significant when compared to other rows in group

The importance of Twitter for politicians may be more easily understood when broken down by group as shown in Table 1. Perhaps most surprising is the large percentage of Black, Non-Hispanic people who use Twitter relative to other race and ethnic groups. Partially

explaining the discrepancy between races and ethnicities is geographic location. Urban locations tend to have a high percentage of Black, non-Hispanic people than rural areas. Urban locations also tend to have a higher percentage of Hispanic people than rural; however, Twitter use amongst Hispanic people is much less pervasive indicating that a factor besides geographic location must be contributing to race and ethnic groups' use of Twitter. Not surprisingly, 18 to 29 year olds compose the largest percent of Twitter users by age group (Smith and Brenner 2012). Based on the demographic information provided in Table 1, one can imagine how Twitter might be used to target specific groups of constituents. News sources often mention the "youth vote" and the "Black vote," both groups that are highly represented on Twitter. It not only logical that politicians would attempt to give themselves all possible advantages through reaching highly represented groups of constituents via Twitter, but that politicians would tailor their tweets to meet the interests and demands of their "followers." Politicians could do so easily by examining the demographics of their constituents.

In order to determine how politicians use Twitter, we examine tweets made by governors, representatives, and senators from February 2013 to April 2013 at both a state and district level. While it is uncertain who specifically crafted each tweet associated with a politician's Twitter account, we treat all tweets as equal opportunities for politicians to shape their image. "Retweets", while not always containing additional commentary from politicians, have also been included for their image shaping capabilities. All tweets were divided into three categories: political, constituency services, and fluff. Political tweets include policy, opinion, executive, legislative, and judicial matters. Constituency service tweets focus on how a politician serves his or her district. Finally, fluff tweets are apolitical in nature and may include topics such as the weather and sports. Taken together with constituent demographics and politicians'

demographics, we hope to gain a clearer picture of why politicians tweet in the manner that they do.

Hypothesis

Given the ease of obtaining constituent data, we hypothesize the politicians will tailor their tweets in order to meet the demands of their constituents. Given politicians' history of slow technology adaptation (Ferber, Foltz, and Pugliese 2003, 2005), however, we expect the impact of constituent demographics to be small relative to the overall number of tweets made by a politician. We do not believe that four years has been enough time for politicians to successfully use Twitter as a strategic constituent outreach tool, thus we expect constituent demographics to play an even smaller role in tweeting within each category. We do not expect politicians from predominantly Black states and districts, for example, to recognize that their many of their constituents use Twitter, thus the politicians will not focus tweets on constituency services, an area in which Black people have routinely expressed interest given their historically low levels of income (Griffin 2011). Instead, we hypothesize that a politician's own demographics will more strongly influence his or her tweeting. We expect older politicians to tweet less overall, and we expect older politicians to tweet without a strategic pattern. Correlated with this is the number of terms that a politician has served in office, We expect freshman politicians to tweet both more overall and more strategically given the high likelihood that Twitter somehow played a role in their election campaigns. Oppositely, incumbents who have served for five terms, for example, have little to no need to adapt to Twitter. Furthermore, we expect Republicans to tweet more overall given that Republicans have used Twitter more frequently in the past (Glassman, Straus, and Shogan 2010).

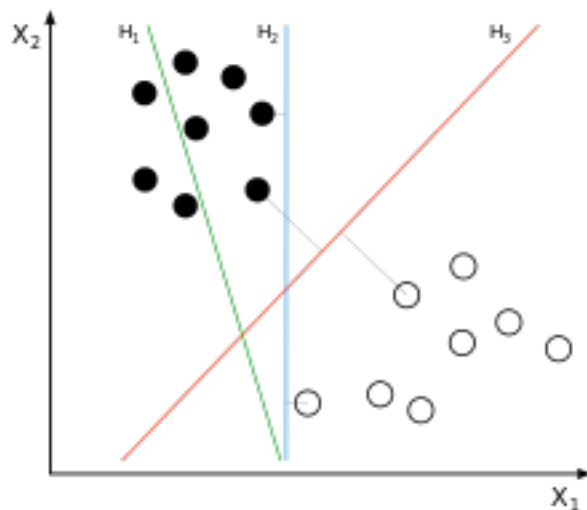
Data and Methods

In order to test our hypothesis, we needed to collect a variety of data. This data falls into two major categories: Twitter data and demographic data. As our dependent variables represent ways in which politicians use Twitter, we needed to acquire a large body of tweets produced by politicians. Since we are testing the hypothesis that personal characteristics of politicians and of their constituents will affect their behavior on Twitter, our other data set is naturally a collection of demographic data about politicians and their constituents.

Twitter Data

Data from Twitter was collected from February 2013 to April 2013. The total amount of governors, representatives, and senators in the United States is nearly 600, and in order to streamline the process of finding all politicians with Twitter accounts, we turned to C-SPAN's curated list of politicians on Twitter, specifically *cspan/members-of-congress* and *cspan/governors*. These lists provide an up-to-date reference regarding national politicians on Twitter. While Twitter provides an interface for collecting tweets directly from lists, this interface uses heuristics to filter tweets based on end-user preference. This would have the effect of artificially shaping and limiting our data set. Instead, we turned to the Twitter Stream API, and set up a listener to watch all Twitter IDs referenced by these C-SPAN lists. This listener collected all tweets produced by these politicians as they were created, and saved them to a database for later analysis. This rough data set was enough to provide the information on raw tweet volume, and also stood as the basis for our dataset on tweet type.

Fig 1. Illustration of Support Vector



H_1 fails to separate classes. H_2 succeeds with a low margin. H_3 is the vector with highest margin.

In order to analyze the content of politicians' tweets, each tweet was coded as political, apolitical or constituency service. With a corpus of over 50,000 tweets, coding all the tweets by hand was out of the question, thus we turned to supervised machine learning techniques for automated categorization of our data. The method we chose consisted of selecting a small, random subset of all tweets, and then categorizing them by hand. We then used this data to feed a support vector machine (SVM) which attempted to fit uncategorized tweets into groups. SVMs are likely to be slightly erratic at first and unable to categorize a majority of data. We used an iterative correction process to compensate for these inaccuracies. Human input was used to select words from the coded corpus with high probability of falling into a specific category. Word such as "Obama," for example, were categorized as political while "basketball," a frequently occurring word given that our data collection coincided with the NCAA tournament, was categorized as fluff. This increased accuracy and allowed for a higher number of tweets to be coded. With this additional input, the machine was run again, producing a larger and more accurate dataset. This process was repeated until all data was successfully coded. The yielded

data set gives us our dependent variables of total tweet volume, political tweets, constituency services tweets and fluff tweets.

Demographic Data

While the data acquired from Twitter represents our dependent variable, one of our goals is to measure the correlation between attributes of politicians and their Twitter usage. This data is broken down into two categories. First, we use politicians' demographic data to examine the effect of politician's personal characteristics on their Twitter behavior. This is useful for analyzing how personal characteristics of politicians affect their use of new technology. These variables also help to provide important controls. Second, we examine demographic information about the constituents of the politicians we are examining. This data is important for determining whether politicians are using Twitter as an effective way to reach constituents and to meet their needs.

Politician-level demographic data consists of five factors, all of which are examined at the state and district level:

Followers. This factor was collected directly from Twitter and represents the number of people that subscribe to a politician. Followers receive all tweets made by the politician that they are following. Follower count represents a user's clout and significance with the Twitter network, thus we include this a control variable.

Freshman dichotomously assesses whether or not it is a politician's first term. This variable was originally coded as the number of terms served, but was recoded in order to avoid collinearity with politicians' age. We expect this variable to be positively associated with the number of tweets made overall and within each category.

Age measures the age of a politician as of April 2013. We expect an increase in age to be negatively correlated with the number of tweets made overall and to not be correlated with any category.

Gender is measured dichotomously with 1 being female and 0 being male. We do not expect a strong negative or positive relationship between gender and tweeting.

Party measures political affiliation. Independents were recoded as Democrats given the necessity of caucusing. A dummy variable was used to measure party with Republicans being 1 and Democrats being 0. We expect this variable to be positively correlated with all measures of tweeting given that the Republican Party is currently trying to reshape its image, thus creating an unclear goal for Twitter use.

Follower data was gathered straight from Twitter at the end of tweet collection. All other information on the demographics of individual politicians was gathered from the Biographical Directory of the United States Congress for senators and representatives, and from the National Governor's Association (cite me) for the state governors.

Demographic data for constituents was gathered from the 2013 Census Bureau's American Communities Survey, which provides detailed data on the demographics of Americans at both a state and district level. ACS data provided the following factors for our analysis:

Median Income. Use of the Internet is no longer limited to the wealthy elite, thus we include a measure of medium income with an eye towards tweeting categories. More specifically, we expect politicians from areas of low median income to focus on constituency service tweets (Griffin 2011). Oppositely, we expect high medium incomes to be more correlated with political tweets.

Achieved BA measures the proportion of constituents who have achieved a Bachelor's degree. We expect politicians from states and districts with a high percent of Bachelor's degree holders to tweet more about politics, as people who are more educated tend to be more interested in politics. We do not expect this to influence the overall number of tweets produced, though.

Black measures the percentage of the population that identifies as being Black. Given that Black, Non-Hispanic people use Twitter much more frequently than any other race or ethnicity (Smith and Brenner 2012), we expect this variable to be highly correlated with the overall number of tweets made as well as with the number of constituency service tweets made (Griffin 2011).

Hispanic measures the percentage of the population that has identified itself as Hispanic. We expect this variable to be significant for the number of constituency services tweets made given that Hispanic people, on average, earn less than non-Hispanic people, and would thus have a greater interest in constituency services (Griffin 2011).

Age 15-34 measures the percent of the population within this age range. This category was formed based off of research completed by the Pew Internet and American Life Project (Smith and Brenner 2012). Given that, on average, people within this age group use Twitter more frequently, we expect this variable to be positively correlated with the overall number of tweets made.

Age 35-59 measures the percentage of the population aged 35 to 59. While people within this age group use Twitter more than people older than them but less than people who are younger, we expect this variable to have a small, positive impact on the number of tweets made.

Age 59 and up measures the percentage of the population within a state or district that is age 59 or older. Given that, on average, people within this age group do not use Twitter in large numbers, we expect this variable to be negatively correlated with the overall number of tweets made.

Population measures the total population of a state and district. We do not hypothesize that the number of people a politician has to tweet to will substantially change their tweeting patterns, as the number of followers that a Twitter user has matters more.

Results

Table 2: Total Tweets

	State-Level	District-Level
<i>Constituent Data</i>		
Median Income	0.001	0.001
Achieved BA	449.330	119.633
Black	27.788	32.048
Hispanic	19.887	10.855
Age 15 to 34	940.719	155.445
Age 35 to 59	1,193.140	−597.457
Age 59 and up	606.788	130.165
Population	0.000	−0.000
<i>Politician Data</i>		
Followers	0.000	0.000
Freshman	−14.494	19.315*
Age	−0.649	−1.222**
Gender (M=1)	10.574	−10.130
Party (R=1)	13.314	5.128
N	135	385
<i>Note:</i> *p<0.1; **p<0.05; ***p<0.01 <i>Source:</i> Twitter, ACS, Bioguide and NGA <i>OLS Regression</i>		

In order to establish a baseline pattern, we first examine how a politician's personal demographics as well as the demographics of his or her constituents influence tweeting. Results can be seen in Table 2 in which N equals the total number of politicians. As expected, constituent demographics play no significant role in influencing politicians' overall tweeting. While some demographic factors, particularly age groups, appear to greatly increase the number of overall tweets, there is not enough information to determine if there is a relationship between age of constituents and overall volume of tweeting. Overall, these consistent with our hypothesis

that politicians fail to consider constituent demographics; if politicians did, we would expect to see Age 15 to 34 increase the number of the of tweets made in a statistically significant manner, as this age group most frequently uses Twitter. In terms of politician demographics, only age and freshman status are related to the overall number of tweets produced within representatives. Interestingly, freshmen representatives of all ages routinely tweet more than incumbents. One possible explanation for this is that in order to get into office, freshmen representatives must, amongst other things, adapt to new technology more quickly in order to get their messages out. This trend may then continue once freshmen attain office. Why this does not apply to freshman governors and senators is not clear.

Table 3: Political Tweets

	State-Level	District-Level
<i>Constituent Data</i>		
Median Income	0.000	0.000
Achieved BA	−0.241	0.118
Black	−0.050	−0.026
Hispanic	−0.102	0.041
Age 15 to 34	0.190	0.384
Age 35 to 59	1.055	0.255
Age 59 and up	−0.253	0.147
Population	0.000	0.000
<i>Politician Data</i>		
Followers	0.000	0.000
Freshman	−0.013	0.011
Age	0.001	0.001
Gender (M=1)	−0.008	−0.019
Party (R=1)	−0.021	0.015
N	136	386
<i>Note:</i> *p<0.1; **p<0.05; ***p<0.01 <i>Source:</i> Twitter, ACS, Bioguide and NGA <i>OLS Regression</i>		

Categorization of tweets in political, constituency service, and apolitical can be seen in Tables 3, 4, and 5, respectively. In all tables, N equals the number of politicians within each category. In Table 3, neither constituent demographics nor politician demographics influence the production of political tweets, including a politician's age and freshman status. Given that an overwhelming majority of tweets made were political, it is surprising that not one variable is statistically significant.

Table 4: Constituency Service Tweets

	State-Level	District-Level
<i>Constituent Data</i>		
Median Income	0.000	0.000
Achieved BA	−0.097	−0.154
Black	0.081	−0.020
Hispanic	0.152*	−0.002
Age 15 to 34	0.042	−0.089
Age 35 to 59	−0.869	0.117
Age 59 and up	0.917	0.023
Population	0.000	0.000
<i>Politician Data</i>		
Followers	0.000	0.000*
Freshman	0.022*	0.000
Age	−0.001	0.000
Gender (M=1)	0.020	0.013
Party (R=1)	0.003	−0.023*
N	136	386
<i>Note:</i> * p<0.1; ** p<0.05; *** p<0.01 <i>Source:</i> Twitter, ACS, Bioguide and NGA <i>OLS Regression</i>		

In Table 4, no constituent demographic variables influence constituency services tweets with the exception of the Hispanic population at a state level. On average, a 1% increase in the percentage of the state population will result in .152 more constituency service tweets made. In the grand scheme of tweeting, however, this increase is miniscule and thus negligible. Similar results occurred regarding politician demographics, While the number of followers that a politician has is statistically significant at a district level, the influence that the number of followers has on constituency service tweets is negligible. Interestingly, freshman status is

significant at the state level, but not the district level, a reversal in pattern from overall tweets. As with Hispanic at a state level and Followers at a district level, the impact of being a freshman at a state level is minimal. While the same holds true being a Republican, a negative sign was unexpected. Since Twitter has been tracked, Republicans have used Twitter more heavily (Glassman, Straus, and Shogan 2010), thus we expected a positive relationship at a minimum.

Table 5: Apolitical Tweets

	State-Level	District-Level
<i>Constituent Data</i>		
Median Income	0.000	0.000
Achieved BA	0.338*	0.025
Black	-0.030	0.034
Hispanic	-0.050	-0.022
Age 15 to 34	-0.232	-0.211
Age 35 to 59	-0.186	-0.288
Age 59 and up	-0.664	-0.081
Population	0.000	0.000**
<i>Politician Data</i>		
Followers	0.000	0.000
Freshman	-0.009	-0.010
Age	0.000	0.000
Gender (M=1)	-0.012	0.002
Party (R=1)	0.018	0.005
N	136	386
<i>Note:</i> *p<0.1; **p<0.05; ***p<0.01 <i>Source:</i> Twitter, ACS, Bioguide and NGA <i>OLS Regression</i>		

As with Tables 3 and 4, Table 5 shows little to no relationship between demographics and apolitical tweets. The only exception is constituents having achieved a Bachelor's degree at the state level, which is associated with a 0.338 increase in apolitical tweets. This relationship is

surprising given that people with higher levels of education tend to be more politically knowledgeable. As a result, we would expect politicians from areas with high education levels to tweet more about political issues than about apolitical topics. While there is a relationship between the population of a district and apolitical tweeting, the relationship is nearly zero, and thus insignificant to this study.

Conclusion

Studies regarding politicians' use of new media, particularly television and the Internet, have frequently noted the slow rate of technology adaptation in the creation of political tools (Ferber, Foltz, and Pugliese 2003, 2005). Twitter is no exception. Our study has proved not only that a majority of senators, governors, and representatives use Twitter, but also that these politicians fail to use Twitter as a political tool. Through an analysis of political, constituency service, and apolitical tweets made from February 2013 to April 2013, we found that politicians do not make tweets based on constituent demographics nor based on their own demographics. Age of the politician and freshman status proved to be the only two significant factors regarding the overall volume of tweets made. Taken together, our results suggest that there is no discernable pattern in how politicians create tweets.

The lack of pattern that we found may, in part, be attributable to the noisiness of Twitter, and the machine learning process. Tweets consist of small sequences of 140 characters or less, and the amount of actual content contained in any one is small, while the intended message may be large. This lends towards tweets being difficult to interpret even for humans, can make it very difficult to find underlying patterns in the noisy data. Our process of supervised machine learning adds to this difficulty, since this technique is subject to more uncertainty with these

small pieces of data than it would be with larger chunks. This lowers our confidence to a certain extent, and explains the added difficulty of determining patterns in the behavior of politicians.

While we found that politicians tweet at random, it may be that politicians use Twitter strategically during election periods. Politicians may tailor their tweets to meet the needs and demands of constituents when doing so matters most, especially in states and districts with highly competitive races. A variable measuring the competitiveness of electoral competition would need to be included in any study regarding Twitter use at election time. Oppositely, politicians may continue to underuse Twitter as a political tool. Further research into Twitter use during election times is needed to gain a better picture of politicians' relationships to Twitter overall.

Should Twitter continue to increase in popularity, we expect an increase in the strategic use of tweeting by politicians. Politicians from with a high percentage of Black people in their states and districts, for example, may take advantage of the fact that Black people use Twitter overwhelmingly more than other races and ethnicities. Knowing this, politicians may target their tweets towards constituency service in an attempt to gain favor. Readily available information such as demographics information and information regarding who uses Twitter should become more valuable in tandem with growth in the number of Twitter users. Twitter, while only offering 140 characters per message, has the potential to help a politician shape his or her public image in a strategic manner, something that we ultimately expect politicians to realize and utilize should the social media platform continue to grow in popularity.