SOCIAL NETWORKING WEBSITES AND VOTER TURNOUT

A Thesis
submitted to the Faculty of the
Graduate School of Arts and Sciences
of Georgetown University
in partial fulfillment of the requirements for the
degree of
Master of Public Policy

By

Bryan Boroughs, B.S.

Washington, DC
April 14, 2010
SOCIAL NETWORKING WEBSITES AND VOTER TURNOUT

Bryan Boroughs, B.S.

Thesis Advisor: John T Christian, Ph.D.

ABSTRACT

Social networking websites (SNWs) like Facebook and Twitter have played an increasing role in American politics. The social networking phenomenon has been discussed heavily in the press, but little scholarship has examined the direct effects, if any, of SNWs on voter turnout. This paper investigated whether accessing political content through SNWs affected the likelihood that individuals would vote in the 2008 presidential election. To examine this question, the paper used the Pew Center dataset titled November 2008 Post-Election Tracking Survey. This random-digit telephone survey dataset examined both voting behavior during the presidential election and online activity prior to the election. Using a probit regression model, this paper determined that a very significant relationship existed between accessing political content on SNWs and the likelihood of voting. The findings estimated that voter turnout increased by 1 vote for every 11 to 25 people who accessed political content on SNWs, depending on the specification used. The practical implications of this finding for political campaigns are powerful.
I would like to thank the Georgetown Public Policy faculty, staff, and students.

Special thanks to my thesis advisor, Dr. John T. Christian, and the fellow students who provided feedback and counsel: Brian Bartlett, Emily Becker, Marshal Erwin, Steven Grimes, Rory O’Sullivan, and Kristen Silverberg.

My greatest thanks goes to Adair.
# Table of Contents

Introduction .......................................................................................................................... 1  
Literature Review ............................................................................................................... 3  
Conceptual Model ............................................................................................................. 8  
Data and Methods ............................................................................................................. 9  
  Dataset Details .................................................................................................................. 9  
  Methods ............................................................................................................................ 11  
  Summary Statistics ............................................................................................................ 15  
Findings ............................................................................................................................... 16  
  Estimates of Predictive Power ......................................................................................... 16  
  Substantive Significance Of Accessing Political Content On SNWs ....................... 18  
  Significant Control Variables ......................................................................................... 21  
Conclusion and Next Steps ............................................................................................... 23  
Appendix: Data Transformation Detail ............................................................................... 26  
  Dependent Variable ........................................................................................................ 26  
  Independent Variable of Interest .................................................................................. 26  
  Control Variables ........................................................................................................... 27  
  Basic Demographics ...................................................................................................... 28  
References .......................................................................................................................... 29
LIST OF TABLES

Table 1: Summary Statistics for Indicator Variables........................................ 15
Table 2: Summary Statistics For the Age and TrustsGovernment Variables ...... 15
Table 3: Probit Z-Coefficients........................................................................... 17
Table 4: Estimated Effects.................................................................................. 19
Introduction

Social Networking Websites (SNWs) have allowed users to communicate with their individual contacts regarding shared interests, activities, or events. Recent political campaigns and lobbying groups have extensively used SNWs like Facebook (2010) and Twitter (2010). Indeed, during the 2008 presidential election, major media outlets frequently discussed President Obama’s leverage of SNWs, compared to Senator McCain’s (Wilburforce, 2009). However, existing literature does not empirically show how much, or if, SNWs impacted voter turnout. This paper found that accessing political content on SNWs had a statistically and substantively significant impact on individuals’ likelihood of voting in the 2008 presidential election.

SNWs likely affect a given person’s likelihood of voting in two ways. First, simply viewing political content on SNWs could directly persuade a voter to go to the polls. Second, political content could help organize and inform activists or volunteers, making them more effective at motivating other potential voters and altering the political discourse. While these two features of SNWs are related, this paper examined the effect of the first.

This paper used the November 2008 Post-Election Tracking Survey, a dataset from the Pew Center, for a quantitative analysis. The Pew Center created this dataset by commissioning a random digit dial telephone survey, and it released the data to the
public domain in 2009. The dataset examined voters’ Internet and media uses in relation to the 2008 presidential election.

This paper used 3 probit regression specifications to estimate what effect accessing political content on SNWs had on individuals’ likelihood of voting. The coefficient for that effect was significant (p < 0.005) for each of the three specifications. The effect was also substantively significant, suggesting that one extra vote was elicited for every 11 to 25 people who accessed political content on SNWs. The practical implications of that result are potentially powerful, particularly given the minimal expense of distributing political content through SNWs.
Literature Review

In 2008, nearly 40% of adults in the United States communicated with others about politics using the Internet: 11% posted political content on a blog, 20% shared their political views on Twitter (2010), and 33% posted political content on social networking websites (SNWs) such as Facebook (2010) and LinkedIn (2008) (Pew Center, 2008). Scholarly research examining this field has focused on the impact of these technologies on the political discourse. The literature debates whether they improve informed public discourse, or if they Balkanize the electorate. The evidence suggests both have happened to some degree, but the implications for voter turnout are unclear.

In his book, Republic.com, Cass Sunstein (2001) argued that the Internet threatened the very fabric of our democracy because it could Balkanize political discourse. He asserted that citizens would soon be able to filter which news they learned about and, with this ability, would only visit websites authored by like-minded people. (Sunstein, 2001). In this theorized scenario, political discourse could devolve into a thousand tiny echo chambers – where people hear only what they want to hear and extremism flourishes.

Other scholars emphasized that New Media could create a virtual forum for discussion, debate, and dissemination of previously unavailable information (Robbin,
2008). Some argued that the Internet improves discourse by allowing citizens to learn about issues beyond those covered in most 30-minute television news broadcasts (Lievrouw, 2001). Indeed, an optimist could imagine the Internet fostering a transcendent national debate, in place of the television-dominated news cycle of the 1990s.

The data suggest that neither theoretical framework has proven complete. Certainly, some of Sunstein’s predictions have come true. People increasingly look to the Internet for political content. In 2008, 44% of adults in America used the Internet to learn about or discuss politics, while only 18% of adults surveyed claimed to have done so in 2000. Twenty-six percent now say that they get most of their political news online, up from 10% in 2000 (Pew Center, 2008).

However, Sunstein also predicted that people would only seek news from sources that agreed with their personal perspectives. That prediction has only proven partly true. In 2008, 33% of adult online political users reported that they got their news from web sites that shared their point of view. However, 21% said that they mostly visited sites that challenged their point of view (Pew Center, 2008). These numbers certainly contain some bias because they are based on self-reported data. Nonetheless, the data do not seem to show that the American population only reads agreeable news, contrary to Sunstein’s fears.
Furthermore, it seems unlikely that one person’s self-filtered Internet feed would fully define the information that person sees on a daily basis. Even if people predominately get news and analysis from a self-filtered web, they will still have family, coworkers, neighbors, and friends who disagree about many issues. And, for all the filtering modern SNWs and blogs can provide, they also provide remarkably easy access to the opinions of those with whom we might disagree.

Some evidence also supports the idea that New Media has improved informed discussion in the country. A 2004 study by the Institute for Politics Democracy & The Internet examined the relationship between political bloggers, which they termed Online Political Citizens (OPCs), and the population at-large. The study showed that OPCs were disproportionately likely to be influential in their non-Internet communities. In fact, OPCs were seven times more likely to be influential persons than the general population. Rather than being isolated from their communities, they were more likely to be leaders of civic organizations, attend public meetings about local affairs, and attend political rallies (IPDI, 2004).

Still, there are plenty of anecdotal reasons to think that the SNWs and blogs have not necessarily raised the quality of the national debate. Many extremists have been given a New Media megaphone; some garnered a great deal of attention by
claiming that President Bush planned the September 11th bombings, President Obama supports infanticide, and the trip to the moon was an elaborate hoax.

Agre (2002) offered a model that synthesizes these competing theories. He suggested that we should think of the Internet and New Media in the framework of an amplification model. He argued that New Media simply provides numerous new options for institutions and individuals to pursue existing political goals and strategies. These technologies make it inexpensive for anyone to express ideas to a massive audience, coordinate activities across great distances, and respond rapidly to developments in the political sphere (Agre 2002).

In that sense, the Internet and New Media are neither doomsday devices nor silver bullets for political discourse. Instead, Agre argues that they make the political discourse louder, faster, and – yes – more chaotic (2002).

However, neither Agre's model nor Sunstein's predictions suggests what, if any, impact these SNWs have on voter turnout. One recent experimental study noted that direct unsolicited e-mail communication did not increase voter turnout (Nickerson 2007). But social media, unlike email, requires affirmative action by the consumer, and no scholarly research has examined whether social media increases or decreases motivation to actually vote, which may be the most important question for those
interested in using social media to advance any number of policy agendas. This paper examined that question.
Conceptual Model

Under Agre’s amplification model, New Media technologies encourage louder and faster political discourse. However, his model does not expressly establish a theoretical framework that would explain whether exposure to New Media would increase or decrease the likelihood of voting.

As acknowledged by both Sunstein and Agre, these New Media technologies are more active, in that the consumer has to take affirmative action to view the content. Whereas users passively view other traditional political marketing tools like television ads, radio, or direct mail. Social networking websites (SNWs) also differ from traditional media in that users often receive content referred by friends or associates, which could lend credibility to the content. As a result of the SNWs active and social qualities, those technologies may have capacity for greater effect than traditional media.

This empirical study examined whether that hypothesis is accurate. It asked whether accessing political content on SNWs increased the likelihood that individuals actually invested the time and effort to vote in 2008. By controlling for demographics that are common to users of New Media, as well as characteristics that impact voter turnout generally, the study isolated the direct effect of accessing SNWs’ political content on voting behavior.
Data and Methods

This analysis was based on the November 2008 Post-Election Tracking Survey dataset published by the Pew Center. The Pew Center released the data approximately 6 months after publishing its own initial reports.

Dataset Details

The Pew Center’s dataset contained basic demographic and voting behavior data, as well as 117 questions focusing largely on how people used various media to gather information about the election. It also asked whether the participants used social networking websites (SNWs) and whether they used those sites for political content.

The Pew Center’s dataset included 2,254 total responses. The margin of error for the total sample was 2%. The dataset contained a subset of 1,591 respondents who are categorized as Internet users, and 1,186 who were categorized as online political users. For both subsets, the margin of error was 3%. This paper used the entire sample (the model, in part, compared the online political users to everyone else).

Sampling Methods

The Pew Center hired Princeton Survey Research Associates International to conduct the surveys between November 20 and December 4, 2008. The sample was created by random digit dialing of United States telephone numbers. At least 10 attempts were made before giving up on any given household’s phone number, over a
mix of times and day of week. Each sampled number was contacted at least once during daytime hours.

For each contacted household, the interviewer used a method the Pew Center claims has been shown to create a sample that closely mirrors the population in terms of age and gender. Specifically, the interviewer first asks for the youngest male currently at home. If none is at home, then the interviewer asks for the youngest female available.

**Success/Response Rates**

The random dialer attempted to contact 26,690 total numbers. Of those, 9,704 were successfully contacted. The remainder were non-working numbers, businesses, fax numbers, etc. Of those contacted, 6,653 declined to participate, 457 more were excluded due to language barriers, and 340 surveys were interrupted and could not be completed.

The remaining 2,254 respondents completed the survey and were included in the dataset (a response rate of 23.3% for working numbers).

**Dataset Limitations**

This dataset did not have pre- and post-election data. Longitudinal data, that gauged respondents’ interest in the election over time, would be more informative. The dataset can speak somewhat to general interest in voting because it surveyed both
whether the respondents trust the government and whether the respondents were satisfied with the direction of the country.

However, prior voting behavior, per se, was only surveyed for people who voted in 2008. The survey’s question asking for prior voting behavior read: “[w]as this year the first time you voted in an election, or have you voted in elections in past years?” The lack of data on prior voting behavior for those who did not vote in 2008 is a substantial limitation. American National Election Studies Data suggests that people who voted in 2004’s presidential election were 61 percentage points more likely to vote in 2008 than those who did not vote in 2004 (ANES 2008). The current study addressed this limitation using two different methods for imputing missing prior voting behavior data, discussed in further detail below.

**Methods**

This paper used three probit regression specifications that were identical except for how they deal with prior voting behavior. As noted above, data for prior voting behavior was only available for those who did vote in the 2008 election. As a result, the three different specifications addressed that dataset limitation from different angles.

The dependent variable in each model was whether the respondent voted in 2008’s elections, and the independent variable of interest was whether the person accessed political content on social networking websites (SNWs). Some control
variables were discussed briefly in this section, and all variables were discussed more fully in Appendix A. STATA 11 was used to perform the statistical analysis.

Specification 1 did not account for prior voting behavior. Specifications 2 and 3 included prior voting behavior by imputing missing values. Imputing missing data was necessary because, otherwise, all respondents who did not vote in 2008 would have missing data and would be ignored by the regression; a regression on voting probability in 2008 then becomes impossible because the data would only include those who did vote.

No form of imputation is perfect, so two forms were used for comparison purposes. Specification 2 used a 2-stage regression method that tends to over-estimate the magnitude of coefficients and predictive power, while specification 3 used a random assignment imputation that tends to under-estimate predictive power and coefficients.

**Specification 1**

This specification did not account for prior voting behavior:

\[
P(Vote = 1) = \Phi(\beta_0 + \beta_1 UsesSNWsPolitically + \beta_2 UsesSNWs \\
+ \beta_3 UsesInternet + \beta_4 UsesInternetDaily + \beta_5 GetsMostNewsFromNet \\
+ \beta_6 Age + \beta_7 Male + \beta_8 BlackNotHisp + \beta_9 Hisp + \beta_{10} RaceOther \\
+ \beta_{11} TrustsGovernment + \beta_{12} SatisfiedWithDirection + \varepsilon)
\]
**Specification 2**

This specification included all the same control variables as specification 1. It added a control variable for prior voting behavior ($PriorVoteImputedWithProbit$). To impute the missing prior voting data, a probit model was first run using all the same independent variables. That first-stage probit regression was then used to predict the probability of prior voting for those respondents with missing data.

However, this method of imputation is only valid if you can assume that the functional form is consistent between those who did and those who did not vote in prior years. If that is not the case, the models predictive power and coefficients will likely be inflated. The resulting specification was as follows:

\[
P(Vote = 1) = \Phi(\beta_0 + \beta_1 UsesSNWspolitically + \beta_2 UsesSNWs
+ \beta_3 UsesInternet + \beta_4 UsesInternetDaily + \beta_5 GetsMostNewsFromNet
+ \beta_6 Age + \beta_7 Male + \beta_8 BlackNotHisp + \beta_9 Hisp + \beta_{10} RaceOther
+ \beta_{11} TrustsGovernment + \beta_{12} SatisfiedWithDirection
+ \beta_{13} PriorVoteImputedWithProbit + \epsilon)
\]

**Specification 3**

Specification 3 also included all the same control variables as specification 1, but it used a random imputation method to account for missing prior voting behavior data. The ANES study showed that, of those who did not vote in 2008, 16.75% did vote in 2004. Accordingly, 16.75% of the respondents who did not vote in 2008 were
randomly assigned to the group of people who did vote in prior elections, and 85.25% were randomly assigned to the group who did not vote in prior elections.

In contrast to the imputation used in specification 2, random assignment washes out the effect of the imputed variable. Thus, it tends to underestimate the magnitude of both coefficients and explanatory power.

$$P(Vote = 1) = \Phi(\beta_0 + \beta_1 UsesSNWsPolitically + \beta_2 UsesSNWs$$
$$+ \beta_3 UsesInternet + \beta_4 UsesInternetDaily + \beta_5 GetsMostNewsFromNet$$
$$+ \beta_6 Age + \beta_7 Male + \beta_8 BlackNotHisp + \beta_9 Hisp + \beta_{10} RaceOther$$
$$+ \beta_{11} TrustsGovernment + \beta_{12} SatisfiedWithDirection$$
$$+ \beta_{13} PriorVote ImputedRandomly + \epsilon)$$

**Variables of Interest**

The dependent variable, *Vote*, was based on a simple question regarding whether the respondent voted in the 2008 elections. The independent variable of interest was: *UsesSNWsPolitically*, an indicator variable for whether the respondent accessed political content through SNWs.

Several control variables, described in the appendix, included whether the respondent used the Internet generally, and whether the respondent was satisfied with the direction of the country. Demographic control variables included age, gender, and race.
Summary Statistics

The vast majority of the variables used were indicator variables, summarized in Table 1, below. *Age* and *TrustsGovernment* were continuous variables, summarized in Table 2, below.

**Table 1: Summary Statistics for Indicator Variables**

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>TOTAL OBSERVATIONS</th>
<th>PERCENTAGE 'YES'</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voted in 2008</td>
<td>2252</td>
<td>85.17%</td>
</tr>
<tr>
<td>Uses SNWs Politically</td>
<td>2254</td>
<td>10.25%</td>
</tr>
<tr>
<td>Uses SNWs Generally</td>
<td>2254</td>
<td>20.72%</td>
</tr>
<tr>
<td>Uses Internet Generally</td>
<td>2254</td>
<td>70.59%</td>
</tr>
<tr>
<td>Used Internet Yesterday</td>
<td>2250</td>
<td>50.22%</td>
</tr>
<tr>
<td>Gets Most News From Internet</td>
<td>2254</td>
<td>23.03%</td>
</tr>
<tr>
<td>Male</td>
<td>2254</td>
<td>47.16%</td>
</tr>
<tr>
<td>Black Not Hispanic</td>
<td>2212</td>
<td>9.45%</td>
</tr>
<tr>
<td>Hispanic</td>
<td>2212</td>
<td>6.19%</td>
</tr>
<tr>
<td>Race Other</td>
<td>2212</td>
<td>3.39%</td>
</tr>
<tr>
<td>Satisfied With Direction Of Country</td>
<td>2254</td>
<td>15.71%</td>
</tr>
<tr>
<td>Prior Vote Imputed With Probit</td>
<td>2219</td>
<td>81.00%</td>
</tr>
<tr>
<td>Prior Vote Imputed Randomly</td>
<td>2254</td>
<td>80.30%</td>
</tr>
</tbody>
</table>

**Table 2: Summary Statistics For the Age and TrustsGovernment Variables**

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>TOTAL OBS</th>
<th>AVERAGE</th>
<th>STD. DEV</th>
<th>MIN</th>
<th>MAX</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>2254</td>
<td>55.19077</td>
<td>18.93353</td>
<td>18</td>
<td>99</td>
</tr>
<tr>
<td>Trusts Government</td>
<td>2146</td>
<td>2.884436</td>
<td>6.187965</td>
<td>1</td>
<td>4</td>
</tr>
</tbody>
</table>
Findings

Accessing political content from social networking websites (SNWs), like Facebook (2010) or Twitter (2010), increased individuals’ likelihood of voting in the 2008 presidential election. The effect was statistically significant across each of the 3 specifications (p<0.005).

The magnitude of the effect was strong across each of the specifications examined. Accessing political content from SNWs (UsesSNWsPolitically) showed a z-coefficient ranging from 2.93 to 2.98. This suggests that voter turnout increases by 1 vote for every 11 to 25 people who view political content on SNWs. Regarding specific subpopulations (20 year olds, people who voted before, etc), voter turnout increases by 1 vote for every 6 to 112 persons reached.

Estimates of Predictive Power

The pseudo R-squared for the three models ranged from 0.11 and 0.56. Specification 1, which did not account for prior voting behavior, showed the lowest pseudo R-squared and only explains about 11% of the variations in voting behavior. When the missing data for prior voting behavior is imputed using a separate probit regression (Specification 2), the model explains approximately 56% of the variations in voting behavior. When missing prior voting behavior is randomly assigned, the model explains approximately 45% of the variations in voting behavior.
Table 3: Probit Z-Coefficients

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>SPEC 1</th>
<th>SPEC 2</th>
<th>SPEC 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pseudo R-squared</td>
<td>.11</td>
<td>.56</td>
<td>.45</td>
</tr>
<tr>
<td>Uses SNWs Politically</td>
<td>2.93 (0.003)</td>
<td>2.98 (0.003)</td>
<td>2.94 (0.003)</td>
</tr>
<tr>
<td>Uses SNWs Generally</td>
<td>-2.21 (0.027)</td>
<td>-0.43 (0.670)</td>
<td>-1.21 (0.225)</td>
</tr>
<tr>
<td>Uses Internet</td>
<td>6.53 (0.000)</td>
<td>3.20 (0.001)</td>
<td>3.42 (0.001)</td>
</tr>
<tr>
<td>Used Internet Yesterday</td>
<td>2.20 (0.027)</td>
<td>0.88 (0.380)</td>
<td>0.79 (0.432)</td>
</tr>
<tr>
<td>Gets Most News From Internet</td>
<td>2.27 (0.023)</td>
<td>2.17 (0.030)</td>
<td>2.11 (0.035)</td>
</tr>
<tr>
<td>Age</td>
<td>8.67 (0.000)</td>
<td>2.82 (0.005)</td>
<td>1.75 (0.081)</td>
</tr>
<tr>
<td>Male</td>
<td>-1.25 (0.219)</td>
<td>-0.56 (0.574)</td>
<td>-0.44 (0.659)</td>
</tr>
<tr>
<td>Black Not Hispanic</td>
<td>3.81 (0.000)</td>
<td>3.75 (0.000)</td>
<td>4.38 (0.000)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>-0.88 (0.380)</td>
<td>-0.96 (0.338)</td>
<td>-0.68 (0.499)</td>
</tr>
<tr>
<td>Race Other</td>
<td>-3.01 (0.003)</td>
<td>-2.12 (0.034)</td>
<td>-1.90 (0.057)</td>
</tr>
<tr>
<td>Trusts Government</td>
<td>-1.50 (0.135)</td>
<td>-0.74 (0.460)</td>
<td>-1.93 (0.054)</td>
</tr>
<tr>
<td>Satisfied With Direction Of Country</td>
<td>-1.37 (0.172)</td>
<td>-2.34 (0.019)</td>
<td>-2.70 (0.007)</td>
</tr>
<tr>
<td>Prior Vote Imputed With Probit</td>
<td>*</td>
<td>21.43 (0.000)</td>
<td>*</td>
</tr>
<tr>
<td>Prior Vote Imputed Randomly</td>
<td>*</td>
<td>*</td>
<td>21.56 (0.000)</td>
</tr>
</tbody>
</table>
Substantive Significance Of Accessing Political Content On SNWs

Probit regressions provide coefficients in the form of Z-scores.

UsesSNWsPolitically has a z-coefficient between 2.93 and 2.98 across each specification. This means that, for any given person, visiting a social networking site will move that person’s likelihood of voting by between 2.93 and 2.98 standard deviations from the mean.

This coefficient is most meaningful when placed in context of a specific set of characteristics. In Table 4 below, this coefficient is used to estimate how much UsesSNWsPolitically will impact the likelihood of voting by an average person. To do so, average values were first applied to all other variables (age, demographics, etc). Then, UsesSNWsPolitically was changed from true to false, and the change in the estimated likelihood of voting is measured.

The same process estimates the impact of UsesSNWsPolitically on various subpopulations including those who voted in prior years, men, women, various age groups, and various races.

This process yields estimates that are statistically significant, in that the likelihood that this effect would occur by chance is less than .5%. However, this paper does not analyze differences between groups for statistical significance. Thus, the differences between subgroups (males/females, blacks/whites, young/old, etc) are not necessarily statistically significant.
Table 4: Estimated Effects

<table>
<thead>
<tr>
<th>CHANGE IN THE PROBABILITY OF VOTING BASED ON USESSNWSPOLITICALLY FOR:</th>
<th>SPEC 1</th>
<th>SPEC 2</th>
<th>SPEC 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>An ‘Average’ Respondent</td>
<td>+0.091</td>
<td>+0.039</td>
<td>+0.071</td>
</tr>
<tr>
<td><strong>Prior Voting Behavior</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Respondents Who Had Not Voted Before</td>
<td>n/a</td>
<td>+0.220</td>
<td>+0.217</td>
</tr>
<tr>
<td>Respondents Who Voted In Prior Years</td>
<td>n/a</td>
<td>+0.011</td>
<td>+0.036</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female Respondents</td>
<td>+0.086</td>
<td>+0.037</td>
<td>+0.069</td>
</tr>
<tr>
<td>Male Respondents</td>
<td>+0.096</td>
<td>+0.042</td>
<td>+0.073</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20 year olds</td>
<td>+0.165</td>
<td>+0.071</td>
<td>+0.093</td>
</tr>
<tr>
<td>30 year olds</td>
<td>+0.146</td>
<td>+0.060</td>
<td>+0.086</td>
</tr>
<tr>
<td>40 year olds</td>
<td>+0.124</td>
<td>+0.051</td>
<td>+0.079</td>
</tr>
<tr>
<td>50 year olds</td>
<td>+0.100</td>
<td>+0.042</td>
<td>+0.073</td>
</tr>
<tr>
<td>60 year olds</td>
<td>+0.077</td>
<td>+0.035</td>
<td>+0.067</td>
</tr>
<tr>
<td>70 year olds</td>
<td>+0.056</td>
<td>+0.029</td>
<td>+0.061</td>
</tr>
<tr>
<td>80 year olds</td>
<td>+0.039</td>
<td>+0.023</td>
<td>+0.056</td>
</tr>
<tr>
<td><strong>Race/Ethnicity</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Whites (non-Hispanic)</td>
<td>+0.094</td>
<td>+0.043</td>
<td>+0.077</td>
</tr>
<tr>
<td>Hispanics</td>
<td>+0.107</td>
<td>+0.059</td>
<td>+0.091</td>
</tr>
<tr>
<td>Blacks (non-Hispanic)</td>
<td>+0.042</td>
<td>+0.009</td>
<td>+0.017</td>
</tr>
</tbody>
</table>

**Estimated Effect On An ‘Average’ Person**

In the spirit of full disclosure: the ‘average’ person does not exist. This estimation first takes the average values of all other variables, resulting in a person
who is 57 years old, 9% Black, and 48% male, among other characteristics. Nonetheless, this is a helpful test in that it shows the expected effect of $\text{UsesSNWsPolitically}$ on a randomly selected person or group of persons.

Based on the results in Table 4 above, viewing political content on SNWs increased the average person’s likelihood of voting by between 9 and 4 percentage points. To put another way, for every 11 to 25 randomly selected persons who accessed political content on SNWs, voter turnout increased by 1 vote.

**Estimated Effect On Those Who Had Voted in Prior Elections**

Accessing political content had a much larger impact on those who had not voted before than on those who had. Viewing political content on SNWs increased the probability of voting by approximately 22 percentage points for those who had not voted in prior years. In contrast, the probability only increased by between 1 and 4 percentage points for those who had voted in prior elections.

**Estimated Effect For Different Genders**

The effect of accessing political content on SNWs was larger for male voters, but the difference between genders was small. Between the three specifications, the increased likelihood of voting for each group ranged from between 3.7 and 9.6 percentage points, and male and female voters were within 1 percentage point of each other in each specification.
Estimated Effect For Different Ages

Younger voters’ turnout was more strongly affected by accessing political content on SNWs. Twenty-year olds were 9.3 to 16.5 percentage points more likely to vote if they had accessed political content on SNWs, while 80 year olds have an estimated impact of between 3.9 and 5.6 percentage points.

Estimated Effect For Different Races

The effect of political content from SNWs on voter turnout varied substantially based on the respondents’ race. Accessing political content on SNWs increased the probability that blacks would vote by only 1.7 to 4.2 percentage points, whereas doing so increased the probability that whites would vote by 4.3 to 7.7 percentage points. Accessing political content on SNWs had an even larger effect on Hispanics, increasing the probability that they would vote by 5.9 to 10.7 percentage points.

Significant Control Variables

The most significant, and unsurprising, predictor for 2008 voting behavior was prior voting behavior (p<0.001). Indeed, including that variable increases the model’s explanatory power from 11% to 45% or 56%. The next most significant variable, across all specifications, was whether or not the respondent was Black; Blacks were substantially more likely to vote in the 2008 election than whites, with a z-score between 3.75 and 4.38 (p < 0.001).
Other significant predictors of voting behavior from the control variables included whether or not the respondent 1) uses the internet (p < 0.005 across all specifications); 2) gets most of their news from the internet (p < 0.05 across all specifications); and 3) is satisfied with the direction of the country (p < 0.05 across specifications 2 and 3). Finally, age had a statistically significant and positive relationship with voting behavior across specifications 2 and 3 (p < 0.005).
Conclusion and Next Steps

This paper examined whether accessing political content on social networking websites (SNWs) increases the likelihood of voting. The findings suggest that it does. They also estimate that the effect is particularly large for those who have not voted before, suggesting that SNWs may be potent tools for Get-Out-The-Vote efforts and other efforts to reach first-time voters.

These findings lend statistical credence to the intuitions of political campaigns that have been using SNWs as a valued communications tool, particularly for young people. This study may not generalize directly to non-presidential campaigns, but the strength of the effect suggests that SNWs may be a persuasive tool for advancing other policy agendas as well.

The practical implications of these findings are vast. Current political campaigns make a host of strategic decisions based on how much it costs to get any single vote by using various media. Since SNWs are so new, that data on how much it costs to gain a vote via social media is simply not widely available. However, this study provides an initial framework for estimating how many people must be reached to increase voter turnout generally and to increase voter turnout for various subpopulations.
Better Accounting for Prior Voting Behavior

The current study imputes missing data for prior voting behavior in order to include that control variable in the analysis. Future research could yield more reliable estimates if prior voting behavior is fully recorded by future datasets.

Future Research Questions

This study’s findings suggest a host of future research questions. First and foremost, future studies could examine whether accessing political content on SNWs increases the likelihood that a voter will vote for a particular candidate. In other words, do SNWs simply motivate people to vote for the candidate they already preferred, or do they inform voters’ decisions about which candidate to support? Also, it would be interesting to learn more about the relationship between voting behavior and the regularity or intensity with which voters consume political content on SNWs.

One other potential area of inquiry would examine the different vectors for political content on SNWs. SNWs make it easy for users to seek out content about a particular candidate, and they make it easy to refer that content to friends and contacts. Future studies could examine whether political content is most effective if it is sought out directly, or if the more powerful effect comes from having that content referred to the voter by an associate.
Compensating for Non-Response Bias

The Pew Center’s study, *The Internet’s Role in Campaign 2008*, compensated for non-response bias by weighting the responses during analysis. They asserted that this weighting was useful because certain population subgroups systematically underrespond to telephone surveys, and that those same subgroups could vary with many other variables (Pew Center, 2008).

They used the Census Bureau’s March 2007 Annual Social and Economic Supplement to estimate population demographics (for those who own a telephone). Those population demographics were then used in the analysis of the sample to give extra weight to the underreported subgroups of the population (Pew Center, 2008). Ideally, future research will achieve more accurate estimates by including a similar weighting to its analysis.
Appendix: Data Transformation Detail

The italicized words below represent variables from the transformed dataset. Fields noted parenthetically are from the original dataset from the Pew Center.

Dependent Variable
Vote = 1 if the respondent answered yes to whether the respondent voted. The survey’s question was worded as follows: “A lot of people have been telling us they didn't get a chance to vote in the elections this year on November 4. How about you... did things come up that kept you from voting, or did you happen to vote?” (vote = 1)

Independent Variable of Interest
UsesSNWsPolitically = 1 if answered ‘yes’, that they have ever: (any of the below)

- Use Twitter to post your thoughts or experiences related to the campaign (q19h <= 5)
- “Posted comments, queries or information about the campaign or the elections . . . On a social networking site such as Facebook, MySpace or LinkedIn.” (q26c = 1)
- “Gotten any campaign or candidate information on the sites” (q27a = 1)
- “Started or joined a political group, or group supporting a cause on a social networking site” (q27b = 1)
- Revealed on a social networking site which presidential candidate you voted for this year. (q27c = 1)
• Discovered on the sites which presidential candidate your friends voted for this year. (q27d = 1)

• Signed up as a “friend” of any candidates on a social networking site. (q27e = 1)

• After you voted in this fall’s elections, did you share with others your experiences casting your ballot by . . . Posting your experiences on a social networking site. . . [or] . . . posting updates to Twitter or another status update service” (vote11e or vote11f = 1)

**Control Variables**

*UsesSNS* = 1 if answered that they ever “Use a social networking site like MySpace, Facebook or LinkedIn.com” or “Use Twitter or another service to share updates about yourself or to see updates about others” (activ87 <= 2, or activ112 <= 2)

*UsesInternet* = 1 if answered yes to: “Do you use the internet, at least occasionally?” (i_user = 1)

*UsedInternetYesterday* = 1 if answered yes to the question “did you happen to use the Internet yesterday?” (q7 = 1)

*GetsMostNewsFromNet* = 1 if answered ‘Internet” to the question: “How have you been getting most of your news about the November elections . . . from television, from newspapers, from radio, from magazines, or from the Internet?” This question was asked twice. This variable is set to 1 if either time resulted in “internet” (q15_1 or q15_2 = 1)
TrustsGovernment = 1 if answered yes to the question “How much of the time do you think you can trust the government in Washington to do what is right… just about always, most of the time, or only some of the time?” (q2)

SatisfiedWithDirectionOfCountry = 1 if answered yes to the question “Overall, are you satisfied or dissatisfied with the way things are going in this country today?” (q1)

FirstTimeVoting = 1 if answered yes to the question: “Was this year the first time you voted in an election, or have you voted in elections in past years?” (vote05 = 1)

PriorVoteImputedWithProbit = 1 if FirstTimeVoting = 0, but missing values were imputed using a probit regression

PriorVoteImputedRandomly = 1 if FirstTimeVoting = 0, but missing values were randomly assigned, based on the NES estimate that non-voters and a 17% probability of having voted in prior elections.

**Basic Demographics**

Age (age), Male (sex = 1), BlackNotHisp (racethn = 2), Hisp (racethn = 3),

RaceOther (racethn = 4)
References


