Title
Essays in American Political Behavior

Author
Bond, Robert M.

Publication Date
2013

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Essays in American Political Behavior

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Political Science

by

Robert Bond

Committee in charge:

Professor James Fowler, Chair
Professor Charles Elkan
Professor David Huber
Professor Thad Kousser
Professor Gary Jacobson

2013

The dissertation of Robert Bond is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego

2013

DEDICATION

My grandparents – Harry Bycroft, Betty Bycroft, Ronald Bond, and Lucy Stockton – did not live to see the completion of this dissertation. It is dedicated to their lives and their memory.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

Many of my colleagues, friends and family have contributed to the creation of this dissertation. My advisors – James Fowler, Charles Elkan, David Huber, Gary Jacobson and Thad Kousser – have been generous with their time and advice since the beginning of this project. Many other members of the UCSD faculty have contributed to this project, especially Sam Popkin, Zoltan Hajnal, Scott Desposato, Keith Poole and Phil Roeder. My fellow graduate students have been a consistent source of advice, especially Chris Fariss, Micah Gell-Redman, Jeremy Horowitz, Alex Hughes, Jason Jones, Lindsay Nielson, Michael Rivera, Paul Schuler, Jaime Settle, Devesh Tiwari and Neil Visalvanich.

Previous versions of several chapters have benefited from the suggestions of participants in the 2011 Midwest Political Science Association Annual National Conference, 2011 Political Networks Conference 2011 American Political Science Association Annual National Conference, 2012 Political Methodology Summer Meeting, 2012 Midwest Political Science Association Annual National Conference and the 2012 Political Networks Conference.

This dissertation would not have been possible without the constant support of my parents, Margaret and Ryan, my sister, Micaella, and the love of my life, Elizabeth.

VITA

2006	B.A. in Political Science, Arizona State University
2007	M.A. in Political Science, Arizona State University
2013	Ph.D. in Political Science, University of California-San Diego

PUBLICATIONS

"The Social Origins of Adult Political Behavior," *American Politics Research* (2011) (with Jaime Settle and Justin Levitt).

"A 61-Million-Person Experiment in Social Influence and Political Mobilization," *Nature* (2012) (with Christopher J. Fariss, Jason J. Jones, Adam D. I. Kramer, Cameron Marlow, Jaime Settle and James H. Fowler).

"Inferring Tie Strength from Online Directed Behavior," *PLoS One* (2013) (with Jason J. Jones, Jaime Settle, Christopher J. Fariss, Cameron Marlow and James H. Fowler).

"Yahtzee: An Anonymized Group Level Matching Procedure," *PLoS One* (2013) (with Jason J. Jones, Christopher J. Fariss, Jaime Settle, Adam D.I. Kramer, Cameron Marlow and James H. Fowler).

ABSTRACT OF THE DISSERTATION

Essays in American Political Behavior

by

Robert Bond

Doctor of Philosophy in Political Science

University of California, San Diego, 2013

Professor James Fowler, Chair

How does the social environment in which people are embedded impact their political behavior and attitudes? This dissertation provides substantive and methodological advances in answering this key question in political science research. Chapter 1 analyzes a get-out-the-vote field experiment involving more than 61 million individuals. The results show that the messages influenced political self-expression, information seeking, and real world voting behavior of millions of people. The effect of social information versus non-social information differed by characteristics of the treated individual such as age, education, relationship status, and the number of social contacts the individual has. These results suggest that while social information increases participation for overall, it is especially effective for subsets of the population. Chapter 2 analyzes the effect of one individual's turnout on that of her social contacts. Results indicate that

when a friend votes an individual is about 7% more likely to vote. Chapter 3 develops a statistical model to estimate the ideology of politicians and their supporters using Facebook data about which users publicly support which political figures by ‘liking’ them on the site. Then, using this measure, I study the topography of ideology across a social network of more than 6 million people, and show that those individuals who are embedded in diverse ideological networks are less likely to turnout to vote than those in homogeneous social networks.

Chapter 1

Social Information and Participation

1.1 Introduction

A growing literature suggests that voting has a substantial social component and that decisions to turnout are impacted by our knowledge of the voting behavior of members of our social networks. We know that turnout is highly correlated between friends and family (Berelson, Lazarsfeld and McPhee 1954; Huckfeldt and Sprague 1995). However, this correlation may be due to many factors. For instance, families may socialize voting behavior, people may observe and imitate the voting behavior of their social contacts, or people may become friends with people who have similar attitudes toward voting. All of these reasons, however, owe to the existence of a social norm that participating in elections is positive. Without a social norm concerning voting, social interactions would not affect our decisions to vote.

Here we employ new data and methods from the growing field of computational social science (Lazer et al. 2009) to study how a GOTV message affects subpopulations differentially. Social science, and political science in particular, are seeing the benefits of using large-scale data sources to study phenomena in new ways. For example, using large-scale data sources Andolabehere and Hersh (2012) study vote reporting, Bond et al. (2012) study peer effects in a GOTV message, and Bonica (2013) uses such data to measure ideology and its consequences. These scholars are using new data and methods to expand our inquiry in ways that were previously not possible due to data constraints. We do the same, by using a GOTV message that was delivered to more than 61 million individuals we are able to gauge not only its effectiveness, but also among which groups the message was more or less effective.

This experiment we designed departs from prior work on the subject of social network information and voting behavior by conducting an experiment designed to prime voters to think about the voting behavior of their friends to differing extents. Subjects in our experiment were exposed to one of two treatment conditions or a control. In one treatment condition, users were reminded of election day and encouraged to turnout. In the second condition, users were given the same information and additional information about the voting behavior of friends. This social information was intended to prime the normativeness of voting within that individualÕs social network. That is, users who saw that friends had vote were encouraged to think of voting as a socially desirable act. As such, we were able to distinguish the impact that informational appeals to voting have from appeals that also include an appeal to the normativeness of voting through the inclusion of friendsÕ voting behavior.

This study makes several important contributions. First, we provide evidence that online appeals to voting can increase turnout. While previous work has suggested that many methods of GOTV contact can be effective (Gerber and Green 1999; Green, Gerber and Nickerson 2003; Gerber and Green 2001; Nickerson 2006; Gerber, Green and Green 2003; Krasno and Green 2008), studies of online appeals to civic participation have shown no effect (Nickerson 2007; Bennion and Nickerson 2010). Prior work has suggested that online appeals are too easily ignored and are therefore not useful methods of contact. Here we show that online appeals can be effective, which is an important finding considering the low-cost nature of online contact.

Second, we provide strong statistical evidence that priming the normativeness of voting within the individualÕs social network increases participation. This finding is consistent with previous work that has shown that priming the normativeness of voting through shaming increases turnout (Gerber, Green and Larimer 2008), but our experiment shows that shaming is not necessary to impact turnout behavior. As those authors pointed out, the social psychology literature has found that in some instances, people are likely to comply with appeals to social norms (Cialdini and Goldstein 2004), but in others people reject such appeals (Brehm and Brehm 1981; Ringold 2002). We therefore test whether the shaming is a necessary component of the appeal to voting through social norms and find that it is not.

Third, because our sample is large enough to do so, we are able to divide our sample in order to investigate the characteristics of individuals for whom social information is particularly important. Our sample consisted of 61 million individuals, by far the largest experiment on voter turnout to date. This allows us to answer not only the question of whether or not this type of contact in an online environment can be effective, but also to investigate the characteristics of people for whom it is particularly effective. Aral and Walker (2012) find that individuals vary in the extent to which they are susceptible to social influence based on their characteristics, finding that susceptibility decreases with age, women are less susceptible than men, and that those who are in more committed relationships are more susceptible to influence for product adoption. We find that the degree to which individuals respond to the inclusion of social information in appeals to participate civically varies by a similar set of individual characteristics: age, education, number of friends and relationship status.

1.2 Social norms and voting behavior

Recent work in social psychology has shown that social norms are capable of explaining, predicting and inducing behavior (for a good review, see Cialdini and Goldstein (2004); Cialdini and Trost (1998)). Such norms are believed to affect behavior in three ways. First, people must know that a norm exists. Second, one must accept the norm as a desired rule for behavior. Third, there must be a mechanism by which norms are enforced. Campaigns that wish to influence behavior through appeals to social norms affect at least one of these three mechanisms.

Voting in U.S. elections is widely seen as a social norm, though participation rates in many elections are still low. As an to attempt to induce greater awareness of and adherence to the norm to increase turnout, researchers have turned to field experiments. In Gerber, Green and Larimer's (2008) study on social pressure and turnout the authors sent postcards to potential voters that appealed to voting in one of four ways: by appealing to civic duty, by mentioning that voting is a public record and that the researchers will examine their behavior (the Hawthorne effect), by sending a postcard with the vote history of members of the household (social pressure within the household), the vote history of neighbors including the recipient (social pressure within the neighborhood). In the last two conditions, subjects were told that an updated mailing would be sent after the election. While groups in each of the conditions voted more than the control group, the largest increase was in the "neighbors" treatment. This suggests that the social pressure of neighbors being aware of voting behavior had a significant impact on turnout decisions.

Further studies of social pressure have examined how different types of appeals to norm compliance may affect behavior. Gerber et al. (2010) found that disclosing past turnout behavior through a mailing increases turnout, especially when the mailing disclosed a recent abstention. Panagopoulos (2010) investigated whether publishing either the names of voters (inducing pride) or abstainers (inducing shame) in the newspaper influenced turnout, finding that both treatments increased turnout. Importantly, not all voters were equally mobilized: the shame treatment mobilized both high- and low-propensity voters, while the pride treatment mobilized only high-propensity voters.

Our experimental treatment is intended to alter the extent to which subjects have information about the turnout decisions of their social contacts. In contrast to the work of Gerber, Green and Larimer (2008), we simply provide information about what social contacts are doing, while Gerber, Green, and Larimer also provide their are subjects with an incentive to comply with the social norm. As Goldstein and Cialdini (2011) put it, the experiment we conduct induces descriptive norms (what is done), while Gerber et al. induce injunctive norms (what ought to be done). While inducing both types of norms may be preferable, there are additional costs and risks associated with inducing injunctive norms. Inducing injunctive norms may require additional mailings or may backfire if the description of the norm misaligned with the desired behavior. For instance, if in the "neighbors" treatment described above the majority of neighbors did not vote, then the norm being induced may be that voting is not normative and could actually decrease turnout, the exact opposite of the desired result.

The most effective campaign that appeals to social norms would be one in which those who are most responsive to descriptive norms are exposed to them, those who are most responsive to injunctive norms are exposed to them, and those who are most responsive to both are exposed to both. However, often we do not know a priori which type of norm will appeal to which type of person. This study, therefore, helps us to understand both the circumstances in which a simple, descriptive norm is effective at increasing participation and also the types of individuals that are most responsive to descriptive norms. This should assist both researchers who are looking for the types of individuals who are responsive to social information as well as campaign managers who are designing campaigns based on which potential supporters to contact.

1.3 Experimental Process and Results

To test the hypothesis that voting behavior is influenced through a GOTV message delivered via an online social network, we conducted a randomized controlled trial with all users who are at least 18 years of age in the United States who accessed Facebook.com on November 2, 2010, the day of the U.S. Congressional elections. Users were randomly assigned to a “social message” group, an “informational message” group, or a control group. The social message group (N=60,055,176) was shown a statement at the top of their “News Feed” (the home page that greets users upon entering the site). This message encouraged the user to vote, provided a link to find local voting poll locations, showed a clickable button reading “I voted” with a counter indicating how many other Facebook users had previously reported voting and displayed up to six small randomly selected “profile pictures” of the userÕs Facebook friends who had clicked the “I voted” button earlier that day (Figure 1.1). The informational message group (N=611,044) was shown the message, poll information, counter and button, but they were not shown any faces of friends. The control group (N=613,096) did not receive any message at the top of their News Feed. Balance tests showed no significant differences between the three groups in age, sex, ideology, or identification as a partisan, suggesting the random assignment procedure produced groups with no significant differences (see Table 6 for more details).

Because the social message group is shown the self-reported voting behavior of friends, users in this condition are encouraged to think of voting as a social act. They are aware of the voting behavior of some of their friends, and they are aware that their (self-reported) voting behavior will be broadcast to their friends. For these reasons, we argue that the social message condition increases the likelihood that an individual thinks of the social norms related to voting. Users in the social message condition are more aware of the norms related to voting in their social networks, and are thus more likely to take them into account when making their decisions about whether or not to participate.

The design of the experiment allows us to assess the impact that the treatments had on three dependent variables: clicking the “I voted” button, clicking the polling information link, and validated turnout. Clicking the “I voted” button is most similar to traditional measures of self-reported voting. However, unlike most instances of

Figure 1.1: Examples of the treatment messages. (a) Example of the social message condition. (b) Example of the informational message. Note that the social message and the informational message are identical save for the faces of up to six friends who had previously self-reported voting, the names of up to 3 additional friends who had previously self-reported voting and the number of previously voting friends shown at the bottom of the social message.

self-reported voting in which the respondent reports their voting behavior to a survey administrator, in our experiment users are self-reporting their voting behavior to their social community. Therefore, because clicking the "I voted" button is a public action, our measure of self-reported vote also measures the extent to which a It is important to note that because a userÕs self-reported voting is communicated to her online community, she may decide to report voting when she has not voted (an over-report) due to social desirability. Therefore, we view self-reported voting, not as a direct of measure of voting behavior, but in this context as a measure of political communication. Self-reporting measures the extent to which she desires to be seen as a voter and the extent to which she wishes to share voting behavior information with social contacts. Subjects may change their vote reporting due to a desire to be seen favorably in their social circles. Social desirability should not affect our other measures of participation in the same way, as they are private actions that are not reported to usersÕ social communities. Clicking the polling place link takes users to a separate website where they may search for information about where they may vote and clicking this link was not reported to friends. Therefore, this variable measures a userÕs desire to seek information required to vote and should not be influenced by social desirability. Finally, we are able to assess the extent to which the treatments affected validated turnout. Validated turnout is, of course, not communicated to friends through the website.

To assess the effect of the treatment on political communication, we focus on rates of reported voting using the "I voted" button. Because the control group did not have the option to click an "I voted" button, we cannot compare the treatment group to the control group that received no message at all. However, we can compare the proportion of users who clicked the "I voted" button between the two treatment groups to estimate the causal effect of exposure to social information (the faces and names of friends who had previously self-reported voting) on self-reporting. As previously reported by Bond et al. (2012), users who received the social message (self-reported turnout = 20.23%) were 2.09% (SE 0.05%, t-test $p<0.01$) more likely to report voting to their social contacts than those who received the informational message (self-reported turnout = 18.14%). This result is consistent with previous work showing that social information influences the decision to vote (Gerber, Green and Larimer 2008; Panagopou-

Table 1.1: Contingency table showing the relationship between validated voting and self-reported voting. "I voted" click represents the group that did click the "I voted" button on Election Day. No "I voted" click represents the group that was presented with the "I voted" button on election day but did not click it. It is important to note that the majority of individuals are on the diagonals and that the majority of those off the diagonal represent under-reports of voting rather than over-reports.

	"I voted" click	No "I voted" click
Validated Voter	1,591,192	1,663,944
Validated Abstainer	240,626	2,907,128

los 2010), though it is also possible that the faces in the social message merely draw attention to the message itself.

The difference in self-reported voting does not necessarily indicate a change in actual voting behavior. In order to evaluate the link between exposure to social information and actual voting, we measure the effect that the treatments had on seeking information about the election and real voting behavior. When evaluating the causal effect of social information on information seeking we found that users who received the social message were 0.26% (SE 0.02%, $p<0.01$) more likely to click the polling place information link than users who received the informational message. Second, we used a group-level matching procedure (Jones et al. 2012) to match 6.3 million publicly available voter records to measure the effect of the social message on real voting behavior. For these individuals we had measures of both validated voting and self-reported voting. The Pearson correlation between the two measures is 0.46 (SE 0.03, $p<0.01$). We found that only 3.8% of these individuals overreported voting, and that among those who did not correctly report their voting behavior, only 12.6% overreported (see Table 1.1). In short, the majority of the difference between the measure of self-reported voting and validated voting is a result of users underreporting their vote rather than overreporting. Nonetheless, we urge the reader to view the dependent variables as separate concepts. While we argue that all three dependent variables are important pieces of the voting process, they are not proxies for each other and should not be viewed as such.

Comparison of validated turnout rates shows that users who received the social message were 0.39% (SE 0.19%, t-test $p=0.02$) more likely to vote than users who received no message at all. Similarly, the difference in voting between those who received the social message and those who received the message was 0.39% (SE 0.17%, t-test p=0.02), suggesting that seeing faces of friends significantly contributed to the overall effect of the message on real world voting. In fact, turnout among those who received the informational message was identical to turnout among those in the control group (treatment effect 0.00%, SE 0.28%, p=0.98), which raises doubts about the effectiveness of information-only appeals to vote in this context.

The above results show that online political mobilization can have effects on voter behavior in the aggregate. Our focus in this paper, however, is how these effects vary based on the characteristics of the individual. The large-N nature of our study allows us to disaggregate the effects more easily than in previous studies due to the increase in statistical power we have to detect differences in behavior. Owing to our ability to only match 6.3 million of our experimental subjects to the validated voting record, and the small effect size we found on validated voting, we do not have enough power to detect subpopulation analyses on validated voting behavior. We therefore restrict our analysis of effects on subpopulations to the other two dependent variables, self-reported voting and information seeking, for which we have sample sizes large enough to detect significant effects in subpopulations.

To test for heterogeneous effects we begin by simply subsetting the sample based on pre-treatment characteristics and conduct t-tests on the resulting groups. While researchers have described new methods that help to find heterogeneous effects (e.g., Imai and Strauss (2011); Green and Kern (2010)), in our study we have enough observations to simply divide the sample and rely on the raw data to detect differential effects. This approach has two advantages. First, the analysis is much simpler and more tractable, especially given the size of the data. Second, the simplicity of the analysis makes it easier to explain and easier for others to evaluate.

All of our pre-treatment covariates come from user-supplied Facebook data. When people create a Facebook account they must enter their birth date to ensure they meet the minimum age requirement of 13 years. Users are also encouraged, though not required, to enter their sex at this stage as well. Therefore, for most of our sample we know the age and sex of the users. We began by looking for differences in treatment effect in based on these characteristics. For sex, we found no difference in treatment

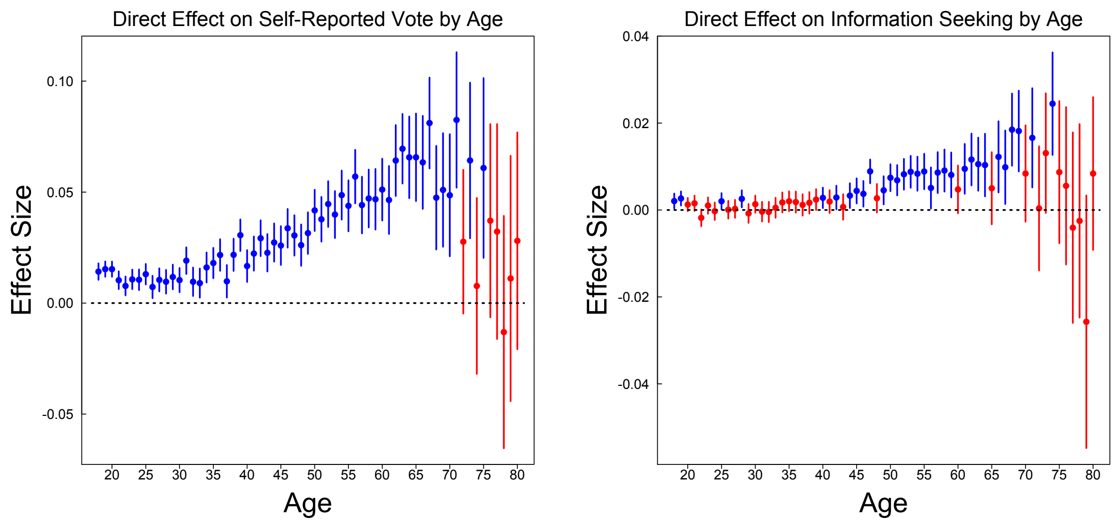

Figure 1.2: Self-reported voting and information seeking effects disaggregated by age. The left panel shows the difference in treatment effect on self-reported voting due to the inclusion of social information by age. The right panel shows the difference in treatment effect on information seeking due to the inclusion of social information by age. In each, points indicate the average treatment effect for the group and bars represent 95% confidence intervals.

effect between men (1.955%, 95% CI 1.806% to 2.103%) versus women (2.186%, 95% CI 2.056% to 2.316%) for self-reported voting. Similarly, there was no difference in treatment effect on information seeking for men versus women. We did, however find a difference in treatment effect by age. As users age, the inclusion of social information has a larger effect on self-reported voting (Figure 1.2). In fact, the effect size for those 50 years of age and older versus that of those aged 18-24 is nearly four times as large for self-reported voting and nearly 8 times as large for information seeking.

We next analyzed differences in treatment effect by education level. Facebook allows users to enter their education history, including the school attended by name and type (high school, college or graduate school), the year graduated (if applicable) and the degree obtained. We coded anyone who listed a graduation year from 2010 or prior as a graduate from that type of school and classified each user as a high school, college or

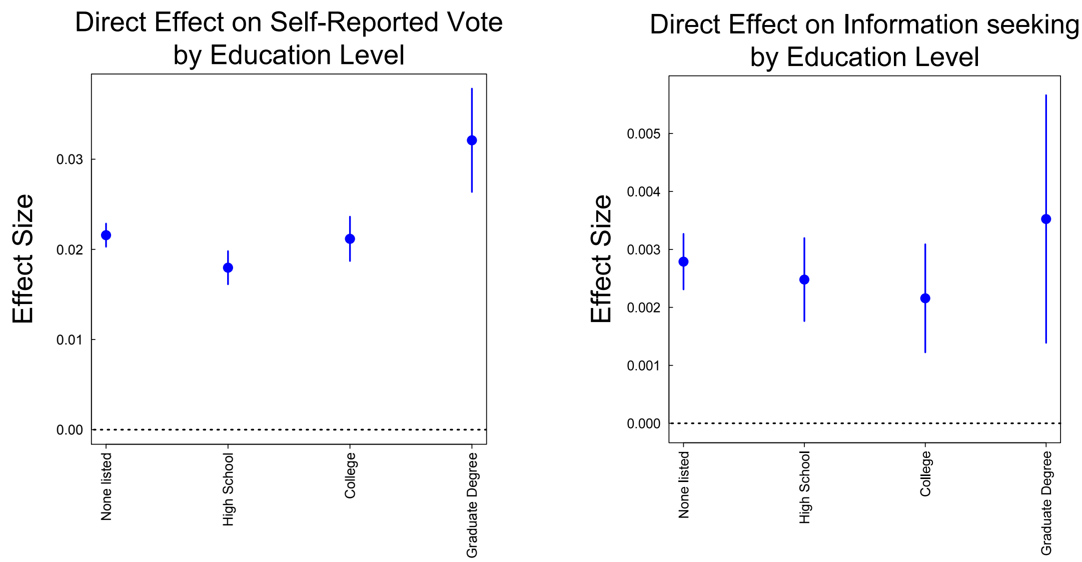

Figure 1.3: Self-reported voting and information seeking effects disaggregated by education. The left panel shows the difference in treatment effect on self-reported voting due to the inclusion of social information by education level. The right panel shows the difference in treatment effect on information seeking due to the inclusion of social information by education level. In each, points indicate the average treatment effect for the group and bars represent 95% confidence intervals.

graduate school graduate, taking the highest degree listed for each user. For instance, a user may list a graduation year in the future when they expect to graduate, so a user who listed a high school graduation year of 2008 and a college graduation year of 2012 would be coded as a high school graduate (as they would not yet have graduated college by the time of the election, while a user who listed a high school graduation year of 1992 and a college graduation year of 1998 would be coded as a college graduate. While we found no significant differences in treatment effect on information seeking by education level, the treatment effect on self reported voting seems to increase as education level increases (Figure 1.3).

Next, we studied how treatment effect varies based on the number of friends an individual has. We hypothesized that users who have more friends should be more responsive to social treatment as they would have a larger number of friends to whom they expect their behavior will be reported. However, as the number of friends overall increases, the likelihood that the faces and names shown in the social message will be those of a close, influential friends decreases, as the faces are randomly drawn from the set of all friends who have previously reported voting. For users with many friends, the social treatment may actually be less influential because the faces shown in the treatment are less likely to be close friends whose behavior is likely to influence the userÕs decision.

We were interested in the possibility that the treatment effect varied both by the number of friends an individual has on Facebook and by the number of close friends that an individual has. The meaning of friendship on social networking sites is yet unclear, so we utilize both the total number of friends that a user has as well as a more restrictive measure that helps us to identify closer friendship relationships. To identify close friendships we used photo "tagging" behavior. On Facebook users can upload photos and then "tag" them with the names of their friends (akin to writing "Grandma Lucy, Mom, and Micaella" on the back of a photo in a physical photo album). We defined close friends as people who identified and tagged one another in at least one Facebook photo (Lewis et al. 2008; Christakis and Fowler 2009) during the 365 days prior to the election. Tagging indicates that the friends are more likely to by physically proximate and suggests a higher level of commitment to the friendship, more positive affect between the friends and a desire for the friendship to be socially recognized (Lewis et al. 2008). Not all of these friendships will be close, but we expect them to be closer on average than those who do not tag each other in photos. We found a curvilinear relationship between the effect of the social message versus the message for both the number of friends and the number of close friends a user has (Figure 1.4). For both friends and close friends, we found that an increase in effect size for from the group with few friends or no close friends to the group with a moderate number of friends or close friends. However, among those who have at least a moderate number of friends or at least one close friends, more social contacts is related with less of an effect from the inclusion of social information. This pattern is true for both self-reported voting and for information seeking.

We next examined if treatment effect varied based on the relationship status of

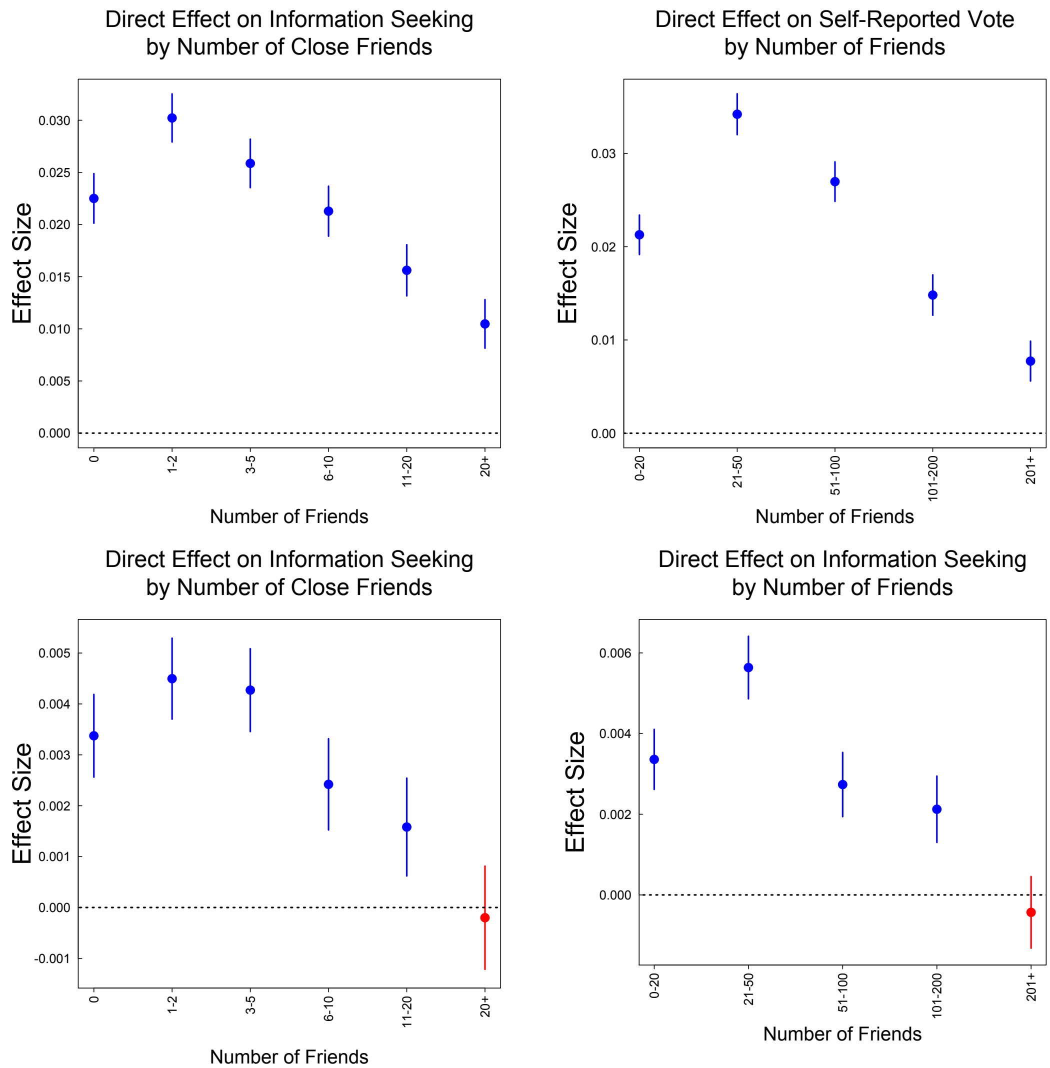

Figure 1.4: Self-reported voting and information seeking effects disaggregated by number of friends and close friends. The upper left panel shows the difference in treatment effect on self-reported voting due to the inclusion of social information by the number of close friends. The upper right panel shows the difference in treatment effect on information seeking due to the inclusion of social information by the number of close friends. The lower left panel shows the difference in treatment effect on self-reported voting due to the inclusion of social information by the number of friends. The lower right panel shows the difference in treatment effect on information seeking due to the inclusion of social information by the number of friends. In each, points indicate the average treatment effect for the group and bars represent 95% confidence intervals.

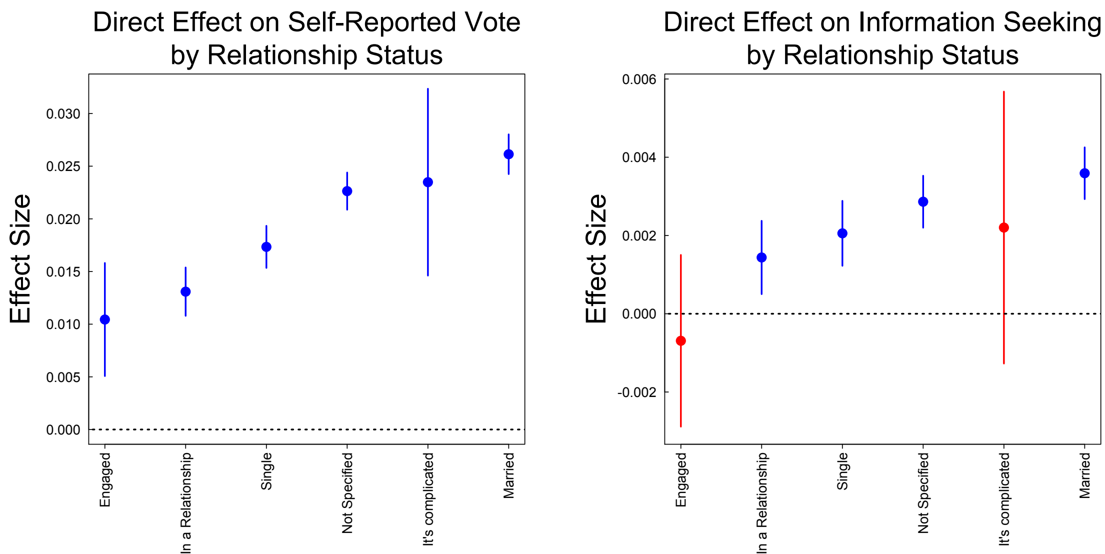

Figure 1.5: Self-reported voting and information seeking effects disaggregated by number of friends and close friends. The upper left panel shows the difference in treatment effect on self-reported voting due to the inclusion of social information by the number of close friends. The upper right panel shows the difference in treatment effect on information seeking due to the inclusion of social information by the number of close friends. The lower left panel shows the difference in treatment effect on self-reported voting due to the inclusion of social information by the number of friends. The lower right panel shows the difference in treatment effect on information seeking due to the inclusion of social information by the number of friends. In each, points indicate the average treatment effect for the group and bars represent 95% confidence intervals.

the individual. Facebook allows users to state the status of their romantic relationships (e.g., “Married”, “Engaged”, “In a Relationship”, “ItÕs complicated” or “Single”). More than 72% of the individuals in our sample have indicated their relationship status on their profile pages. We found that there are substantial differences in treatment effect based on relationship status, as shown in Figure 1.5. The results show that married individuals are most sensitive to the inclusion of social information in the appeal to vote, while those who are engaged are least sensitive for both self-reported voting and information seeking. There does not, however, seem to be a pattern in which stronger relationships are necessarily more predictive of treatment effect.

Table 1.2: OLS regression results using pre-election demographic variables available from Facebook to predict each of the three political behaviors.

	Validated Vote			Political communication			Information seeking		
	Estimate	SE	P-value	Estimate	SE	P-value	Estimate	SE	P-value
Age	0.010	0.002	0.000	0.003	0.000	0.000	0.000	0.000	0.000
Female	-0.011	0.004	0.005	0.005	0.001	0.000	-0.002	0.000	0.000
Republican	0.009	0.058	0.883	0.075	0.016	0.000	0.004	0.006	0.559
Democrat	0.066	0.051	0.190	0.065	0.014	0.000	0.010	0.006	0.076
College	0.077	0.004	0.000	0.065	0.001	0.000	0.006	0.001	0.000
Conservative	0.188	0.025	0.000	0.142	0.007	0.000	0.013	0.003	0.000
Liberal	0.064	0.027	0.016	0.102	0.008	0.000	0.014	0.003	0.000
Number of close friends	0.000	0.000	0.850	0.001	0.000	0.000	0.000	0.000	0.000
Number of friends	0.000	0.000	0.834	0.000	0.000	0.000	0.000	0.000	0.000
Employed	0.026	0.011	0.019	0.040	0.003	0.000	0.005	0.001	0.000
Married	0.097	0.004	0.000	0.059	0.001	0.000	0.001	0.000	0.007
Constant	0.086	0.007	0.000	0.012	0.002	0.000	0.010	0.001	0.000
N	63,095			603,558			603,558		
R-squared	0.105			0.038			0.002		

Finally, we use the procedure outlined by Enos, Fowler and Vavreck (2012) to test for the effect that the treatment has on inequality in participation for each of our three dependent variables. First, we estimate an OLS regression model for each of the three dependent variables using pre-treatment characteristics without including treatment status, using only observations from the Ôcontrol groupÕ. Note that the Ôcontrol groupÕ is different when we study validated vote than when we study self-reported vote and information seeking. For validated vote the control group is the group that received no GOTV message at all. For self-reported vote and information seeking the control group is the group that received the informational appeal to vote without social information. Thus, for validated vote we are able to compare the control group to both the informational message group and the social information message group, while for self-reported vote and information seeking we are only able to compare the social message group to the informational message group. See Table 1.2 for the results of the regression of pre-treatment characteristics on the three dependent variables for the control groups. Using this regression model we compute a propensity score for each user to engage in each of the three participatory behaviors. The propensity score measures the individualÕs propensity to participate if they were not contacted at all (validated vote) or if they were contacted only with information about the election rather than social information (self-reported vote and information seeking). Finally, we rescale the propensity variable to have mean 0 and standard deviation 1 for ease of interpretation in the subsequent model in which it is an independent variable.

Using the propensity score, we then estimate a second regression in which we include the propensity to engage in the dependent variable without treatment as estimated

Table 1.3: OLS regression results of treatment status interacted with predicted probability of each of the three political behaviors.

	Validated Vote			Political communication			Information seeking		
	Estimate	SE	P-value	Estimate	SE	P-value	Estimate	SE	P-value
Constant	0.504	0.002	0.001	0.182	0.001	0.000	0.022	0.000	0.000
Social message	0.005	0.002	0.017	0.021	0.001	0.000	0.003	0.000	0.000
Predicted probability	0.161	0.002	0.001	0.070	0.001	0.000	0.006	0.000	0.000
Social message * Predicted probability	-0.001	0.002	0.502	0.008	0.001	0.000	0.000	0.000	0.791
Informational message	-0.006	0.003	0.823						
Informational message * Predicted probability	-0.002	0.003	0.435						
N	6,316,163			59,895,087			59,895,087		
R-squared	0.103			0.038			0.002		

by the propensity score from the models in Table 2 with the treatment. These models take the form

$$Participation = \alpha + \beta_1 \times treatment + \beta_2 \times propensity + \beta_3 \times treatment \times propensity + \varepsilon$$

In this model the coefficients we are most interested in are β_1 and β_3. β_1 represents the treatment effect for a user with an average propensity score (that is, an average likelihood of participation). β_3 represents the extent to which the effect of the treatment varies based on an individualÕs propensity to participate without treatment. If β_3 is significant and greater than 0, then the treatment is more effective among individuals who were already likely to participate, which means that the treatment increases inequality in participation. If β_3 is significant and less than 0, then the treatment is more effective among individuals who were less likely to participate without treatment, which means that the treatment decreases inequality in participation.

The results of the three models are shown in Table 1.3. For validated vote we find a negative but insignificant coefficient for the social message (β = -0.001, p = 0.502) and the informational message (β = -0.002, p = 0.435), indicating that the treatment affected low and high propensity individuals equally. For self-reported vote we find a positive, significant coefficient for the effect of social information (β = 0.008, p < 0.001). This indicates that for self-reported voting, the social message was more effective among those who were likely to self-report their vote with an information only message than those who were not. This means that the social message made participation in self-reported voting less equal than had users only received an informational appeal. For information seeking, we find a positive, but insignificant coefficient (β = -0.002, p = 0.435), indicating that the treatment affected low and high propensity individuals equally.

1.4 Discussion

In this study we used a large-scale GOTV experiment to examine how information about the voting behavior of an individualÕs social network influences the decision to participate. We disaggregated the main treatment effect on the pre-treatment sex, age, education and social connectedness. We observed no difference in treatment effect by sex, but the results of disaggregating by age showed that as people age they become more sensitive to the inclusion of social information in appeals to vote. We found that as users become more educated they seem to become more responsive to social information for self-reported voting, though this result is only suggestive. Finally, we found that for most people, those with more social contacts are less responsive to the inclusion of social information. Overall, while we found that providing social information increases participation more than providing information alone in the aggregate, which types of individuals are responsive varies considerably.

Next, we showed that the differences in treatment effect on self-reported voting led to those who were already likely to self-report to be more likely to do so if treated with social information. We found no such relationship for either information seeking or validated voting, indicating that for these behaviors the treatments affected users equally. This suggests that online GOTV treatments that are related to social information may create less equality of representation among those who self-report voting, but indicates that equality of representation in real-world voting is not affected by the treatment.

One significant limitation of the current study is the use of dependent variables, self-reported voting and information seeking, rather than validated voting, for most of the study of differential effects. As Vavreck (2007) points out, relying on self-reported measures of turnout can yield over-estimates of treatment effects. However, it should be noted that using validated turnout is not a panacea, as a recent comparison of self-reported turnout and validated voting concluded that, “(u)sing government records in lieu of self-reports, which can be both time-consuming and expensive, appears to inject more error than accuracy into measurements of registration and turnout" (Berent, Krosnick and Lupia 2011s). We feel that both measures have flaws, but both tap the concept we wish to study, political participation. We are somewhat fortunate in this regard, as we have measures of validated voting for a subset of our sample. While we do find a larger treatment effect for self-reported voting than for validated turnout, we have the ability to estimate the correlation between the two measures of participation. While we find that self-reported voting and validated voting correlate highly, we also found that Facebook users both under- and overreported voting to a substantial degree. Future research should investigate whether trends we find in self-reported voting and information seeking are similar for validated voting.

We also note that the treatment administered was an election-day message delivered through a social networking site. A more broad-based campaign, perhaps using multiple methods of contact, may be considerably more powerful. Nevertheless, our findings underscore the power of social information in affecting participation. In a more broad-based campaign, one may expect that social information would play a role not only for voting, but also for other forms of participation, such as registering to vote or donations to campaigns.

In addition, as with any randomized experiment, one must consider the replicability and generalizability of the study. Given the power of the experiment and random assignment we are confident that the effects we observed are real and are not false positives or the result of sampling error. However, we must be cautious and not assume that these results would hold in other settings. Our sample encompasses a large portion of the voting age population in the United States and more than 50% of U.S. adults are Facebook users, so we feel confident that the results generalize to the U.S. population overall. However, the generalizability to other electoral settings (such as presidential or local elections) is an area for future research. Further, because our treatment relies heavily on social information and social norms differ greatly by culture, it is not clear that treatment effects would be similar in other countries. Likewise, individuals in other countries may use technological tools, such as social networking sites, in different ways that may change how they react to GOTV campaigns such as these.

Overall, the above limitations aside, this study indicates that GOTV campaigns conducted online can have significant effects. Social information appears to be more effective than information about the election alone. The importance of social information in affecting participation varies greatly according to characteristics of the individual like age, sex, and relationship status. This information should help policymakers to plan successful GOTV campaigns, as they will have a better understanding of what types of appeals to participate are effective for which types of individuals. This should also help to equip researchers to find effects, either by targeting subpopulations when designing studies or by helping to guide data analysis once studies have been conducted.

As political campaigns increasingly use online tools, understanding how they work will be important for social science researchers. In addition, campaigns are collecting more data about people than ever before, making targeting of campaign materials to those who are most likely to respond to it easier. Studies like the present one help us to understand whether and how online appeals may work and, when they do, how campaigns may most effectively target their resources to have maximal effects.

Chapter 2

The Dynamic Spread of Voting

2.1 Introduction

Voting is, at least in part, a social act. People rely on their social contacts for information about how to vote, where to vote and whom to cast a ballot for. Recent experimental studies have examined some of these processes. Nickerson (2008) showed that sixty percent of the effect of GOTV contact in two person households transferred from the person contacted to the other member of the household. Gerber, Green and Larimer (2008) showed that turnout increased when those contacted were promised that their neighbors would be shown whether or not they voted. Bond et al. (2012) conducted a GOTV study on Facebook and showed that a GOTV message that included information about which friends of the contacted individual had voted increased turnout not only for the contacted individual but also among the recipients friends. Social contacts encourage each other to participate (McClurg 2004) and about which candidates they should support (Huckfeldt and Sprague 1995). In fact, information about elections comes not only from news sources and elites, but also from social contacts (Robinson 1976). This literature has struggled, however, to move from the observation of these facts, to a causal argument about how one person's behavior influences that of another. It is well known that individuals sort themselves into networks of likeminded people (Mutz and Martin 2001). This tendency makes distinguishing influence from homophily or the effects of a shared common environment difficult. Fortunately, methods have advanced such that we may find results that are suggestive of causal relationships in observational data.

Experimental research designs are the most conceptually clear method for distinguishing between influence and other factors. Political scientists are beginning to use such designs to study voting. In fact, recent experimental studies have suggested that encouragements to vote spread in social networks (Nickerson 2008; Bond et al. 2012). The experimental nature of these studies allows researchers to be more certain that effects that spread from person to person are causal and not due to other factors. However, while using experimental methods to study turnout gives us confidence in causal inference, it can also be limiting in the types of phenomena that we are capable of studying. These studies are of how an encouragement to vote spreads, rather than the voting act itself.

These studies force us to reconsider an assumption that has long been a core element of voting research, that of the atomistic individual. For decades, scholars have assumed that individuals make their voting decisions as individuals, with little or no influence from their social contacts (Campbell et al. 1960; Rosenstone and Hansen 1993; Wolfinger and Rosenstone 1980). More recent research (the above studies aside) on the effects of get out the vote (GOTV) campaigns (Gerber and Green 2000; Green, Gerber and Nickerson 2003) also focus their effects on individuals. The fact that encouragements spread from person to person mean that these studies have underestimated their effects. The assumption that individuals are atomistic and unaffected by their social environments, while a convenient one due to the nature of the data available to the researchers, seems unrealistic when we know that encouragements to vote spread.

While experimental studies have shown that the encouragement of voting behavior may spread in networks (Nickerson 2008; Bond et al. 2012), it is not possible to create a field experiment in which the voting behavior itself is randomly assigned. Instead, these researchers have randomly assigned the degree to which a social contact is encouraged to vote. While this research contributes to our understanding of how voting encouragements may spread and affect behavior, it does not allow us to answer the question of how one turnout decision affects another. This study does attempts to work around this problem by using observational data. However, the observational nature of this study means that establishing a causal relationship will be more difficult.

Using observational data comes with a cost. With observational data one is never able to completely distinguish influence from other omitted variables (Shalizi and Thomas 2011). Using careful research design, however, we are able to show that causal relationships are far more likely than the alternatives. These methods get us as close as possible to understanding how one turnout decision affects that of another.

2.2 Peer effects and voting

A significant body of research has suggested that the decision to turnout to vote is impacted by social factors. Turnout is highly correlated in social networks of friends, family and co-workers (Lazarsfeld, Berelson and Gaudet 1944; Berelson, Lazarsfeld and McPhee 1954; Campbell et al. 1960; Kenny 1992; Huckfeldt and Sprague 1995). That there is significant homophily in social networks regarding many behaviors, including voting, is well-established (Christakis and Fowler 2007, 2008; Huckfeldt and Sprague 1995). However, none of these studies have attempted to assess the impact that one individual's turnout decision has on those of her social contacts.

The explanatory mechanisms for social factors impacting political attitudes and behaviors, such as turnout, are simple and straightforward. Simply put, people take cues from their social environments about what political attitudes and behaviors are important. People may learn from their social contacts, imitate their behavior, ignore them, or do the opposite of what some others do. The behavior of others is, therefore, one of many factors that impacts an individual's decision-making about politics. In fact, in most instances the behavior of one's social contacts will have countervailing effects on the individual's ultimate behavior.

Recent experimental work has attempted to address the question of how turnout may spread from person to person. Nickerson (2008) showed that when a member of a two-person household was encouraged to vote by a canvasser that 60% of the increase in turnout among those contacted was transferred to the other member of the household. Bond et al. (2012) showed that a Facebook message increased turnout not only among the contacted individual, but also among that individuals friends and friends of friends.

2.3 Data and methods

We study how voting spreads from person to person through the combination of data on voting from public voter records and data on social relationships from the social networking website Facebook. The coupling of dynamic data on both turnout behavior and social ties allows us to characterize the extent to which turnout may spread from person to person. We employ exact matching in order to estimate the causal effect of a friend turning out to vote on self-turnout. Matching methods have been shown to be superior to regression methods when estimating causal effects in a social network (Aral, Muchnik and Sundararajan 2009). In order to do so, we also need control variables with which we may match users to one another. The sources of our data on turnout, social relationships, and control variables are described below.

Through the use of a probability matching method (Jones, Bond, Fariss, Settle, Kramer, Marlow and Fowler 2012) we match public voter records from 13 states. To choose which states to match, we identified those that provided (for research purposes) first names, last names, and full birth dates in publicly available voting records. From these, we chose a set that minimized cost per population. We then matched the turnout behavior from the 2008 and 2010 general elections of all individuals from those 13 states to data from Facebook from those 13 states using first name, last name and date of birth to match. See Jones, Bond, Fariss, Settle, Kramer, Marlow and Fowler (2012) for more information on the matching process and match rates.

Next, we must characterize the nature of social relationships between people. Facebook's data offer myriad ways to describe the social relationships between people. The most straightforward measure of friendship using this data is Facebook friendship. As of November 2010 the average Facebook user has approximately 150 friends on Facebook. While the behavioral similarities of Facebook friends is interesting, we are interested in the relationships between closer, real-world friendships, which we expect are often a subset of Facebook friendships. In order to measure real-world friendships, we use data from Facebook about the interaction between two people to estimate the likelihood that they are close friends. Some studies (Bond et al. 2012; Christakis and Fowler 2009) use photo friendships as a proxy for real-world friendships. Facebook friendships and photo friendships are a good starting point. In fact, Facebook friends have a significant correlation (cor = .05) in their self-reported voting behavior, as do photo friends (cor = .15). Recent work suggests that we are able to characterize the closeness of friendships using Facebook data (Jones, Settle, Bond, Fariss, Marlow and Fowler 2012). Bond et al. (2012) use the number of interactions that one user has with another over some specified time period to define close friendships. We use this method because it offers an improvement over photo tagging in that it includes multiple measures of the closeness of a relationship.

Finally, we use data from the Facebook profiles of users we could match to voting records in order to match them based on characteristics predictive of voting and characteristics of their friends that are predictive of voting. Variables are coded in the following ways:

- Age: user supplied date of birth. All users must input their date of birth when creating an account.
- Gender: user supplied gender. Most users input their gender when creating an account. Those records that did not include gender were removed from the analysis.
- College attendance: users who indicated in their profile that they had attended a post-secondary institution are coded as '1'. All other users are coded as a '0'.
- Marital status: users who indicated in their 'relationship status' that they are married are coded as '1'. All other users are coded as a '0'.
- Religious views: users who input a religious view in their profile are coded as '1'. All other users are coded as a '0'.
- Republican: users whose 'political views' included the word "republican" are coded as '1'. All other users are coded as a '0'.
- Democrat: users whose 'political views' included the word "democrat" are coded as '1'. All other users are coded as a '0'.
- Liberal: users whose 'political views' included the word "liberal" are coded as '1'. All other users are coded as a '0'.
- Conservative: users whose 'political views' included the word "conservative" are coded as '1'. All other users are coded as a '0'.

2.4 Matching

We match pairs of subjects where the friend (the 'alter') of the person whose behavior we are studying (the 'ego') voted in the 2012 presidential election to egos with alters who did not vote in 2012. The purpose of this process is to balance treatment and control groups on covariates that predict both the ego's voting behavior and the alter's voting behavior. By doing this we have treatment and control groups that are as likely as possible to have had an alter who voted, as this is our treatment variable. The criteria for the

Table 2.1: Match criteria.

	Subject Variables	Friend Variables
Age	Coarsened age[1]	Coarsened Age[1]
Gender	Male/Female	Male/Female
Marital Status	Married/Unmarried	Married/Unmarried
Religiosity	Religious views stated/unstated	Religious views stated/unstated
College	College attended/none listed	College attended/none listed
Partisanship	Democrat/Republican/Unknown	Democrat/Republican/Unknown
Ideology	Liberal/Conservative/Unknown	Liberal/Conservative/Unknown
Vote History	Turnout in 2008	Turnout in 2008/2010

[1]Coarsened age categories:
(20:24),(25:29),(30:34),(35:39),(40:44),(45:49),(50:54),(55:59)
(60:64),(65:69),(70:74),(75:79),(80:120)

matching are shown in Table 2.1. The pairs of subjects are matched exactly on a many-to-many basis. This implementation uses the methods described in (Ho et al. 2007). Matched cases (m_T) and controls (m_C) used in the analysis receive weights described in this equation:

$$w_i = \begin{cases} 1, & i \in T^S \\ \frac{m_T^S}{m_C^S}, & i \in C^S, (\times constant = \frac{m_c}{m_t}) \end{cases}$$

The control weights are the ratio of cases to controls in the matched stratum ($\frac{m_T^S}{m_C^S}$), multiplied by the ration of matched controls to matched cases in the trial ($\frac{m_c}{m_t}$). Unmatched cases and controls receive a weight of zero.

Aral, Muchnik and Sundararajan (2009) showed that a matching approach outperforms regression approaches in studying peer effects observationally. The matching criteria directly and indirectly account for a number of important factors which affect the comparability of the two types of dyads (dyads in which the alter voted and dyads in which the alter abstained). As voters are known to cluster in social networks, matching allows us to compare dyads that are as similar as possible. By matching on variables that are known to predict turnout we are best able to account for the fact that voters are likely to be friends with voters and abstainers friends with abstainers.

The vote history lag variables for the ego and alter are catch-all matching terms.

Turnout is known to be influenced by a number of factors that are unmeasured in the present study, including habit (Gerber, Green and Schahar 2003), persistence (Denny and Doyle 2009), economic status, and genetic factors (Fowler, Baker and Dawes 2008), matching on a voting history lag variable partially controls for these factors for both the ego and the alter.

2.5 Calculation of Treatment Effect

Given the exact matching process and the size of the dataset (19.3 million treatment dyads and 18.1 million control dyads), calculation of treatment effects is non-parametric and makes few assumptions. We use average treatment effect on the treated (ATET) to calculate the effect of a friend turning out to vote on self-turnout. The calculation is the turnout rate of egos whose alter turned out to vote minus the turnout rate of egos whose alter abstained. We used generalized estimating equation (GEE) procedures to account for multiple observations of the same ego across ego-alter pairings (Liang and Zeger 1986). We assumed an independent working correlation structure for the clusters (Schildcrout and Heagerty 2005). From these regression estimates we calculate point estimates for the effect of alter turnout on ego turnout, standard errors and 95% confidence intervals. We next calculated the effect of an alter voting on ego turnout by simulating the first difference in the alter's turnout (changing from 0 to 1) using 1,000 randomly drawn sets of estimates from the coefficient covariance matrix.

2.6 Results

After we have matched voters and calculated the average treatment effect for voters who had a friend who voted in the same election, we repeat the procedure for groups of friendship as defined by the closeness of the relationship. The overall results appear in Figure 2.1.

In Figure 2.1, each point represents the estimated turnout effect (average treatment effect on the treated) for those voters who had a friend of that particular decile of interaction who turned out to vote in 2010. The average number of dyads in each decile of interaction was 2.8 million, and the successful match rate for dyads was 78%. Unmatched cases receive a weight of zero and are not included in the estimate (according to the procedure described above from (Ho et al. 2007)). The x-axis is the decile of interaction (normed for the ego's total interactions across all friends) between the two individuals, where the first decile is friends who rarely interact, and the 10th decile is friends who interact often. The y-axis is the estimated effect of the alter voting on the ego as calculated by simulating the first difference using the regression estimates post-matching.

There are three main results. First, turnout rates are higher overall when a friend also turns out to vote. For all of the groups of dyads, when the alter turns out to vote the ego is at least 6.7% more likely to vote. Second, the closest alters (the 10th decile) exhibit the most influence on ego turnout. Tenth decile alters who vote are associated with egos who vote 7.4% more often, while the next highest impact of alter voting is 7.1%. While the differences are statistically significant, we do not detect large differences in influence based on the level of interaction between the individuals. Third, there appears to be a U-shaped relationship between the level of interaction between dyads and the amount of influence that an alter has on ego turnout. While the closest friends exhibit the strongest influence on turnout (7.4%, 95% CI 7.3% to 7.5%), the second largest impact is among friends that interact the least (7.1%, 95% CI 7.0% to 7.2%). This is surprising, given that Bond et al. (2012) found that friend influence from a GOTV experiment increased with the level of interaction between the friends.

Next, we were interested in the predictiveness of alter characteristics on ego turnout. First, we estimated a simple model of ego turnout using ego characteristics as explanatory variables. Next, we estimated a model of ego turnout using *only alter characteristics*. We began by using only characteristics of the closest friend, then added in the characteristics of the next closest friend, and repeated the procedure until the model using alter characteristics only was more predictive using only alter characteristics than the model using only ego characteristics (using the area under ROC curves as a measure of the predictiveness of the model). Finally, we estimated a model using both the ego characteristics and the alter characteristics. The results of this process are shown in Figure 2.2.

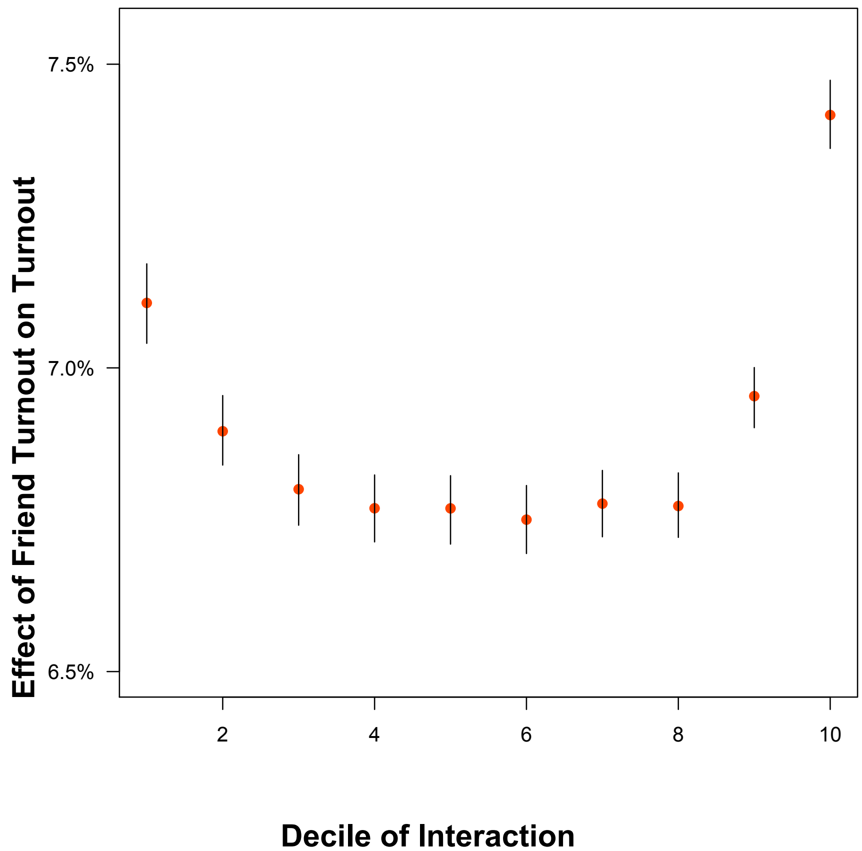

Figure 2.1: Simulated effects of alter turnout on ego turnout, separated by decile.

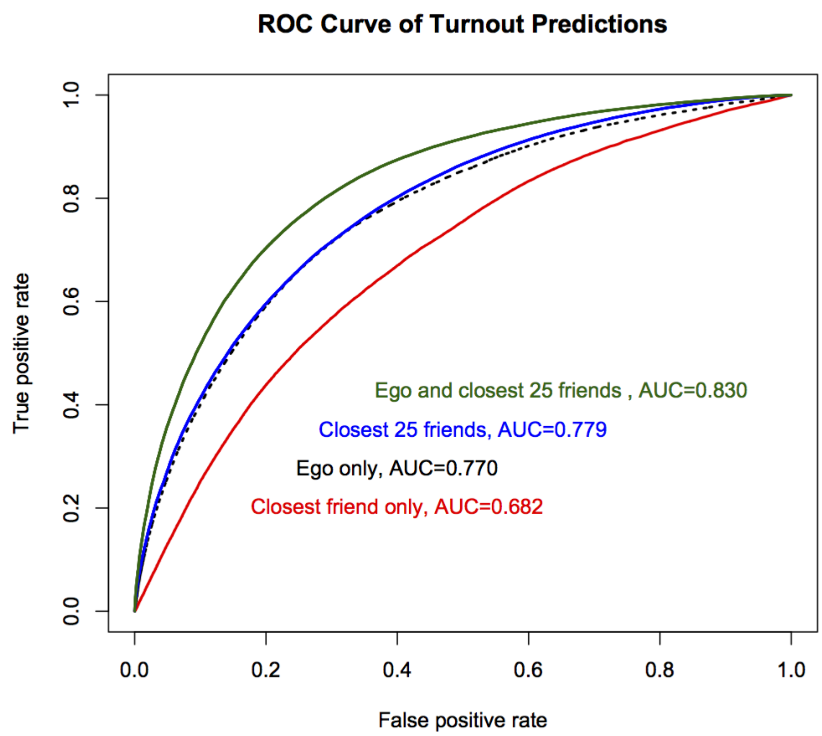

Figure 2.2: ROC curves for predictions of voting.

There are three main results from this analysis. First, the results show that we are able to predict ego turnout more effectively with alter characteristics alone than with ego characteristics, but that we need many alters (at least 25) in order to do so. It is important to note that only 6.6% of individuals have at least 25 alters who were matched to turnout records and for whom we have information with which we can predict ego turnout. Second, having just one alter (the closest friend only curve in Figure 2.2) does a good job of predicting ego turnout. This is not particularly surprising given that we know that there is significant homophily in turnout. Third, including alter characteristics with ego characteristics improves the predictiveness of the model, but by relatively little considering the costs of collection of alter data. Overall, while using alter characteristics alone to predict the behavior of an individual is possible, these results indicate that, when possible, using ego characteristics is far preferable.

2.7 Discussion

Here, we estimate that when a friend turns out to vote an individual is approximately 7% more likely to vote. Moreover, the influence of friends differs by the amount of interaction that the friends engage in prior to the election, with the closest friends exhibiting more influence than more distant friends. We find some evidence that the relationship between the closeness of a friendship relationship and the influence that friend has on the individual's turnout may be curvilinear, but that finding deserves further investigation. Finally, we find that friend characteristics are highly related to individual turnout. However, we note that while friend characteristics are predictive of turnout, using the characteristics of the individual is preferable. Without extensive knowledge of an individual's friendship network, the predictiveness of an individual's own characteristics far outweighs those of friends.

Political behavior is related to the social environment. While we do not test for the mechanisms by which a friend's turnout behavior affects the individual's turnout, it is clear that social factors play a role in determining whether or not an individual casts a ballot. We note here that this finding does not adjudicate a long standing theory of political science, that of rational voting. Specifically because we do not test the mechanisms by which friend turnout affects individual turnout we cannot say whether this finding supports a rational choice view of turnout or not. It may be that friend voting affects individual turnout by serving as an information shortcut (Popkin 1994) about the costs and benefits of turnout out to vote. In contrast, individuals may be turning out to vote because they anticipate social costs and benefits for their social contacts.

To put these results we find in perspective, Gerber, Green and Larimer (2008) estimate a treatment effect of 8 percentage points in a GOTV experiment using social pressure to increase turnout. Hobbs, Christakis and Fowler (forthcoming) find that widows are approximately 10 percentage points less likely to vote after a spousal death. Here, we find that the effect of a friend voting is similar, a 7 percentage point increase in turnout from friend turnout.

There are several limitations of this study. First, is that our sample is a convenience sample. Those who are on Facebook, and in particular are in the 13 states for which we could match voting records, may not be representative of others. That is, our results are specific to the population that uses Facebook as well as those who are located in specific states, which limits the external validity of our study. Second, we are not certain of the reasons that individuals are more likely to vote when friends vote. While previous explanations (social encouragement, imitation, information sharing, etc.) seem plausible, we do not test any of these mechanisms in this study. Third, our study is limited to non-presidential elections. While we have data on turnout in a presidential election included here, it is only used as a match variable. Future work should investigate if these processes differ in presidential elections, where the incentives to engage in social mobilization may differ.

2.8 Conclusion

A friend turning out to vote significantly increases one's own propensity to vote. While previous experimental studies have shown that experimental treatments may influence not only the individual who is encouraged to vote, but also the social contacts of the individual, our results estimate the effect that an individual's turnout decision has on her social contacts. This is a significant way in which our work departs from previous research on social mobilization by directly testing the effect of one individual's turnout on another's. This approach differs significantly from previous studies in that we are able to estimate not how an experimental treatment may affect social contacts, but how an indvidual's actions affect those of others. Further, we are able to differentiate the effects based on the closeness of the relationship.

We hope that the findings we have discussed here serve as a encourage further research into social mobilization and the social aspects of political behavior. While scholars have long researched the political behavior of individuals, new techniques and data sources are making studying social influences on political behavior more fruitful. The growing availability of large-scale data sources that contain information about not only the attitudes and behavior of individuals, but also about the nature of social connections between those individuals should encourage more discovery about how politics functions in our social system.

Finally, we note two areas that merit future study: the mechanisms by which friend turnout affects individual turnout and the electoral consequences of clustering in voting in social networks.

As we have made clear throughout the paper, we do not test any mechanisms by which friend turnout affects individual turnout. This is an important area for future researchers to study as it will tell us how the transmission of one person's political behavior to another occurs. This may occur for many reasons, from the functional explanation of one friend giving a ride to another to the polling place, to psychological reasons (such as those laid out by Hobbs, Christakis and Fowler (forthcoming)), to rational reasons such as cues on the rationality of an action. We feel that this is likely to take place for different reasons among different types of people, so multiple avenues of research should investigate these possibilities.

In this paper we have restrained from speculating about the electoral consequences of our findings. However, we should note that our work and the work of many others has shown that political behaviors and attitudes cluster in our social networks. For instance, Huckfeldt, Johnson and Sprague (2004) show that ideology is clustered in social networks. With both voting and ideology clustered in our social networks we may be at risk for having social networks that are highly divided with those participating in our electoral processes only in contact with like-minded people. The findings we present here are only one step towards demonstrating that possibility, and future work should continue down that path.

Chapter 3

Estimating Ideology using Facebook's 'Like' Data

3.1 Introduction

Many theories in political science rely on ideology at their core, whether they are explanations for individual behavior and preferences, governmental relations, or the links between the two. Because ideology plays such an important role in these theories, a great deal of research has been conducted on techniques to measure and estimate the ideologies of various political actors. While reliable techniques for measuring individual ideology, such as on a 7-point survey question, or legislator ideology, such as through the use of roll-call vote analysis, are well-established, methods for jointly estimating the ideologies of the general public and elite actors have only recently been developed. To a large extent, the lack of joint estimation of individuals and elites can be attributed to a lack of data that can be used for such estimations. While scholars continue to develop methods that use text as data for measuring ideology (Laver, Benoit and Garry 2003; Monroe and Maeda 2004; Monroe, Colaresi and Quinn 2009) that should one day be broadly applicable to a diverse set of political actors, data such as roll-call votes (Poole and Rosenthal 1997), cosponsorship records (Aleman et al. 2009), campaign finance contributions (Bonica 2013), and – as in this paper – public showing of support for elites by members of the general public remain the most promising avenues for estimating ideology.

In this paper, we contribute to the effort to jointly measure elite and the general public's ideology. While previous methods for jointly scaling elites and the general public using campaign finance records are promising (Bonica 2013), they suffer from the fact that only donors who donate at least $250 to a particular campaign are named in campaign finance reports. Many donors, however, give to campaigns in much smaller amounts. In fact, 46% of the donations to Barack Obama's 2008 presidential campaign were for $200 or less (Malbin 2009). Thus, CFscores are able to estimate ideological positions for politicians and PACs, but the estimates of ideology for individuals are largely for an elite class of donors who give large amounts to campaigns in any given election. In contrast, the data we use (Facebook 'likes', explained in more depth below) do not cost the user anything, and any Facebook user may like any page she wishes. Furthermore, the set of pages that a user can like is limited only by the set that have been created, meaning that a analysis of this data allows for estimation of a broad range of political actors.

Typically, one cannot estimate the ideologies of a diverse set of political actors from separate political institutions (e.g., Supreme Court justices and Senators), as the choices they make are disjoint. That is, Senators do not vote on Supreme Court cases, nor do Supreme Court justices vote on Senate bills. In order to estimate the ideologies of actors with (primarily) disjoint choice sets one must have some set of choices that 'bridges' the choice divides (Gerber and Lewis 2004; Poole and Rosenthal 1997; Bailey 2007; Bafumi and Herron 2010). That is, to estimate the ideology of actors that typically do not decide on the same choices, one must take advantage of the opportunities that that arise when they are required to make such decisions. These decisions 'bridge' the two sets of choices, making joint estimation possible. As with campaign contributions, Facebook's data on which users support which candidates represent the type of bridge actors necessary to jointly estimate the ideology of politicians and the general public. Users are able to 'like' pages regardless of their political affiliation, meaning that all pages are part of the same choice set. Bridge actors, such as Facebook users or donors, serve two important purposes. First, they bridge political actors that otherwise would not be connected, such as members of Congress and mayors of cities. Second, they span the divide between politicians and individuals. While most previous measures of ideology are based on the decisions of political actors, using Facebook users as bridge actors we are able to estimate the ideology of both politicians and the general public together.

In order to put elite political actors and the general public on the same ideological scale we use data from Facebook that mitigates the limitations that previous measures of ideology faced in terms of bridging diverse sets of elite actors and including measures of the general public's ideology. For Facebook users, showing support for a political figure is simple, relatively costless, and requires little cognitive effort or knowledge about issues. These factors should alleviate some of the difficulties that previous researchers faced. Our approach uses singular value decomposition (SVD) on the transformed matrix of user to political page connections on Facebook to estimate the ideological positions of Facebook users and the political pages they support. These estimates are consistent with the first ideological dimension recovered from roll-call data (Clinton, Jackman and Rivers 2004; Poole and Rosenthal 1997) and with users' self-reported political views indicated on the user's Facebook profile page.

The paper proceeds as follows. Section 2 explains how the Facebook like data are structured and the extent to which they are applicable to the estimation of ideology. Section 3 describes our estimation technique and describes the results of the estimation. Section 4 compares the estimates we obtain to other measures of ideology for subsets of both the elite and the general public. Section 5 shows an application of the data to test a hypothesis about the structure of ideology in social networks. Section 6 discusses the contribution this technique makes to the study of ideology as well as the opportunities for future work using this data.

3.2 Facebook 'Like' Data

Social networking sites like Facebook have given individuals the opportunity to connect with political figures and elites more readily than ever before. On Facebook users create profile pages where they list characteristics that they wish to share with their friends online. These may include their gender, birthday, political views, current place of residence, hometown, etc. In addition, users may list things that they 'like', such as television shows, books, movies, musicians, and well known figures or celebrities, each of which links the user to a separate page for the book, movie, etc., making the user a 'fan' of that page. The act of liking a political figure is communicated to the user's social community via the political figure's page, the user's page, and via the 'news feed' homepage that friends of the user will see when the log into the site. Further, Facebook maintains an accounting for which pages are 'official'[1]. The connection from a Facebook user to a political figure through the user liking the political figure's official page is the data we use to scale ideology for both the user and the political figure.

In considering the applicability of Facebook's like data to the estimation of ideology, one must first consider the data generating process. That is, what processes account for users liking political figures' pages? If Facebook users were to rely exclusively on

[1] For instance, there are many pages that are about president Obama in one way or another, but only one page (www.facebook.com/barackobama) is official and is maintained by the President (or, more likely, by his staff on his behalf).

spatial models of politics to determine which political figures to support, they would rank political figures based on their ideological proximity and like the set that are closest to them until the costs of liking another figure outweigh the benefit the user saw in liking that figure. Were users to behave in this way, modeling ideology from these data would yield ideological estimates with a great deal of precision.

However, Facebook users certainly do not behave in this fashion. In reality, the ideological proximity of the user to the figure is but one of many factors that individuals take into account when deciding which figures to support. A great deal of research has shown that voters evaluate candidates on many factors aside from ideology (e.g., race, gender, personality, electability, etc.). In addition to these characteristics, the way that a user interacts with the Facebook site no doubt influences her liking behavior. Some users prefer to keep their political preferences private and do not share such information on the site, while many others are willing to do so. Some users are also more active on the site, making them more likely to come into contact with the pages of political figures they support. Because Facebook users vary in the extent to which they interact with the site overall and the extent to which they engage politically with the site, a substantial number of Facebook users do register their support for political figures, but this number is far from a full accounting of users' preferences for political figures. As a result there is likely to be substantial under-reporting of preferences for candidates in the data, making estimates of ideology that result from this process in no way a random sample of Facebook users or even necessarily a full accounting of the preferences of those who do make their preferences known. Fortunately, the fact that so many users *do* report their preferences means that the sample will still constitute a substantial portion of the user population.

While the data-generating process is complicated because users vary in how they decide which pages they connect to and the extent to which they express those preferences to their social community, the structure of the data is quite manageable. The data can be organized as a contingency matrix, where the rows are pages maintained by political figures, the columns are individual Facebook users and each cell represents the presence or absence of a liking relationship from the user to the figure. This structure means that scaling techniques already familiar to political scientists are suitable for

Facebook political figure liking data.

As with previous methods that put a diverse set of political actors on a common scale, this method will allow for tests of theories from spatial models of politics that require data on the relative positioning of political actors. This method has the potential to uncover ideological estimates from legislators, the candidates for office they have defeated, bureaucrats, ballot measures and issues. This type of data will be important for our ability to study political phenomena concerning the interaction between legislator and constituent ideology, such as representation or vote choice.

One of the greatest benefits of conducting scaling using Facebook data is the wealth of information that the site stores about users. Such data holds tremendous potential for measuring how ideology varies across region, age, education, by profession, or by other groups available in the Facebook data. Further, the information about the social relationships that people have will allow for research into the ways that ideology clusters in social space. While much is known about the patterns of ideology in geographic space, social relationships do not always align with these patterns. Some neighbors are very close socially, while others are rarely interact, and some of our closest friends may live in distant locations. Using Facebook's data about social relationships we are able to study patterns in the ideological makeup of what people are exposed to through their social relationships, which we call social space.

For most of the results we show here, we used data from all U.S. users of Facebook over the age of 18 who had liked at least two of the official political pages on Facebook. For most of the results shown here, we used data from March 1, 2011. This constituted 6.2 million individuals and 1,223 pages. We also separately collected data from March 1, 2012 to see who ideology changes year over year below. Finally, we separately collected data from November 2, 2010, the day of the Presidential election that year, in order to calculate an ideology score that we could use to study its effects on voting, as explained below.

3.3 Using Facebook data to scale ideological positions

While roll call data are ready-made for scaling, and campaign contribution data require some processing in order to be ready for scaling, Facebook support for candidates data lies somewhere between the two. Roll call data are particularly clean because votes are coded as either 'yea' or 'nay', and abstentions may be simply treated as missing data. For both Facebook and campaign contribution data, the presence of a relationship is clear, but the absence of a relationship is ambiguous. The lack of a relationship may mean that the individual has decided against initiating a relationship with the political figure or it may owe to some other factor. In each case, individuals who choose not to supportto a candidate may do so because of ideological considerations or they may do so because they simply do not know about the candidate or are unwilling to make their supportive relationships public[2]. Additionally, for campaign finance data the fact that contributors may give at various levels makes the empirics more complex, a problem that Facebook's like data do not face.

As with campaign contribution data, Facebook users may choose to support any combination of political figures. While this is true, for campaign contribution data, giving to candidates is influenced by an individual's budget, which often may preclude an individual from giving to the full set of candidates they support, or from giving to any candidates at all. A Facebook user may choose to support no candidates, some subset of candidates, or may support all candidates. This differs from a set of choices among competing candidates in that there is nothing that precludes a user from supporting a candidate for office and her opponent. It would be tempting to simplify the analysis by treating the data as choices between incumbent-challenger pairs. However, many of the political figures we wish to scale run for office against minimal opposition or do not run for office at all, meaning that their opposition has few or no supporters.

[2]As noted above, Facebook liking relationships are made public through the profile of the user, the page of the political figure and the news feed. Campaign contributions of more than $250 are made public through campaign finance records that are publicly available.

3.3.1 Model of liking

Suppose n users choose whether or not to like m candidates. Each user $i = 1 \ldots n$ chooses whether or not to like candidate $j = 1 \ldots m$ by comparing the candidate's position at ζ_j and the status quo located at ψ_j, both in $\boldsymbol{R}^d$, where d = dimensions of policy space. Let

$$y_{ij} = \begin{cases} 1 & \text{if user } i \text{ likes candidate } j \\ 0 & \text{otherwise} \end{cases} \tag{3.1}$$

User i receives utility for supporting candidates close to her own ideal point $\boldsymbol{x}_i$ in $\boldsymbol{R}^d$ policy space. We can specify this ideological utility as a quadratic loss function that depends on the location of the candidate and the status quo:

$$\begin{aligned} U_{ij}^{candidate} &= -\|\boldsymbol{x}_i - \boldsymbol{\zeta}_j\|^2 \\ U_{ij}^{status\ quo} &= -\|\boldsymbol{x}_i - \boldsymbol{\psi}_j\|^2 \end{aligned} \tag{3.2}$$

The net benefit of liking is then the difference in these two utilities

$$\begin{aligned} U_{ij}^{like} &= U_{ij}^{candidate} - U_{ij}^{status\ quo} \\ &= -\|\boldsymbol{x}_i - \boldsymbol{\zeta}_j\|^2 + \|\boldsymbol{x}_i - \boldsymbol{\psi}_j\|^2 \end{aligned} \tag{3.3}$$

Notice that the utility of liking is *decreasing* in distance between the candidate and the user, but *increasing* in the distance between the status quo and the user.

Finally, suppose also that the utility of liking is increased by candidate-specific factors ϕ_j that govern how desirable each political page is (some pages are simply more popular and, perhaps, easier to find on the site) and user-specific factors η_i that govern each user's propensity to support candidates (some users simply get greater utility from the act of liking, and are thus more likely to engage in the activity than others). Putting these together with ideological utility yields:

$$U_{ij} = -\|\boldsymbol{x}_i - \boldsymbol{\zeta}_j\|^2 + \|\boldsymbol{x}_i - \boldsymbol{\psi}_j\|^2 + \eta_i + \phi_j \tag{3.4}$$

If we group row terms and column terms into new variables, this simplifies to

$$U_{ij} = -2\boldsymbol{x}_i\boldsymbol{\beta}_j + \eta_i + \theta_j \tag{3.5}$$

where $\boldsymbol{\beta}_j = \boldsymbol{\psi}_j - \boldsymbol{\zeta}_j$, and $\theta_j = \boldsymbol{\psi}_j^2 - \boldsymbol{\zeta}_j^2 + \phi_j$.

We do not observe direct utilities, but we do observe likes. Suppose that observing a like means that the true utility of liking is high, while not observing one means that the true utility is low (without loss of generality, suppose the utilities are 1 and 0, respectively). Not all likes yield exactly the same utility, so we can think of the true utility as being equal to a function of the observed like (y_{ij}) minus an error term (ν_{ij}):

$$U_{ij} = y_{ij} - \nu_{ij} \tag{3.6}$$

Substituting, we get:

$$y_{ij} = -2\boldsymbol{x}_i\boldsymbol{\beta}_j + \eta_i + \theta_j + \nu_{ij} \tag{3.7}$$

Define the *double-center operator* $D(.)$ for a matrix $\boldsymbol{Z}$ to be each element minus its row and column means plus its grand mean divided by -2:

$$D(z_{ij}) = (z_{ij} - \bar{z}_{i.} - \bar{z}_{.j} + \bar{z}_{..})/(-2) \tag{3.8}$$

In the literature that utilizes roll call votes to estimate ideology, Poole (2005) and Clinton, Jackman and Rivers (2004) discuss the use of the double-center operator and they use it on a *squared distance* matrix, not on the roll call matrix itself. The effect of this operator is to generate a new matrix with all row and column means equal to zero. As a result, any term that does not interact with *both* a row and column variable will factor out of the matrix.

Suppose ν_{ij} is an independent and identically distributed random variable drawn from a stable density. Suppose further, without loss of generality, that the dimension-by-dimension means of $\boldsymbol{x}$ and $\boldsymbol{\beta}$ equal 0. If so, then applying the double-center operator in equation (3.8) to both sides of equation (3.7) yields:

$$D(y_{ij}) = \boldsymbol{x}_i\boldsymbol{\beta}_j + \varepsilon_{ij} \tag{3.9}$$

where the new error term ε_{ij} is also a stable density defining the stochastic component of the identity. We can now use singular value decomposition (SVD) of the double-center matrix of likes to find the best d dimensional approximation of $\boldsymbol{x}_i$ and $\boldsymbol{\beta}_j$ (Eckart and Young 1936):

$$D(\boldsymbol{Y}) = \boldsymbol{X}\Sigma\boldsymbol{B} \tag{3.10}$$

where $\boldsymbol{Y}$ is the observed matrix of likes, $\boldsymbol{X}$ is an $n \times n$ matrix of user ideology locations, Σ is a $n \times m$ matrix with a diagonal of singular values, and $\boldsymbol{B}$ is a $m \times m$ matrix of βs. The d largest singular values correspond to the d columns of $\boldsymbol{X}$ and d rows of $\boldsymbol{B}$ that generate the best fitting estimates of $\boldsymbol{x}_i$ and $\boldsymbol{\beta}_j$ (Eckart and Young 1936).

While it is possible to analyze the full matrix of likes from users to political pages, the candidate (ϕ_j) and user (η_i) specific factors mentioned above bias the estimation. Because some pages are so much more popular than others, and to a lesser extent because some users like many more politicians than average, the estimation yields estimates of ideology that are weighted by the relative popularity of the candidates. For instance, Obama is to the *extreme* left of the distribution and Romney and Palin are to the *extreme* right, while candidates who have few likes are in the middle of the distribution regardless of their ideological views. To account for this, we create a distance matrix of political pages using the like data, as described below, that accounts for the candidate and user specific factors.

3.3.2 Estimation of ideology from liking

We begin by creating a matrix in which each column represents a user and each row represents a political figure's official Facebook page. We limit our data to Facebook users in the United States who are over the age of 18 who like at least two political pages[3]. We selected the set of pages that Facebook has determined are maintained by the person that it purports to represent, rather than pages created and maintained by others. This leaves us with approximately 6.2 million users and 1,223 pages and 18 million like actions from users to pages. An example of the first ten rows and first ten

[3]We exclude users who like only one page and pages with only one supporter as they do not add any additional information to the matrix, as explained below.

Table 3.1: The first ten rows of the user by political page matrix. Entries in the matrix are dichotomous, where a 1 means that the user has liked the page and a 0 means that the user has not.

	user 1	user 2	user 3	user 4	user 5	user 6	user 6	user 8	user 9	user 10
Barack Obama	1	1	1	0	0	0	0	0	0	0
Mitt Romney	1	0	0	0	0	0	0	1	0	0
Howard Dean	1	0	0	0	0	0	0	0	0	1
Joe Biden	0	0	0	0	0	0	0	0	0	0
Mike Bloomberg	1	0	1	0	0	0	0	0	0	0
Anthony Weiner	1	0	0	0	0	0	0	0	0	0
Deval Patrick	1	0	0	0	0	0	0	0	0	0
Diane Feinstein	1	0	1	0	0	0	0	0	1	0
Sarah Palin	0	0	0	0	0	0	0	0	0	0
Nancy Pelosi	0	0	0	0	0	0	1	0	0	0

columns of the bipartite matrix is in Table 3.3.2. A few things should be clear from this example and the summary statistics described here. First, there are few likes relative to the size of the matrix overall, making the matrix sparse. Second, users vary in the number of candidates they support. While we limit the data to include users who like at minimum two pages, the average number of likes is 3.04 and the maximum number of pages liked is 625. Figure 3.1 shows the full distribution of the number of likes for users and the number of fans per page. Clearly, most users like only a few pages, but a few like many. Similarly, political pages vary greatly in the amount of support from users they attract. Some pages attract few supporters; we only collect information from pages with at least 2 fans from the set of users we have selected. The maximum number of fans of a page is 3.67 million (Barack Obama), with an average of 15,422.5 fans per page. Most pages have a few thousand fans.

To estimate ideology our approach is quite similar to the approach used by Aleman et al. (2009) to estimate the ideal points of legislators in the United States and Argentina using cosponsorship data. Estimating separate parameters for users and pages, as is typical of estimation techniques like W-NOMINATE or Bayesian analysis (Clinton, Jackman and Rivers 2004), would be time consuming and difficult on a large, sparse matrix such as the Facebook like matrix. Instead, we construct an affiliation matrix between the political pages in which each cell indicates the number of users that like both pages. We do not use the original (two-mode) dataset of connections between users and

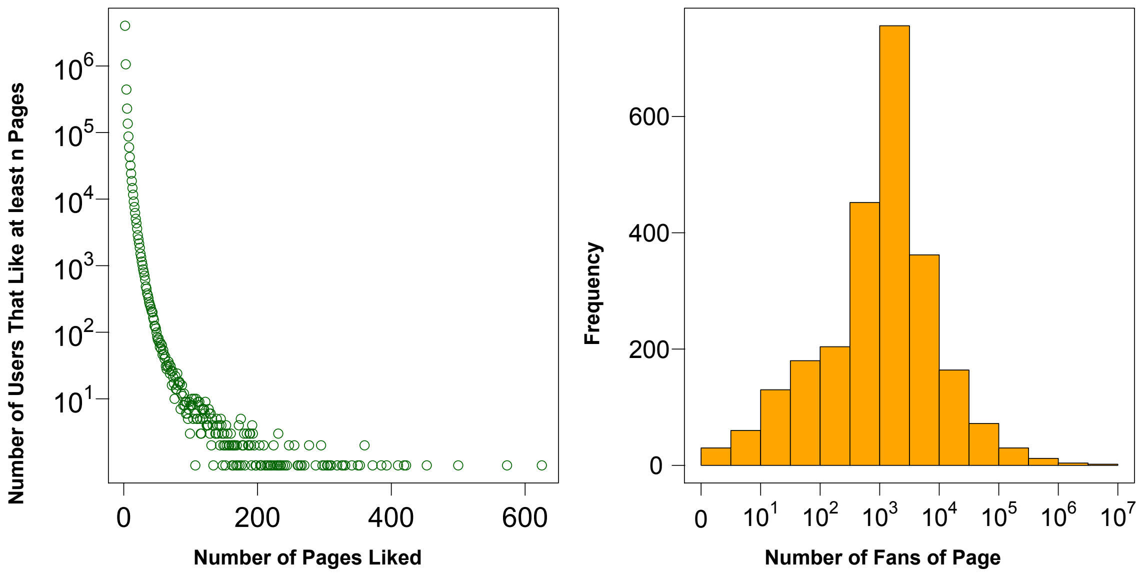

Figure 3.1: The left panel shows the distribution of the number of pages that each user likes. The right panel shows the distribution of the number of fans that each page has.

political figures, which is organized as an $X = r \times c$ matrix, with $r = 1, 2 \ldots R$ users and $c = 1, 2 \ldots C$ pages, but instead use an affiliation matrix (Table 3.3.2), $A = XX'$. In this affiliation matrix, the diagonal entries are the total number of users that like each page and the off diagonal entries are the number of times an individual user likes both pages. Table 3.3.2 shows the first ten rows and ten columns of the affiliation matrix, A. The table shows that there are very significant differences in the total number of users that like each page, as well as notable differences in the number of users that like each pair of pages.

Next, we calculate the ratio of shared users by dividing the number of users that like *both* pages by the total number of users that like *each* page independently, which produces an agreement matrix, $G = a_{ij}/diag(a_i)$, as depicted in Table 3.3.2. Because each page has a different number of fans, the denominator changes and the upper and lower triangles of the new square matrix, G, are not identical. For instance, Barack Obama has many fans on Facebook (the most of any page in the data set) so the values in the first row are small as they are all divided by the number of fans that Obama has.

Table 3.2: The first ten rows of the affiliation matrix. Diagonal entries are the number of user that are a fan of the page. Off-diagonal entries are the number of users who like both pages.

	Obama	Romney	Dean	Biden	Bloomberg	Weiner	Patrick	Feinstein	Palin	Pelosi
Barack Obama	3,675,192									
Mitt Romney	73,280	943,005								
Howard Dean	20,160	924	25,764							
Joe Biden	219,955	4,487	8,095	230,554						
Mike Bloomberg	12,873	2,658	831	2,408	20,076					
Anthony Weiner	31,169	2,158	4,618	7,437	1,270	42,211				
Deval Patrick	17,523	1,076	1,619	3,751	562	1,164	21,608			
Diane Feinstein	8,590	484	1,359	2,727	308	1,084	518	11,589		
Sarah Palin	142,022	627,793	1,229	6,356	2,748	3,066	1,006	634	1,649,936	
Nancy Pelosi	3,3467	2,562	7,280	10,185	1,196	6,117	1,660	2,363	3,429	41,690

Table 3.3: The first ten rows of the ratio of affiliation matrix. Because each page has a different number of fans, the denominator changes and the off diagonal entries are not identical.

	Obama	Romney	Dean	Biden	Bloomberg	Weiner	Patrick	Feinstein	Palin	Pelosi
Barack Obama	1.000	0.020	0.005	0.060	0.004	0.008	0.005	0.002	0.039	0.009
Mitt Romney	0.078	1.000	0.001	0.005	0.003	0.002	0.001	0.001	0.666	0.003
Howard Dean	0.782	0.036	1.000	0.314	0.032	0.179	0.063	0.053	0.048	0.283
Joe Biden	0.954	0.019	0.035	1.000	0.010	0.032	0.016	0.012	0.028	0.044
Mike Bloomberg	0.641	0.132	0.041	0.120	1.000	0.063	0.028	0.015	0.137	0.060
Anthony Weiner	0.738	0.051	0.109	0.176	0.030	1.000	0.028	0.026	0.073	0.145
Deval Patrick	0.811	0.050	0.075	0.174	0.026	0.054	1.000	0.024	0.047	0.077
Diane Feinstein	0.741	0.042	0.117	0.235	0.027	0.094	0.045	1.000	0.055	0.204
Sarah Palin	0.086	0.380	0.001	0.004	0.002	0.002	0.001	0.000	1.000	0.002
Nancy Pelosi	0.803	0.061	0.175	0.244	0.029	0.147	0.040	0.057	0.082	1.000

However, many of the fans of other candidates also like Obama (at least in the case of the other Democratic candidates in the example), so those values are relatively high.

It is notable that there is overlap in fans across partisan lines. Take for instance, the Barack Obama column in Table 3.3.2: this column represents the proportion of each of the other politicians' fans who are also fans of Obama. While it is not surprising that the other Democratic politicians have fans that are also fans of Obama, more than 7% of both Romney's and Palin's fans are also fans of Obama. This suggests that, for Facebook users who are fans of at least two candidates, there is not complete polarization among Facebook users.

The agreement matrix provides all of the information required to estimate ideal points from the liking data. From this stage, a number of methods to scale the data may be employed. For simplicity, we use SVD on the centered matrix, G. Note that, because we normalize the agreement matrix, which makes it asymmetric, the results from the left and right singular vectors are not similar. The right singular value is still highly related to the popularity of the page, as its denominator is the number of fans of the page itself. The left singular value does not suffer from the same problem, as its denominator is unrelated to the popularity of the page. Therefore, we retrieved the first rotated left singular value as the measure of ideology for the pages. We re-scaled the values to have mean 0 and standard deviation 1 for ease of interpretation. Figure 3.2 shows the proportion of the variance explained by each dimension. The first dimension explains approximately 19% and the first two explain 25%. This is far lower than the proportions of variance explained by the first two dimensions from scaling of roll-call votes (around 90%) or co-sponsorship data (between 70% and 90%) as calculated by Aleman et al. (2009). That the first dimension of the Facebook data do not explain as much of the variance as the first dimension in other data sources is not altogether surprising given that the Facebook data include many different political actors from different political institutions and levels of governance. The smaller explanatory power of the first dimension may owe to the actions of users as well, especially the sparsity of the matrix which owes to users frequently liking few politicians. Further, when members of Congress vote on bills the act of voting 'nay' on a particular bill is accounted for and easily interpretable, where the lack of a supporting relationship from a user to a particular page may owe to many factors (as explained above). This makes the matrix far more sparse than the roll call matrix, which means the matrix is less likely to be well-explained by a single factor.

We were next interested in estimating ideology scores for the users. If we were to scale the entire user by page matrix we would be able to estimate separate parameters for the users and the pages that should be measures of ideology. Another approach would be to do what we have done with the pages and to create a matrix of connections between users based on shared political pages that they both like. This would create a matrix of approximately 6.2M-squared entries. However, the original matrix is quite large, and the user by user matrix far larger, making the process computationally difficult enough that we take a different approach. Instead, we use the estimates derived from the pages

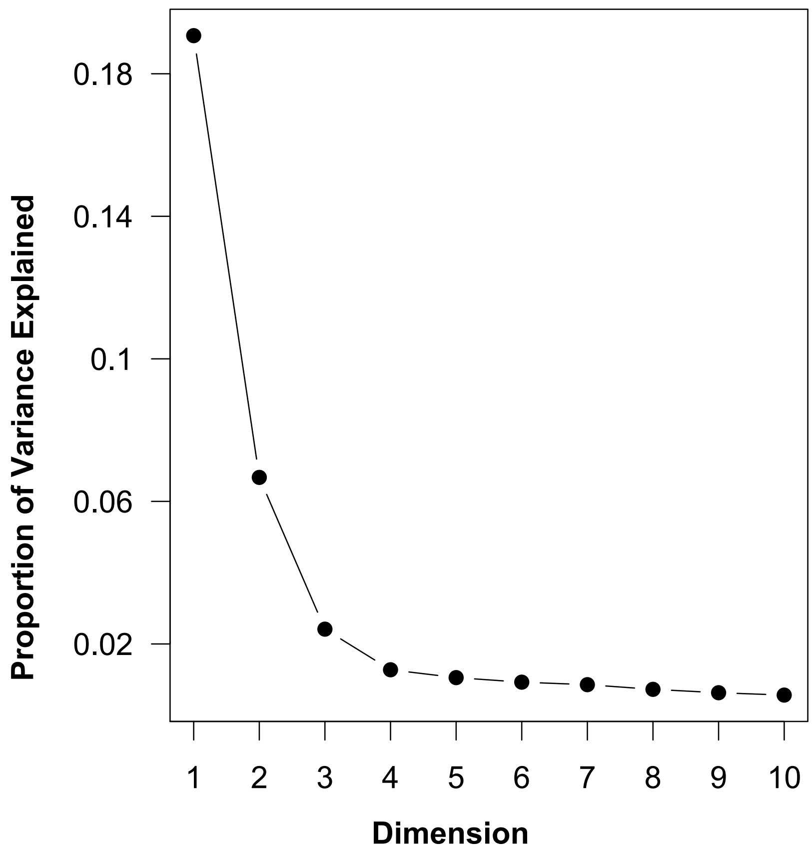

Figure 3.2: Proportion of the variance explained by each dimension of the scaled agreement matrix.

to estimate user ideology. We simply take the average score of all of the scores from the pages that we have already estimated as the score for the user. Doing so makes the calculation of user ideology much more computationally tractable.

We begin our exploration of the results of this process by examining the distributions of the ideology scores of the pages and the users, as shown in Figure 3.3. The figure shows that both users and pages are bimodally distributed, which is similar to Poole and Rosenthal's (1997) results for the U.S. Congress and Bonica's (2010) results for candidates for office, but it contrasts with the unimodal distribution of PACs. The bimodal distribution of the users is consistent with the finding that the American public are polarized as well (Levendusky 2009; Abramowitz and Saunders 2008; Brewer 2005; Hetherington 2001; Bartels 2000). It is also clear that there is a large mass of the distribution on the left, due to the large number of fans of Barack Obama. Finally, the distribution of candidates is more dispersed than the distribution for users. The distributions of both sets of actors show evidence of polarization, though the fact that the pages show more dispersion is suggestive that pages are more polarized than users.

3.4 Validation of the measure

To examine the similarity between legislators' positions as derived from roll-call vote data and those from Facebook liking data, we matched Facebook pages to their corresponding DW-NOMINATE first dimension scores. Of the 1,223 political pages scaled, we matched 465 to a member of the 111th Congress. As the estimates from SVD are rotationally invariant, we fixed Democrats to be on the left of the spectrum (negative numbers) and Republicans to be on the right (positive numbers). The pearson correlation between the two measures is 0.94, and the within-party correlation for Democrats is 0.47 and for Republicans is 0.42. This correlation is quite high given that Aleman et al. (2009) find that the ideological estimates from roll-call data and ideological estimates from cosponsorship data in the U.S. Congress correlate between 0.85 and 0.94. Figure 3.4 provides a visual representation of the relationship between the two ideology scores. The plot shows that the measures cluster legislators into two parties in much the same way and that the correlation within parties is quite high as well. The

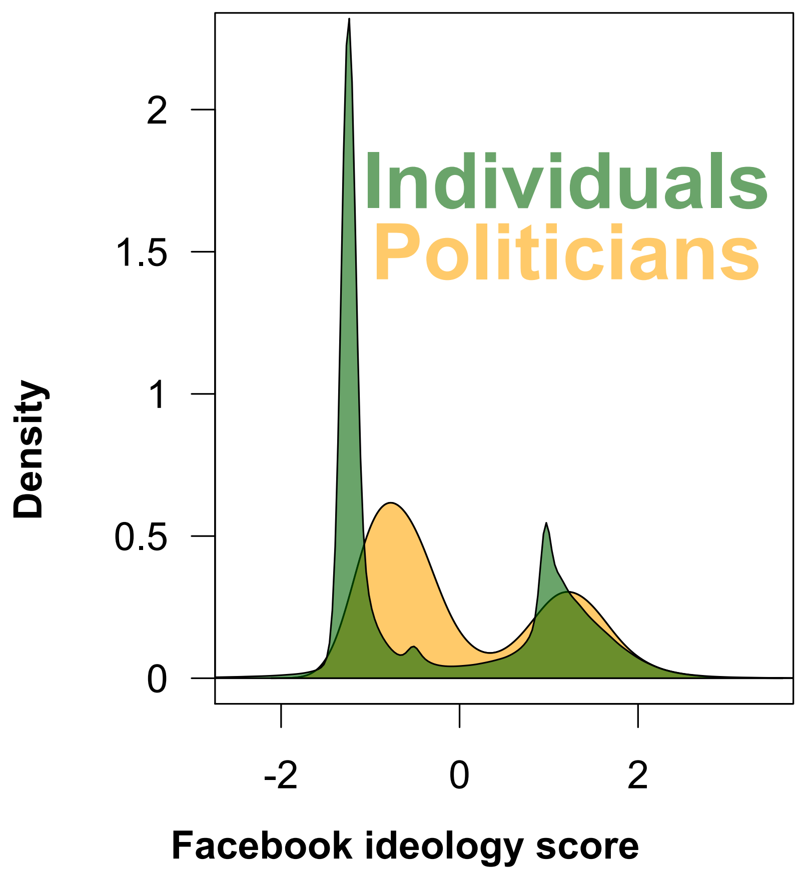

Figure 3.3: Density plots of ideological estimates of 1,223 pages and 6.2 million users.

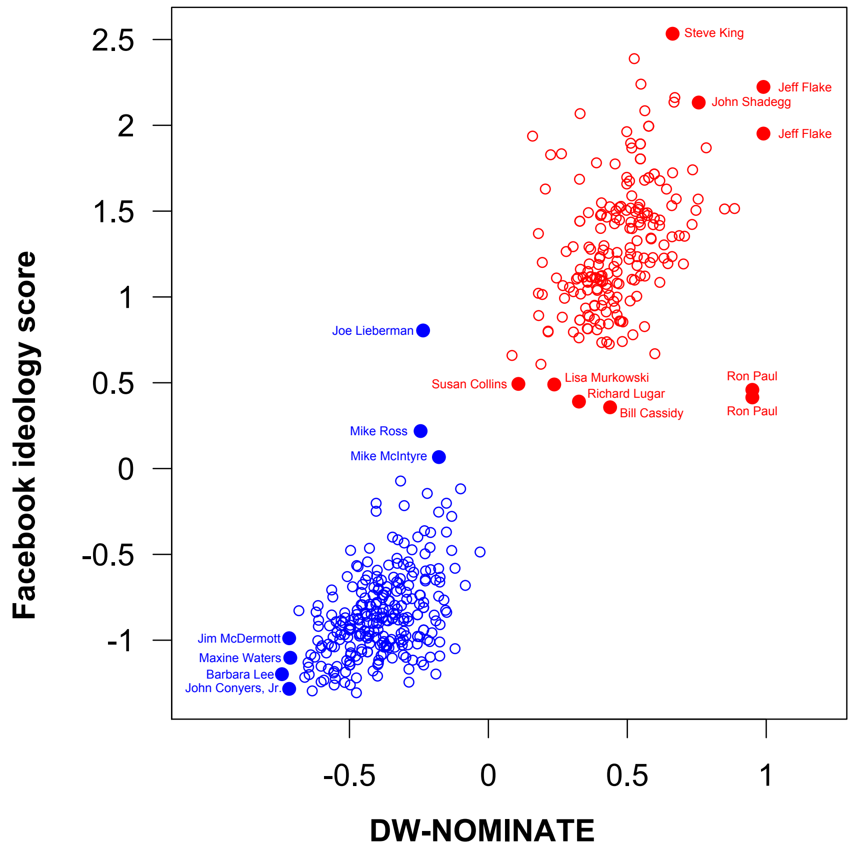

Figure 3.4: Scatter plot showing the relationship between the Facebook based ideology measure and DW-NOMINATE.

correlation among Democrats and the correlation among Republicans are both approximately 0.41. For comparison, Bonica (2013) finds that the overall correlation between DW-NOMINATE and CFscores scores among incumbents is 0.89, with a Democratic correlation of 0.62 and a Republican correlation of 0.53.

We note that there are several outliers in which the ideology score we estimate does not closely match the politician's DW-NOMINATE score. We have labeled some of the members of Congress in the figure for to illustrate where some of the most extreme members and some of the more moderate members lie on both measures. Note also that there are two points for Jeff Flake and Ron Paul in Figure 3.4. While most of the political figures in the dataset had only one official page, a few had more than one. Ron Paul maintained two official pages, one for his presidential candidacy and one for his work in Congress (this one explicitly asked for all discussion related to his presidential candidacy to move to the other page). The reasoning behind Jeff Flake's two official pages is less clear. However, we note that while it is somewhat unfortunate to have multiple pages representing the same individual, it allows us to see whether or not we get consistent estimates of ideology across multiple pages. Because the estimates for both Flake and Paul are so similar, we feel as though the estimates are reliable across separate Facebook pages.

To validate the user-level measure of ideology we computed the average ideology score of users based on their stated political views. On Facebook users' profiles there is a free response field that many users fill out called 'political views'. In this field users may type anything they like (subject to a constraint on the length). Many users type the same things in, such as 'Democrat', 'Republican', 'Liberal', 'Conservative'. We took all responses that more than 20,000 users had written in and calculated the average ideology score for the group, as well as the 95% confidence interval for that estimate. The results are shown in Figure 3.5.

This figure shows that the ideology score predicts users' stated ideology well. There appear to be at least three clear groups – those who state their ideology and are liberal, those who do not state a clearly liberal or conservative ideology, and those who state their ideology and are conservative. There is also substantial variation in the middle group, those that do not state a liberal or conservative ideology. The groups represented are in approximately the order one would expect based on their average ideology, save for the fact that those who self-identify as 'very conservative' are slightly to the left of those who self-identify as 'conservative'.

3.5 Age and Ideology

A substantial literature spanning across disciplines has found a relationship between age and political ideology (Glenn 1974; Ray 1985; Ross 1989; Krosnick and Alwin 1989; Cornelis et al. 2009). Most studies have found that as people age they become more conservative, and that as people age they become less susceptible to attitude change (Jennings and Niemi 1978; Krosnick and Alwin 1989). While the evidence

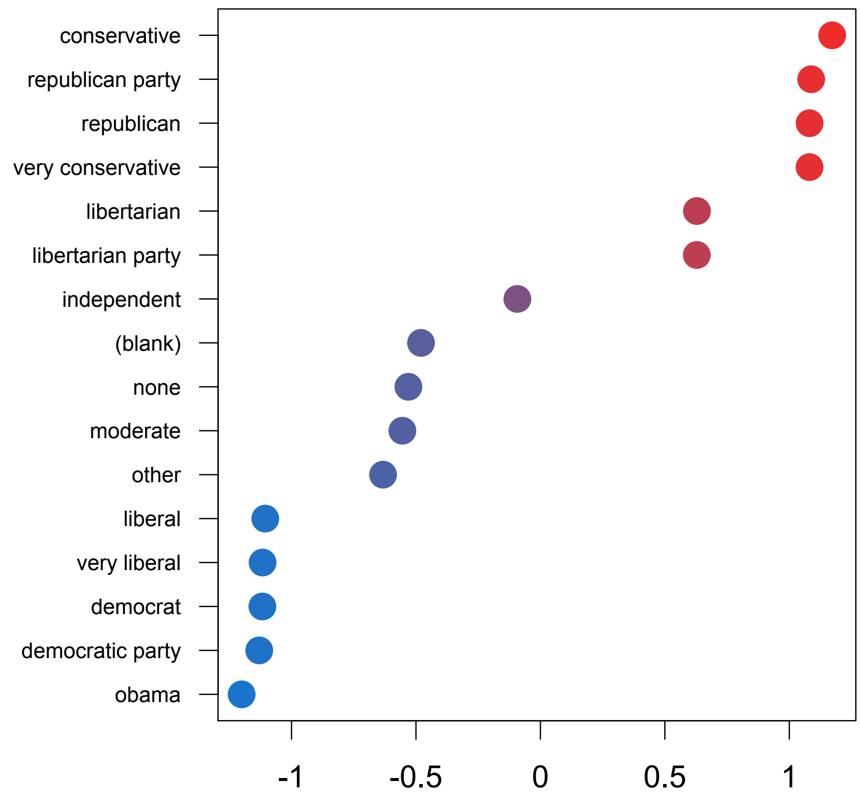

Figure 3.5: Average Facebook ideology score of users grouped by the users' stated political views. Note that the category labeled 'none' is the group of users that actually wrote the word 'none' as their political views. The point labeled '(blank)' is the group of users that has not entered anything in as their political views. Note also that the 95% confidence intervals for each of the estimates is smaller than the point. The color of the points is on a scale from blue to red that is proportional to each group's average ideology score.

for increased conservatism over the life cycle is strong, less is known about how that relationship varies across individual characteristics. Recent studies have begun to investigate how personality plays a role in mediating the relationship between age and conservatism (Cornelis et al. 2009). However, less is known about how the relationship between age and political ideology varies by characteristics such as gender, marital status and educational attainment.

A key advantage of using the Facebook ideology data is the large number of observations. By looking at patterns in the raw data we can understand phenomena that are not as easily understood using standard techniques, such as regression. In order to study ideology across characteristics, we took all individuals for whom we calculated an ideology score and matched the individual's characteristics as listed on their profile (See Supplementary Information for more detail on data coding). Figure 3.6 shows the average ideology of users age 18 through 80, as well as separating the estimates by gender, marital status and college attendance. The figure shows that for all groups, older people are more conservative than younger people. Women are more liberal than men, for all age groups, but a similar pattern in which older women and older men are more conservative than their younger counterparts emerges. While the overall pattern is similar for those who get married and those who do not, young people who get married are more conservative than their younger counterparts and young people who do not get married are more liberal than their younger counterparts. After the age of approximately 35, the pattern of increasing conservatism with age is similar across the groups. Finally, college attendance is not predictive of ideology among the young, but for older people having attended college is related to being more liberal than those the same age who have not attended college.

While these results are consistent with previous research, it should be noted that we have not studied change in ideology over time. The finding that older people are more conservative than younger people is consistent with a population that becomes more conservative over time. However, it is also consistent with recent surveys that have shown that the most recent generation is among the most liberal in recent memory (Kohut et al. 2007). While we have some evidence of how an individual's ideology changes over time, we find that the correlation in ideology from March 2011 to March

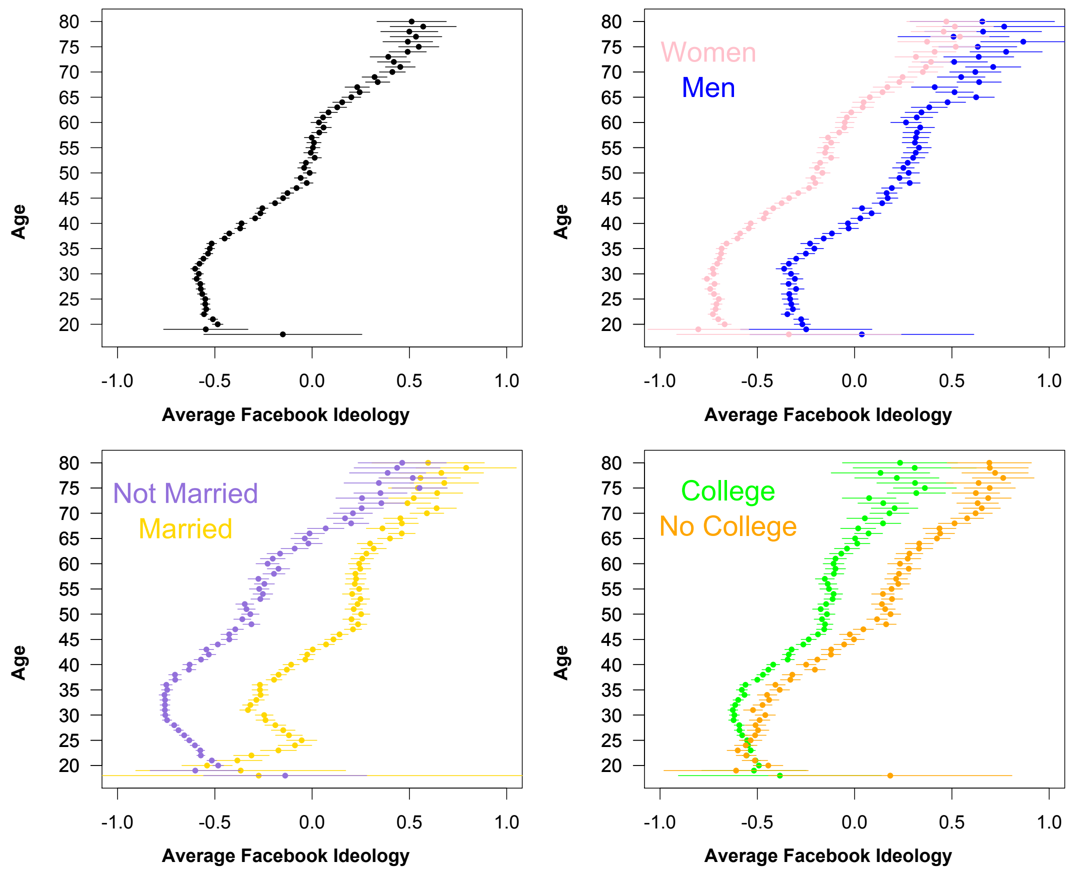

Figure 3.6: In each panel the points show the average ideology of a a specific age group for individuals age 18 through 80, and the lines represent the 95% confidence interval of the estimate. The upper left panel shows the average ideology of all users in our sample. The upper right panel shows the average ideology of men and women by age. The lower left panel shows the average ideology of married and unmarried individuals by age. The lower right panel shows the average ideology of college attendees and those who have not attended college by age.

2012 is 0.99. With more time we should be able to better discern whether the patterns we found are due to cohort effects or change in ideology over time.

3.6 Ideology in Social Space

One of the great advantages of using the Facebook data is the abundance of data about social networks and social interaction that it logs on a daily basis. We were interested in characterizing the ideological correlation among connected individuals. While previous work has shown that ideology does cluster in social networks (Huckfeldt, Johnson and Sprague 2004), we wish to characterize the extent to which the clustering varies based on the strength of the relationship between two individuals. We consider three types of relationships, each with varying types of ties: friendships, family ties, and romantic partners.

Clustering in the network may be due to some combination of three possibilities. First, clustering may be due to exposure to a shared environment. That is, friends may be both exposed to some external factor (for example, a war or economic depression) that influences both to change their ideology in a similar way. We may observe clustering in the network if both friends are both influenced by the same external factors. Second, clustering may be due to homophily. That is, people may be choosing friends based on political ideology. Third, clustering may be due to influence. That is, one friend may argue in favor of an ideological position and change their friend's views. We expect that each of these are more likely to occur between close friends than more distant friends. Closer friends are more likely to be physically proximate, which makes them more likely to experience the same external stimuli. Friends who are similar ideologically are more likely to become close friends as, ceteris paribus, they will have more in common. Finally, close friends will have more opportunities to interact and to influence one another about ideological views, making them more likely to be similar ideologically.

Due to the processes outlined above, we expect that as social relationships become stronger, friends will be more similar ideologically. Similarly, we expect this relationship to hold for romantic relationships as well. Recent work has shown that while people do not usually specifically look for matches based on political ideology when selecting romantic partners, they do base their decisions on other factors that are predictive of ideology (Klofstad, McDermott and Hatemi 2012). Therefore, romantic partners have highly correlated ideological views. We expect that as romantic relationships become more committed, that the correlation in ideology will increase. This could result from one partner influencing the other, from some external factor that both partners are exposed to, or from selection effects, in which relationships in which the partners are more similar ideologically are more likely to last.

Finally, we expect that familial relationships will show evidence of these processes. Members of a nuclear family should be more likely to experience the same external stimuli, due to being more likely to live near one another than friends are. While parents do not typically have the opportunity to select their children (or vice-versa) based on ideological views (or other characteristics related to it), there is evidence that there is a genetic component to ideology (Hatemi et al. 2010, 2011). This genetic component coupled with socialization and similar exposure to factors that influence ideology mean that ideology is likely to be highly correlated within family units.

Since Facebook allows users to identify their familial and romantic relationships, we were able to test for ideological similarity across these social links. We began by pairing all individuals with their siblings, parents, or romantic partners, for every pair that we had an ideological estimate for both users. We then calculated the Pearson correlation for each group. The results are shown in Figure 3.7. The figure shows that married couples have the highest correlation in ideology, while engaged couples have the second highest value. Correlations within the nuclear family have lower values, with parent-child relationships being stronger than sibling relationships.

Next, we were interested in the correlation of ideology between friends. First, we paired all 6.2 million users for whom we estimated ideology in 2012 with every Facebook friend for which we also had an estimate for ideology, for a total of 327 million friendship dyads and an average of approximately 53 friends per user. The overall correlation in 2012 was 0.69, which closely approximates other measures of ideological correlation among friends (Huckfeldt, Johnson and Sprague 2004). We repeated this procedure for 2011, with 6.1 million users who had 238 million friendship connections

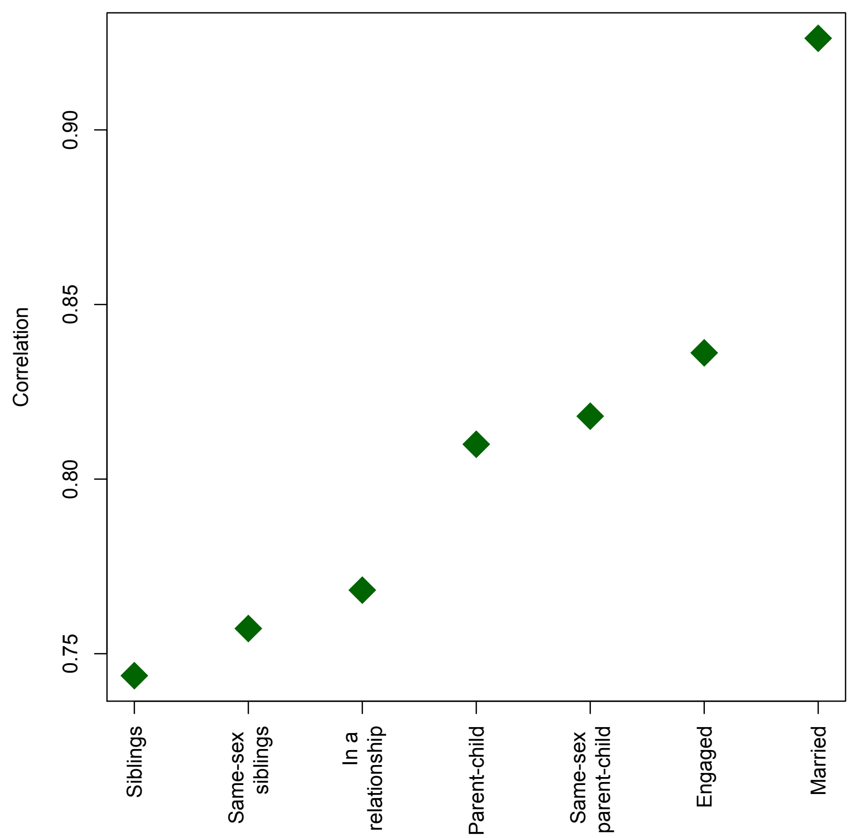

Figure 3.7: The correlation in ideology for familial and romantic relationships.

to other users for whom we could estimate ideology. The correlation over time of ideology within an individual from 2011 to 2012 is 0.99, indicating that by this measure user's ideologies are very consistent. While this is further evidence of ideological constraint, the importance of how the measure is constructed should not be lost on the reader, as it can greatly impact the measure of change over time (Achen 1975; Converse 2006). As few users changed the politicians they liked over the period, such a high correlation in ideology not entirely surprising. The reader should keep in mind that the correlation is not due only to few users changing the pages they like, but also the changes that did occur did not change the ordering of the ideologies of the pages to a significant degree. The overall correlation between ideology among these friends in 2011 was 0.67. The slight increase in correlation of friends' ideology from 2011 to 2012 is suggestive that there is greater polarization among friends in 2012.

Additionally, we categorized all friendships in each year of our sample by decile, ranking them from lowest to highest percent of interactions. Each decile is a separate sample of friendship dyads. For example, decile 1 contains all friends at the 0th percentile of interaction to the 10th percentile while decile 2 contains all friends at the 11th percentile of interaction to the 20th, and so on. We validated this measure of tie strength with a survey (see Jones, Bond, Fariss, Settle, Kramer, Marlow and Fowler (2012) for more detail) in which we asked Facebook users to identify their closest friends (either 1, 3, 5, or 10). We then measured the percentile of interaction between friends in the same way and predicted survey response based on interaction between Facebook friends. The results show that as the decile of interaction increases the probability that a friendship is the userÕs closest friend increases. This finding is consistent with the hypothesis that the closer a social tie between two people, the more frequently they will interact, regardless of medium. In this case, frequency of Facebook interaction is a good predictor of being named a close friend.

Using the decile measure of tie strength, we then calculated the correlation between user and friend ideology on each set of dyads for both 2011 and 2012 (see Figure 3.8). For both years the correlation in friends' ideology increases as tie strength increases. The proportion of interaction between friends is a better predictor of similarity of ideology between friends in 2012 than in 2011, again suggesting that in 2012

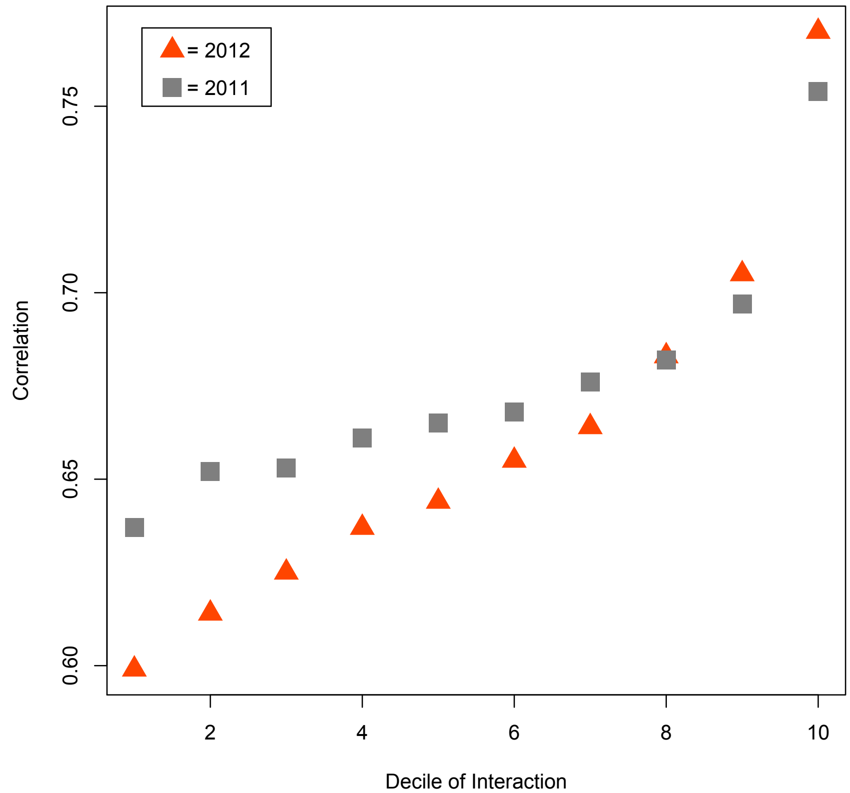

Figure 3.8: The correlation in ideology for friendship relationships. Each decile represents a separate set of friendship dyads. Decile of interaction is based on the proportion of interaction between the pair during the three months before ideology was scaled.

friendships are more politically polarized than in 2011. We caution that this correlation may increase for other reasons, such as better measures of ideology in 2012 owing to users liking more political pages, or perhaps to changes of the makeup of the set of individuals for whom we can estimate ideology. While we can only estimate ideology for about 2% more people in 2012 than 2011, the change in the sample could be enough to account for the difference in 2012.

We emphasize that we do not have a causal story to tell about the closeness of a friendship relationship and the correlation in ideology. Any combination of the three processes outlined above (homophily, influence and environmental factors) may explain the change in ideology that we observe. The data we have analyzed here cannot distinguish between these explanations for the observed relationship with any certainty.

3.7 Effect of Friend Ideology on Turnout

While the composition of ideology in social networks is important on its own, considerable research has been applied to understanding how the makeup of an individual's social network affects political participation (Mutz 2002; McClurg 2003; Robert Huckfeldt 2004; Huckfeldt, Johnson and Sprague 2004; McClurg 2006; Scheufele et al. 2004; Jr and Hively 2009). Scholars have long theorized about how cross-cutting pressures in an individual's social environment may cause an individual to become less interested in politics and to disengage (Campbell et al. 1960; Ithel de Sola Pool and Popkin 1956). More recent work has tested these theories about whether disagreement in an individual's social network affects the propensity to participate in politics (Mutz 2002; Huckfeldt, Johnson and Sprague 2004). This work has consistently found that exposure to disagreement depresses engagement and participation.

Previous studies have relied on snowball samples in order to construct social network measures. Social network sites, such as Facebook and Twitter, have allowed us to observe friendships without asking individuals about who they have discussed politics with. Survey respondents may be biased in recalling their discussion partners. If bias in recalling discussion partners is associated with the difference in ideology between a pair of individuals, which certainly may be the case if such discussions are more likely to be memorable, estimates of its affect on participation may be biased as well.

We seek to add to this literature by studying how exposure to disagreement affects validated voting. To do so, we matched in the validated voting records of all individuals from 13 states (see Jones, Bond, Fariss, Settle, Kramer, Marlow and Fowler (2012) for more information) to the Facebook data. We then matched each individual to their Facebook friends for whom we were able to calculate a measure of ideology and for whom we obtained a validated voting for each side of the friendship pair. Among the 6.2 million individuals for whom we calculated an ideology score, we could match a voting record for 397,815, with a total of 2,410,097 friendship pairs where both the individual and the friend had both an ideology score and a turnout record. Here we test

Table 3.4: The effect of friend ideology on ego turnout. Results of logistic regression of ego validated voting in 2010 on ego covariates and alter characteristics of ideology and turnout. Models were estimated using a generalized estimating equation with clustering on the ego and an independent working covariance structure (Liang and Zeger 1986; Schildcrout and Heagerty 2005). Models with an exchangeable correlation structure yielded poorer fit. Fit statistics show sum of squared deviance between predicted and observed values for the model and a null model with no covariates (Wei 2002).

	Estimate	Robust S.E.	p
Ideology Difference	-0.065	0.002	0.000
Ego Age	0.084	0.001	0.000
Ego Age Squared	-0.001	0.000	0.000
Alter Voted	0.395	0.003	0.000
Ego Ideology	0.187	0.001	0.000
|Ego Ideology|	0.347	0.006	0.000
Ego Female	-0.054	0.003	0.000
Ego Married	0.446	0.004	0.000
Ego College	0.474	0.004	0.000
Ego Friend Count	0.003	0.000	0.000
Intercept	-2.478	0.015	-0.000
N - individuals	397815		
N - dyads	2410097		
Deviance	435239		
Null Deviance	480789		

the relationship between an individual’s (in social network terms, the “ego”) turnout and the difference between the ego’s ideology and that of a friend (the “alter”). As Table B.2 shows, an increase in the ideological distance between friends is associated with lower rates of turnout by the ego.

This result is consistent with previous work showing that disagreement in an individual’s social network is associated with lower turnout rates (Mutz 2002; Huckfeldt, Johnson and Sprague 2004). However, the present work has the advantage of being purely observational. That is, we use validated voting records, use a behavioral measure of ideology, and observe friendships, which help us avoid survey response bias and recall bias for friendships.

3.8 Discussion

This paper makes several important contributions, first it shows a method for measuring ideology using large-scale data from social media, and second to the study of the structure of ideology in society and its possible polarization and its effects on rates of participation. We demonstrate a new method for obtaining estimates of ideology that puts elites and the general public on the same scale. We show that the method produces reliable estimates of ideology that are predictive of other measures of elite ideology, DW-NOMINATE, and individual-level ideology, self-expressed political views. Placement of elites and the general public on the same scale is especially significant not only for the accuracy and validity of this measure, but also for the important contributions to the study of electoral politics and political communication. For instance, one longstanding debate in the literature concerns whether the American public is polarizing (Fiorina, Abrams and Pope 2006; Abramowitz and Saunders 2008) and, if so, whether or not such a divide is driven by elites. Data that put elites and the the general public on the same scale are critical to study of these types of phenomena because they allow for reliable comparison of the ideology of both groups.

We show how this new data can be coupled with the extensive information Facebook has about the nature of our social relationships to investigate how our social networks are structured ideologically. We show that while there is correlation between the ideology between friends, that the correlations between close friends, family, and romantic partners are much stronger. This evidence confirms previous work that showed friends were likely to share similar ideologies and also shows how much more likely shared ideologies are among close friends. While close friends have strongly correlated ideologies, the correlations among family members and romantic partners are stronger still, with married couples having by far the strongest correlation (above 0.92). Further, we show that from 2011 to 2012 there is an increase in the extent to which ideology is associated with the closeness of a friendship, suggesting polarization over the one-year period. This evidence contributes to the debate about polarization by using new evidence about not only the distribution of ideologies in the electorate, but also in social space.

The evidence we show of polarization is notably different than previous evidence in that it uses information about our social ties and the strength of those ties. Prior evidence of polarization typically came from survey responses to questions about ideology, which seek to explain how ideological characteristics are distributed based on some other factor or factors (e.g., partisanship, political engagement, geography). In contrast, we show evidence of polarization based on who we interact with. A better understanding of this type of polarization is critical, as the ideologies of our social contacts can impact the likelihood that we are exposed to new ideas, which is a critical component of democracy (Huckfeldt, Johnson and Sprague 2004). Future work on polarization among the friends that we interact with will be important for understanding the evolution of ideological polarization in the general public.

Finally, we show evidence of the association between disagreement in an individual's social network and decreased rates of turnout. This result is consistent with previous work (Mutz 2002; Huckfeldt, Johnson and Sprague 2004), but has the advantage of being purely observational both in measuring ideology and turnout of individuals and also in observing friendships. This helps to avoid biases that individuals have both in answering survey questions about ideology and turnout and also in recalling friends with whom they have discussed politics. While our results show that individuals in networks with disagreement are less likely to participate in politics, further work should further investigate whether this relationship is causal.

The possibilities for future research using large data sets that contain previously un-studied types of information about people such as Facebook and Twitter should not be underestimated. New measures of ideology allow for opportunities to study phenomena over time and at a scale previously not possible. Previous measures of elite ideology, such as DW-NOMINATE, rely on new votes to update measures of ideology. In fact, in most cases, ideology is measured simply for one term of Congress, and is assumed to be stable throughout the term. New data, such as campaign finance data, allow for a slightly finer grain, as new measures of ideology can be computed quarterly[4]. However, data collected online, such as from Facebook, present the potential to create new measures of ideology down to the second. While in many cases such a fine grained measure

[4]While it is true that measures of ideology using campaign finance records should be able to be updated daily based on the date of a contribution, in practice many campaigns wait until the deadline for a filing period to record some donations, making measures within a filing period subject to bias.

is not necessary, in some applications measures within short periods of time may be very useful. For instance, one may be curious how the perception of a candidate's ideology changes over the course of a debate. In the past, polls or surveys were necessary to see how viewers were reacting to the debate. With measures such as the one described in this paper, we may be able to detect changes in the public perception of a candidate simply by seeing how the ideology of the candidate changes according to the public's changing preferences online throughout the debate.

Aside from the ability to study phenomena in small time steps, the ability to measure the ideology of such a large number of individuals grants us a great deal more statistical power. The increased power that the large-N nature of studies based on data like these affords researchers the opportunity to unobtrusively test theories we previously could not; furthermore it increases precision when we do so.[5]

This research is part of a growing literature in the social sciences in which large sources of data are used to conduct research that was previously not possible (Lazer et al. 2009). We hope that the measure of ideology we use in this paper, and others that measure the ideology of large numbers of the general public (Bonica 2013; Tausanovitch and Warshaw 2012) will contribute to our understanding of ideology in new ways. While there has been a long tradition of research into ideology and its structure, this paper should form a starting point for future research into how our social networks are critical to our understanding of society's ideological makeup.

[5]We caution, though, that with this increase in statistical power that we should be careful to not confuse statistical significance with practical significance. For example, in this paper we find that those who self-describe their political views as 'conservative' are more conservative that those who self-describe as 'very conservative', though the difference between the two groups is statistically significant, it is very small. In this case, the important differences are between groups further apart on the ideological scale.

Appendix A

Supplementary information for “Social Information and Participation”

We use data from the Facebook profiles of users we could match to voting records in order to match them based on characteristics predictive of voting and characteristics of their friends that are predictive of voting. Variables are coded in the following ways:

- Age: user supplied date of birth. All users must input their date of birth when creating an account.
- Gender: user supplied gender. Most users input their gender when creating an account. Those records that did not include gender were removed from the analysis.
- College attendance: users who indicated in their profile that they had graduated from high school prior to 2010 are coded as ‘high school graduate’. Users who indicated in their profile that they had graduated from college prior to 2010 are coded as ‘college graduate’. Users who indicated in their profile that they had graduated from from a post-graduate program prior to 2010 are coded as ‘Graduate Degree’. Users who had not indicated in their profile that they had graduated from any such institutions prior to 2010 are coded as ‘none listed’.
- Relationship status: users who indicated in their ‘relationship status’ that they are ‘engaged’, ‘in a relationship’, ‘single’, ‘it’s complicated’ or ‘married’ are coded as such. All other users are coded as a ‘Not specified’.

- Religious views: users who input a religious view in their profile are coded as ‘1’. All other users are coded as a ‘0’.
- Republican: users whose ‘political views’ included the word “republican” are coded as ‘1’. All other users are coded as a ‘0’.
- Democrat: users whose ‘political views’ included the word “democrat” are coded as ‘1’. All other users are coded as a ‘0’.
- Liberal: users whose ‘political views’ included the word “liberal” are coded as ‘1’. All other users are coded as a ‘0’.
- Conservative: users whose ‘political views’ included the word “conservative” are coded as ‘1’. All other users are coded as a ‘0’.

Table A.1: Comparison of means of clicking on self-reported voting across the two message types by group.

Group	Social Message Mean	Message Mean	Difference of Mean	T statistic	N
Everyone	20.226%	18.141%	2.085%	42.061	60,666,220
Women	21.267%	19.081%	2.186%	32.898	35,989,748
Men	18.875%	16.920%	1.955%	25.846	25,580,374
18-24	13.381%	12.138%	1.244%	16.454	18,724,152
25-29	15.892%	14.840%	1.052%	8.984	7,184,576
30-39	20.167%	18.546	1.620%	15.179	13,286,552
40-49	25.593%	23.338%	2.615%	19.179	9,701,506
50+	31.879%	27.131%	4.748%	33.231	9,730,601
No education listed	16.978%	14.820%	2.158%	32.882	29,483,797
High school graduate	20.563%	18.766%	1.797%	19.128	17,304,196
College graduate	25.850%	23.733%	2.117%	16.870	11,522,434
Graduate degree	30.897%	27.687%	3.210%	10.975	2,355,793
0 close friends	15.687%	13.437%	2.250%	18.537	7,952,882
1-2 close friends	20.213%	17.191%	3.022%	25.645	10,297,109
3-5 close friends	21.216%	18.630%	2.586%	21.791	10,817,185
6-10 close friends	21.549%	19.421%	2.128%	17.388	10,512,566
11-20 close friends	21.395%	19.833%	1.561%	12.466	10,139,957
20+ close friends	20.204%	19.157%	1.047%	8.801	10,946,521
0-20 friends	14.536%	12.408%	2.128%	19.690	9,379,878
21-50 friends	21.626%	18.206%	3.420%	30.557	11,999,052
51-100 friends	22.381%	19.684%	2.697%	24.896	13,519,770
101-200 friends	21.831%	20.350%	1.481%	13.376	13,230,958
201+ friends	19.125%	18.352%	0.773%	7.079	12,536,562
Engaged	15.451%	14.407%	1.044%	3.813	1,660,420
In a relationship	15.586%	14.278%	1.309%	11.176	8,911,749
Single	17.969%	16.235%	1.734%	17.011	13,148,355
Not Specified	17.217%	14.954%	2.263%	25.133	15,794,227
It is complicated	18.209%	15.861%	2.348%	5.190	651,374
Married	26.724%	24.111%	2.614%	27.157	19,818,353

Table A.2: Comparison of means of information seeking across the two message types by group.

Group	Social Message Mean	Message Mean	Difference of Mean	T statistic	N
Everyone	2.423%	2.162%	0.260%	13.9201	60,666,220
Women	2.468%	2.171%	0.297%	12.0691	35,989,748
Men	2.367%	2.165%	0.202%	6.8674	25,580,374
18-24	2.301%	2.194%	0.107%	3.1581	18,724,152
25-29	2.083%	1.996%	0.086%	1.8758	7,184,576
30-39	1.992%	1.879%	0.112%	3.0093	13,286,552
40-49	2.399%	2.051%	0.348%	7.6059	9,701,506
50+	3.590%	2.755%	0.835%	15.8617	9,730,601
No education listed	2.063%	1.784%	0.279%	11.403	29,483,797
High school graduate	2.615%	2.367%	0.248%	6.779	17,304,196
College graduate	2.891%	2.676%	0.216%	4.531	11,522,434
Graduate degree	3.214%	2.862%	0.352%	3.233	2,355,793
0 close friends	1.713%	1.375%	0.337%	8.136	7,952,882
39083 close friends	2.168%	1.718%	0.450%	11.073	10,297,109
39145 close friends	2.320%	1.893%	0.427%	10.276	10,817,185
39242 close friends	2.480%	2.238%	0.242%	5.287	10,512,566
39405 close friends	2.662%	2.503%	0.158%	3.222	10,139,957
20+ friends	3.002%	3.022%	-0.020%	-0.388	10,946,521
0-20 friends	1.696%	1.360%	0.336%	8.846	9,379,878
21-50 friends	2.460%	1.896%	0.564%	14.238	11,999,052
51-100 friends	2.550%	2.276%	0.274%	6.732	13,519,770
101-200 friends	2.596%	2.384%	0.212%	5.061	13,230,958
201+ friends	2.610%	2.653%	-0.043%	-0.953	12,536,562
Engaged	2.037%	2.106%	-0.001%	-0.617	1,660,420
In a relationship	2.225%	2.081%	0.144%	3.010	8,911,749
Single	2.611%	2.40%6	0.206%	4.853	13,148,355
Not Specified	2.119%	1.833%	0.286%	8.452	15,794,227
It is complicated	2.313%	2.092%	0.220%	1.242	651,374
Married	2.660%	2.300%	0.359%	10.643	19,818,353

Table A.3: Comparison of means across the two message types and the control. It is important to note that people rarely self-report political characteristics on their Facebook profile (less than 1%, as shown).

	Social Message		Message		No Message	
Age	34.711	(0.002)	34.703	(0.019)	34.717	(0.019)
Male	41.232%	(0.006%)	41.276%	(0.063%)	41.291%	(0.063%)
Partisan	0.200%	(0.001%)	0.202%	(0.006%)	0.203%	(0.006%)
Ideologue	0.805%	(0.001%)	0.821%	(0.012%)	0.823%	(0.011%)
Liberal	0.384%	(0.001%)	0.388%	(0.008%)	0.394%	(0.008%)
Conservative	0.473%	(0.001%)	0.481%	(0.009%)	0.484%	(0.009%)
Democrat	0.119%	(0.000%)	0.116%	(0.004%)	0.116%	(0.004%)
Republican	0.097%	(0.000%)	0.099%	(0.004%)	0.096%	(0.004%)

Appendix B

Supplementary information for "The Dynamic Spread of Voting"

B.1 Regression results by decile

Table B.1 shows the results of the logistic GEE models post-matching for each decile. In each, the coefficient of interest is the Alter Vote coefficient, which estimates the extent to which an alter's vote influences the individual's vote. We used generalized estimating equation (GEE) procedures to account for multiple observations of the same ego across the two elections and across ego-alter pairings (Liang and Zeger 1986). We assumed an independent working correlation structure for the clusters (Schildcrout and Heagerty 2005).

B.2 Matching to Voting Records

To choose which states to validate, we identified those states that provided (for research purposes) first names, last names, and full birth dates in publicly available voting records. From these, we chose a set that minimized cost per population. The cost of state records varied from \$0 to \$1500 per state. We excluded records from Texas

Table B.1: Results of logistic regression of Alter Vote on ego vote..

Decile	N (dyads)	Intercept	Intercept SE	Alter Vote Coeff	Alter Vote SE
1	2822070	0.399	0.001	0.181	0.001
2	2845937	0.395	0.001	0.175	0.001
3	2830844	0.396	0.001	0.172	0.001
4	2932188	0.397	0.001	0.172	0.001
5	2909098	0.400	0.001	0.172	0.001
6	2806129	0.400	0.001	0.171	0.001
7	2820716	0.404	0.001	0.173	0.001
8	2851846	0.404	0.001	0.173	0.001
9	2984667	0.402	0.000	0.177	0.001
10	2672930	0.397	0.001	0.188	0.001

because they had systematically excluded some individuals from their voting records (specifically, they did not report on the voting behavior of people that had abstained in the four prior elections). The resulting list of states included Arkansas, California, Connecticut, Florida, Kansas, Kentucky, Missouri, Nevada, New Jersey, New York, Oklahoma, Pennsylvania, and Rhode Island. These states account for about 40% of all registered voters in the U.S., and theirrecords yielded 6,338,882 matched observations of voters and abstainers that we could use to compare to treatment categories from the experiment.

About 1 in 3 users were successfully matched to voter records (success depends on many factors, including voting eligibility, rates of registration, and so on). It is important to note that the match rate for our study is lower than the match rates in many other voting studies, in which more than 50% of users are matched. The primary reason for the low match rate is the age distribution of Facebook users; because the population of Facebook users shows positive skew relative to the country in general (i.e., Facebook users are younger), and young people are less likely to be registered voters, we were able to match fewer records. Additionally, as in other studies in which individuals self-enter data2, matches are more difficult due to a lack of consistency in name conventions in the voter file and Facebook (for instance, a voter may be listed as “Lucille” in the voter record and “Lucy” in Facebook). All information was discarded after we finished the data analysis.

In order to match information in Facebook to public voting records, we relied on the “Yahtzee” method (Jones, Bond, Fariss, Settle, Kramer, Marlow and Fowler 2012). This method is a group-level matching procedure that preserves the privacy of individual actions while still allowing statistical analysis to be conducted at the individual level. We matched users to individuals on the registration list in the same state by first name, last name, and date of birth (dropping all instances that had duplicates) and set the level of error in individual assignments to be 5%. This means that a matched user identified as a voter had a 5% chance of being classified as an abstainer, and vice versa.

B.3 Determination of “Close” Friends

We wished to characterize the strength of ties between pairs of Facebook users beyond the mere existence (or not) of a friendship tie. It has been frequently observed that strong ties engage in “media multiplexity.” For example, if two people communicate often by phone, it is likely they also communicate often through email. Boase et al. (2006) summarize their findings by saying, “People who communicate frequently use multiple media to do so. The more contact by one medium, the more contact by others” (p. 23). We used the frequency with which users interacted with each other on Facebook to estimate the overall closeness of their social tie.

We followed the procedure previously studied by Jones, Bond, Fariss, Settle, Kramer, Marlow and Fowler (2012). On Facebook, people can interact by sending messages, uploading and tagging photos, commenting on posts by friends, posting a “like” on another userÕs post in order to show approval, or in a number of other methods. To identify which Facebook friendships represented close ties, we began with the set of friends who interacted with each other at least once during the three months prior to the election. As individuals vary in the degree to which they use the Facebook website, we normed this level of interaction by dividing the total number of interactions with a specific friend by the total number of interactions a user had with all friends. This gives us a measure of the percentage of a userÕs interactions accounted for by each friend (for example, a user may interact 1% of the time with one friend and 20% of the time with another).

We then categorized all friendships in our sample by decile, ranking them from lowest to highest percentage of interactions. Each decile is a subset of the previous decile. For example, decile 5 contains all friends at the 40th percentile of interaction or higher while decile 6 contains all friends at the 50th percentile of interaction or higher, meaning that decile 6 is a subset of decile 5. We validated this measure of tie strength with a survey. We fielded four surveys to Facebook users asking them to name some number of their close friends (1, 3, 5, or 10). Each survey began with the following prompt:

Think of the people with whom you have spent time in your life, friends with whom you have a close relationship. These friends might also be family members, neighbors, coworkers, classmates, and so on.

Who are your closest friends?

We tested the hypothesis that counting interactions would be a good predictor of named closest friends. We constructed a list of closest friends by pairing each survey respondent with the first friend named in response to the prompt. Thus, closest friends were defined as friendships including Person A (the survey-taker) and Person B (the first name generated by the survey-taker when prompted to name his/her closest friends).

The surveys were completed between October 2010 and January 2011. We obtained 1,656 responses. We then counted the number of times respondents interacted with each of their friends over the three months prior to the user taking the survey, and divided that number by the total number of interactions that the user had with all friends over the same three-month period. We split the percentages of interaction into deciles . This is the same procedure we used to create the deciles of interaction for users in the main text.

B.4 Alternate method of analysis

Establishing causality in a purely observational network is difficult, so we show results here from a different method for estimating the effect of a friend's turnout on self-turnout. Researchers have a set of tools that, while not definitive, help to demonstrate that a causal relationship is likely. The following procedures are outlined in Fowler et al. (2011) as the best practices for estimating the causal effect in a large-scale social network. Here, we account for homophily – the extent to which people create friendships with each other based on shared characteristics, such as the propensity to vote. The method requires longitudinal measures of each person's behavior and of their social ties. The statistical model we use takes the form:

$$Y_{t+1}^{ego} = \alpha + \beta_1 Y_t^{ego} + \beta_2 Y_{t+1}^{alter} + \beta_3 Y_t^{alter} + \gamma_{ego}$$

In this equation, the primary coefficient of interest is that of the alter's voting behavior in the current period (β_2), as this is the estimate of the extent to which the voting decision of the alter in the current period affects the voting decision of the ego in the current period. The other variables represent key controls that help establish causality. The inclusion of the ego's voting behavior in the previous election typically eliminates serial correlation in the errors, substantially controls for the ego's genetic endowment, and controls for any stable tendency for the ego to vote. The alter's voting behavior in the previous period helps control for homophily (Carrington, Scott and Wasserman 2005; Fowler and Christakis 2008). We used generalized estimating equation (GEE) procedures to account for multiple observations of the same ego across the two elections and across ego-alter pairings (Liang and Zeger 1986). We assumed an independent working correlation structure for the clusters (Schildcrout and Heagerty 2005).

Table B.2 shows the results of this regression. The results show that when an alter votes in the same election as the ego, the ego is 11.5% (95% CI 11.2% to 11.7%) more likely to vote. The results from the linear model specification suggest a higher level of influence than do the results from matching, but the estimates are fairly consistent.

Table B.2: The effect of friend turnout on ego turnout. Results of logistic regression of ego validated voting in 2010 on ego covariates and alter validated voting. Models were estimated using a generalized estimating equation with clustering on the ego and an independent working covariance structure (Liang and Zeger 1986; Schildcrout and Heagerty 2005). Models with an exchangeable correlation structure yielded poorer fit. Fit statistics show sum of squared deviance between predicted and observed values for the model and a null model with no covariates (Wei 2002).

	Estimate	S.E.	p
Alter Vote 2010	0.109	0.001	0.000
Alter Vote 2008	-0.051	0.001	0.000
Ego Turnout 2008	1.52	0.001	0.000
Ego Age	0.03	0.000	0.000
Ego Female	-0.11	0.001	0.000
Ego Married	0.33	0.001	0.000
Ego Religious	0.11	0.001	0.000
Ego College	0.30	0.001	0.000
Ego Employed	0.130	0.002	0.000
Ego Democrat	0.259	0.003	0.000
Ego Republican	0.638	0.004	0.000
Ego Liberal	0.384	0.003	0.000
Ego Conservative	0.788	0.004	0.000
Intercept	-2.90	0.002	0.000
N - individuals	3,941,219		
N - dyads	105,093,490		
Deviance	2443220		
Null Deviance	3078173		

Appendix C

Supplementary information for "Estimating Ideology using Facebook's 'Like' Data"

C.1 Variable coding

Variables were coded in the following ways:

- *Validated vote*. Respondents who had the same first name, last name, and birthdate as a record in their state's voter file were matched at the group level to allow statistical analysis on the relationship between the treatment and real world behaviour (see below).

- *Married*. Respondents who had listed a marital relationship on their Facebook profile are coded as married and all others are not.

- *College attendance*. Respondents who had listed a college they have attended on their Facebook profile are coded as having attended college and all others are not.

- *Friend count*. A count of the number of friends an individual has who have both a calculated ideology score and a validated voting record.

C.2 Matching to Voting Records

To choose which states to validate, we identified those states that provided (for research purposes) first names, last names, and full birth dates in publicly available voting records. From these, we chose a set that minimized cost per population. The cost of state records varied from $0 to $1500 per state. We excluded records from Texas because they had systematically excluded some individuals from their voting records (specifically, they did not report on the voting behavior of people that had abstained in the four prior elections). The resulting list of states included Arkansas, California, Connecticut, Florida, Kansas, Kentucky, Missouri, Nevada, New Jersey, New York, Oklahoma, Pennsylvania, and Rhode Island. These states account for about 40% of all registered voters in the U.S., and theirrecords yielded 6,338,882 matched observations of voters and abstainers that we could use to compare to treatment categories from the experiment.

About 1 in 3 users were successfully matched to voter records (success depends on many factors, including voting eligibility, rates of registration, and so on). It is important to note that the match rate for our study is lower than the match rates in many other voting studies, in which more than 50% of users are matched. The primary reason for the low match rate is the age distribution of Facebook users; because the population of Facebook users shows positive skew relative to the country in general (i.e., Facebook users are younger), and young people are less likely to be registered voters, we were able to match fewer records. Additionally, as in other studies in which individuals self-enter data2, matches are more difficult due to a lack of consistency in name conventions in the voter file and Facebook (for instance, a voter may be listed as "Lucille" in the voter record and "Lucy" in Facebook). All information was discarded after we finished the data analysis.

In order to match information in Facebook to public voting records, we relied on the "Yahtzee" method (Jones, Bond, Fariss, Settle, Kramer, Marlow and Fowler 2012). This method is a group-level matching procedure that preserves the privacy of individual actions while still allowing statistical analysis to be conducted at the individual level. We matched users to individuals on the registration list in the same state by first name, last name, and date of birth (dropping all instances that had duplicates) and set the level of error in individual assignments to be 5%. This means that a matched user identified as a voter had a 5% chance of being classified as an abstainer, and vice versa.

C.3 Determination of "Close" Friends

We wished to characterize the strength of ties between pairs of Facebook users beyond the mere existence (or not) of a friendship tie. It has been frequently observed that strong ties engage in "media multiplexity." For example, if two people communicate often by phone, it is likely they also communicate often through email. Boase et al. (2006) summarize their findings by saying, "People who communicate frequently use multiple media to do so. The more contact by one medium, the more contact by others" (p. 23). We used the frequency with which users interacted with each other on Facebook to estimate the overall closeness of their social tie.

We followed the procedure previously studied by Jones, Bond, Fariss, Settle, Kramer, Marlow and Fowler (2012). On Facebook, people can interact by sending messages, uploading and tagging photos, commenting on posts by friends, posting a "like" on another user's post in order to show approval, or in a number of other methods. To identify which Facebook friendships represented close ties, we began with the set of friends who interacted with each other at least once during the three months prior to the election. As individuals vary in the degree to which they use the Facebook website, we normed this level of interaction by dividing the total number of interactions with a specific friend by the total number of interactions a user had with all friends. This gives us a measure of the percentage of a user's interactions accounted for by each friend (for example, a user may interact 1% of the time with one friend and 20% of the time with another).

We then categorized all friendships in our sample by decile, ranking them from lowest to highest percentage of interactions. Each decile is a subset of the previous decile. For example, decile 5 contains all friends at the 40th percentile of interaction or higher while decile 6 contains all friends at the 50th percentile of interaction or higher, meaning that decile 6 is a subset of decile 5. We validated this measure of tie strength with a survey. We fielded four surveys to Facebook users asking them to name some number of their close friends (1, 3, 5, or 10). Each survey began with the following prompt:

Think of the people with whom you have spent time in your life, friends with whom you have a close relationship. These friends might also be family members, neighbors, coworkers, classmates, and so on.

Who are your closest friends?

We tested the hypothesis that counting interactions would be a good predictor of named closest friends. We constructed a list of closest friends by pairing each survey respondent with the first friend named in response to the prompt. Thus, closest friends were defined as friendships including Person A (the survey-taker) and Person B (the first name generated by the survey-taker when prompted to name his/her closest friends).

The surveys were completed between October 2010 and January 2011. We obtained 1,656 responses. We then counted the number of times respondents interacted with each of their friends over the three months prior to the user taking the survey, and divided that number by the total number of interactions that the user had with all friends over the same three-month period. We split the percentages of interaction into deciles . This is the same procedure we used to create the deciles of interaction for users in the main text.

C.4 Expertise and estimation of ideology

We hypothesized users who are more involved in politics should provide more accurate data about the ideological positions of the politicians. While we do not have direct measures of the political expertise of the users, we expect that users who like more politicians are more likely to have greater expertise in politics. While greater interest should lead to more reliable estimates, it comes at a cost of fewer data points to use in the process of estimation. To test this hypothesis, we took a subset of the initial matrix in which only users who had liked a specified number of political pages were included. We then re-ran the analysis using only this subset of users and calculated the correlation with DW-NOMINATE's first dimension. We considered a higher correlation with DW-NOMINATE to be evidence that the ideological estimates were better. A visual

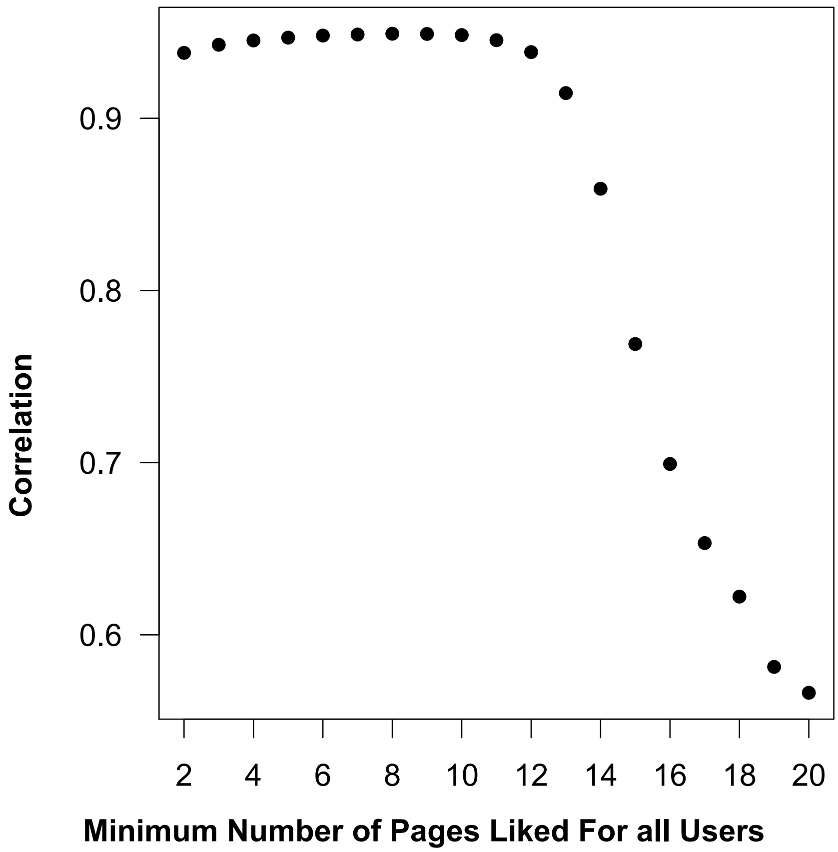

Figure C.1: The correlation of the Facebook measure of ideology with DW-NOMINATE as the minimum number of political pages liked by users increases.

representation of the results of this process can be seen in Figure C.1, and the raw data including the number of users in each category is in Table C.4.

These results show that while there is initially a modest increase in the correlation between the ideology scores we calculate and DWNOMINATE by increasing the expertise that users must have to be included in the data, the difference is minimal (the initial correlation is 0.938, the maximum is 0.949). Increasing the threshold for inclusion eventually causes the correlation to decrease significantly. Indeed, for users who like at least 20 candidates the correlation is only 0.57. This suggests that the gain in efficiency of keeping all users who like at least two candidates outweighs the benefit of selecting a set of users who have liked more candidates.

Table C.1: The correlation of the like measure of ideology with DW-NOMINATE as the minimum number of likes of candidates increases. Note that each group of users is a subset of the previous group of users.

Minimum Number of Pages Liked	Pearson Correlation	Number of Users
2	0.938	6,239,327
3	0.943	2,218,300
4	0.945	1,158,438
5	0.947	715,856
6	0.948	486,865
7	0.949	350,878
8	0.949	263,571
9	0.949	203,461
10	0.948	160,587
11	0.945	128,583
12	0.938	104,366
13	0.915	85,649
14	0.859	70,996
15	0.769	59,352
16	0.699	50,164
17	0.653	42,572
18	0.622	36,389
19	0.581	31,329
20	0.566	26,993

C.5 Ideology Scores

Table C.2 shows the page name and ideology score for all pages for which we estimated an ideology score.

Page name	Ideology Score
Rep John D. Dingell	0.010753819
Congressman John Conyers, Jr.	0.016148247
Charles Rangel	0.018984346
Don Young	-0.010698701
Max Baucus	0.005224771
Chuck Grassley	-0.010927811
Chuck Grassley	-0.01530492
Senator Tom Harkin	0.012373785
George Miller	0.00837573
Rep. George Miller	0.02232428
Henry Waxman	0.008377384
Congressman Ron Paul	-0.006750233
Ron Paul	-0.074275803
Senator Patrick Leahy	0.008379933
Senator Daniel Kahikina Akaka	0.009646644
Norm Dicks for Congress	0.005361049

Norm Dicks	0.009483988
Congressman Dale E. Kildee	0.006832131
Edward J. Markey	0.014713843
U.S. Senator Barbara A. Mikulski	0.009170276
Congressman Nick Rahall	0.003084727
U.S. Senator Richard G. Lugar	-0.010346879
Dan Lungren	-0.010516489
Bill Nelson	0.018405316
Jim Sensenbrenner	-0.010057599
Olympia Snowe for Senate	-0.002173606
Tom Petri	-0.006279777
Carl Levin	0.018989155
Barney Frank	0.047397155
Pat Roberts	-0.006273565
Harold Rogers	-0.007449801
Chuck Schumer	0.031904279
Congressman Frank Wolf	-0.020970446
Senator Ron Wyden	0.019365311
Democratic Whip Steny Hoyer	0.023344533
Jeff Bingaman	0.013211067
Frank Lautenberg	0.010341612
Senator Frank R. Lautenberg	0.010740712
John Kerry	0.0324499
Mitch McConnell	-0.000218878
Senator Mitch McConnell	-0.018979955
Senator Jay Rockefeller	0.004145246
Gary Ackerman	0.007479788
Congressman Gary Ackerman	0.007850617
Congressman Howard Berman	0.010876401
Boucher for Congress	0.005232562
Senator Barbara Boxer	0.007392305
Barbara Boxer	0.065238893
Danny Burton	-0.01037354
Senator Tom Carper	0.00731042
Jim Cooper	0.009614412
Senator Dick Durbin	0.008383192
Rep Sandy Levin	0.00857596
John McCain	-0.117616669
Harry Reid	0.033262239
Senator Harry Reid	0.004426028
Congressman Edolphus Towns	0.007308592
Joe Barton	-0.01995096
Bart Gordon	0.007751484
Congressman Paul E. Kanjorski (PA-11)	0.00423701
Congressman Pete Visclosky	0.003088899
Neil Abercrombie	0.006589149
Ben Cardin	0.01094071
Senator Ben Cardin	0.007802522

Peter DeFazio	0.010539431
Wally Herger	-0.007679636
Senator Jim Inhofe	-0.014791972
Senator Tim Johnson	0.004960523
Jon Kyl	-0.021990269
John Lewis	0.018245401
David Price	0.007837787
Louise Slaughter	0.013745659
Louise Slaughter	0.013274914
Congressman Lamar Smith	-0.015801588
Fred Upton	-0.009210048
Nancy Pelosi	0.08122084
Jerry Costello	0.00403783
Frank Pallone Jr.	0.012418392
Rep. John J. Duncan, Jr.	-0.006733213
Kent Conrad	0.006236054
Eliot Engel	0.009100011
Nita Lowey	0.008275256
Congressman Jim McDermott	0.01374121
Rep Cliff Stearns	-0.013668164
Senator Herb Kohl	0.008402882
Joseph Lieberman	-0.004961326
Congressman Jeff Flake	-0.015810132
Jeff Flake	-0.032069576
John Boozman	-0.01938546
John Boozman	-0.015464007
U.S. Congressman Mike Ross	0.000643378
Mike Honda	0.019944333
Adam Schiff	0.011033582
Darrell Issa	-0.047461818
Congresswoman Susan Davis	0.006360069
Mark Kirk	-0.034610398
Mike Pence	-0.081568275
Representative Stephen F. Lynch	0.002982374
Stephen Lynch for Congress	0.004936828
Mike J. Rogers	-0.011964099
Mike Rogers	-0.006642145
Betty McCollum	0.004588836
Todd Akin	-0.003797302
Congressman Todd Akin	-0.015061152
Sam Graves	-0.008953129
Denny Rehberg for U.S. Senate	-0.004364306
Denny Rehberg, Montana’s Congressman	-0.012616545
Pat Tiberi	-0.007659594
Patrick Tiberi	-0.006970556
John Sullivan	-0.009220128
Congressman Bill Shuster	-0.009002111
Congressman Jim Langevin	0.008877878

Joe Wilson	-0.040633427
Joe Wilson	-0.033342147
Matheson for Congress	0.004537369
Rep. Jim Matheson	0.002700166
Congressman Randy Forbes	-0.016566326
Randy Forbes for Congress	-0.007430811
Eric Cantor	-0.148582194
Congressman Rick Larsen	0.00385468
Rick Larsen	0.007819667
Shelley Moore Capito	-0.010027452
Jo Bonner	-0.010245711
Artur Davis	0.009137604
Congressman Trent Franks	-0.019397499
Raul M. Grijalva	0.021164289
RaIJl Grijalva	0.014426094
Representative Dennis Cardoza	0.008653876
Linda T. Sanchez	0.010850352
Rep. Linda SĞnchez	0.005790933
Phil Gingrey	-0.020551127
Congressman David Scott	0.007419238
David Scott for Congress	0.004984883
Congressman Steve King	-0.032094741
Steve King for Congress	-0.009147289
Congressman Ben Chandler	0.002766292
Ben Chandler	0.006128559
Mike Michaud	0.009948608
Chris Van Hollen	0.017521172
Chris Van Hollen	0.011148974
Candice Miller	-0.010275553
Representative Thaddeus McCotter	-0.035851584
Scott Garrett	-0.014982241
Timothy Bishop	0.010851589
Brad Miller	0.003119776
Mike Turner	-0.009098937
Congressman Tim Ryan	0.015795674
Tim Ryan	0.007848766
Jim Gerlach	-0.010482656
Congressman Jim Gerlach	-0.003415569
Lincoln Davis	0.007488434
Marsha Blackburn	-0.027601587
Jeb Hensarling	-0.022874428
Randy Neugebauer	-0.005869505
Congressman Randy Neugebauer	-0.015837617
Michael Burgess	-0.016091324
Rep. John Carter	-0.012834245
Rep. Rob Bishop	-0.007994652
KENDRICK MEEK	0.030715496
John Salazar	0.007297131

Rep. Connie Mack	-0.013189054
Rep. Debbie Wasserman Schultz	0.021667492
Debbie Wasserman Schultz	0.052194003
Rep. Tom Price	-0.024140734
Congressman Lynn Westmoreland	-0.014421773
Congressman John Barrow	0.006474026
John Barrow	0.007907875
Daniel Lipinski	0.00404177
Congressman Dan Lipinski	0.002850166
Melissa Bean	0.005999779
Charlie Melancon	0.025274289
Charles Boustany Jr	-0.00481451
Congressman Russ Carnahan	0.006625593
Russ Carnahan	0.008453481
Emanuel Cleaver II	0.01496858
Congressman Jeff Fortenberry	-0.011207257
Brian Higgins for Congress	0.004728609
Brian Higgins	0.012252218
Virginia Foxx for Congress	-0.008126409
Rep. Virginia Foxx	-0.006656421
Patrick McHenry	-0.013748326
Congressman Charlie Dent	-0.008876675
Charles Dent	-0.004975372
Louie Gohmert	-0.020183414
Louie Gohmert for Congress	-0.005816165
Ted Poe	-0.015463222
Michael McCaul	-0.014199147
Kenny Marchant	-0.013138828
U.S. Congressman Henry Cuellar (TX-28)	0.003870949
Cathy McMorris Rodgers for Congress	-0.009630051
Congressman Dave Reichert	-0.009590949
Gwen S. Moore	0.013330644
Doris Matsui	0.005720103
Congresswoman Doris O. Matsui	0.012411723
Congressman John Campbell	-0.014930634
Jean Schmidt	-0.005635235
Albio Sires	0.008302881
Gabrielle Giffords	0.051129483
Jerry McNerney	0.013531739
Congressman Jerry McNerney	0.003424746
Kevin McCarthy	-0.02530375
Kevin McCarthy	-0.019518769
Congressman Doug Lamborn	-0.012294477
Ed Perlmutter	0.008615568
Congressman Ed Perlmutter	0.013049367
Joe Courtney	0.007139948
Christopher Murphy	0.013992392
U.S. Rep. Kathy Castor	0.003395935

Congressman Vern Buchanan	-0.015910935
Ron Klein	0.008108086
Congressman Hank Johnson	0.006677059
Rep. Peter Roskam	-0.016772792
Peter Roskam	-0.009822026
Congressman Joe Donnelly	0.003792687
Bruce Braley	0.009860583
Congressman Dave Loebsack	0.01105943
Congressman John Yarmuth (KY-3)	0.004690953
John Sarbanes	0.011248648
Tim Walz	0.008549127
Keith Ellison	0.026260368
Michele Bachmann	-0.202505359
Michele Bachmann	-0.180566098
Dean Heller	-0.01567029
Brett Harrell	-0.000327571
Carol Shea-Porter	0.011580997
U.S. Rep. Yvette D. Clarke	0.007207161
John Hall	0.009287809
John Hall	0.00612377
Kirsten Gillibrand	0.055731377
Heath Shuler	0.008122361
Jim Jordan	-0.027848223
Congressman Zack Space	0.004841878
Mary Fallin	-0.015262073
Jason Altmire for Congress	0.005764704
Jason Altmire	0.00523946
Congressman Tim Murphy (PA-18)	-0.006723787
Congressman Steve Cohen	0.009263378
Bill Foster	0.007595283
Steve Kagen	0.007101876
U.S. Congresswoman Laura Richardson (CA-37)	0.01099215
Congresswoman Niki Tsongas	0.005737196
Rob Wittman for Congress	-0.005072282
Rob Wittman	-0.011430666
Congressman AndrŐ Carson	0.009634742
Steve Scalise	-0.011689168
Congresswoman Jackie Speier	0.019885056
Congresswoman Donna F. Edwards	0.015992927
Ann Kirkpatrick	0.006030622
Congressman Tom McClintock	-0.018616408
Congressman Jared Polis	0.021689532
Congressman Mike Coffman	-0.01241905
Jim Himes	0.009019675
Jim Himes	0.009557141
Alan Grayson	0.054478723
Tom Rooney	-0.01350295
Kosmas for Congress	0.004735337

Debbie Halvorson	0.008856663
Aaron Schock	-0.018081719
Lynn Jenkins for Congress	-0.004162338
Lynn Jenkins	-0.015600635
Brett Guthrie	-0.008595427
Anh "Joseph" Cao	-4.94E-05
Congressman John Fleming	-0.019579409
Bill Cassidy	-0.00533147
Chellie Pingree	0.012471242
Frank Kratovil	0.005896566
Mark Schauer for Congress	0.005756454
Mark Schauer	0.011419335
Gary Peters	0.008817074
Gary Peters	0.006482366
Congressman Erik Paulsen	-0.009291705
Gregg Harper	-0.014712064
John Adler	0.009411891
Congressman Leonard Lance	-0.008112718
Harry Teague	0.006136567
Congressman Harry Teague	0.006240312
Ben Ray Lujan	0.009946823
Congressman Michael E. McMahon	0.006519584
Paul D. Tonko	0.006377157
Dan Maffei	0.006683273
Larry Kissell	0.007984669
Congressman Steve Austria	-0.008125568
Representative Marcia L. Fudge	0.013158413
U.S. Representative Mary Jo Kilroy	0.009459768
Congressman Kurt Schrader	0.005782023
Kurt Schrader	0.007421377
Glenn Thompson	-0.009415739
Dr. Phil Roe	-0.010264237
Pete Olson	-0.014416117
Jason Chaffetz	-0.02524713
Congressman Tom Perriello	0.009612629
Tom Perriello	0.013704901
Congressman Gerry Connolly	0.008682507
Gerry Connolly	0.010908683
Mike Quigley	0.008855633
Judy Chu	0.017396457
Congressman Bill Owens	0.006217831
John Garamendi	0.015855401
Congressman John Garamendi	0.010543957
Congressman Ted Deutch	0.006764567
Congressman Mark S. Critz	0.004406118
Djou for Congress 2010	-0.011767123
Congressman Tom Graves	-0.014742742
Congressman Duncan Hunter	-0.016702137

Duncan Hunter	-0.009040829
Congressman Marlin Stutzman	-0.007330687
Rep. Tom Reed	-0.004136862
Tom Reed	-0.008297239
Maxine Waters	0.013859311
Maxine Waters	0.025352869
Rosa DeLauro	0.01389275
Congressman John Olver	0.009462057
Dave Camp	-0.014218263
Congressman JosŐ E. Serrano	0.01725887
Office of Speaker Boehner	-0.044087103
John Boehner	-0.190666347
Sam Johnson	-0.011385771
Chet Edwards	0.003951965
Bernie Sanders	0.07890138
Congressman Jim Moran	0.014125673
Spencer Bachus	-0.014722184
Congresswoman Lynn Woolsey	0.010842831
Congresswoman Anna Eshoo	0.012168453
Sam Farr	0.012203359
Buck McKeon	-0.010434472
Congressman Xavier Becerra	0.02042416
Rep. Lucille Roybal-Allard	0.00879077
Ed Royce	-0.012232537
Congressman Ken Calvert	-0.008151035
Ken Calvert For Congress	-0.004969863
Congressman Bob Filner	0.007772785
John L. Mica	-0.009310593
Congressman Alcee L. Hastings	0.012369732
Jack Kingston	-0.011651171
Senator Mike Crapo	-0.013915313
Congressman Bobby L. Rush	0.005213555
Congressman Luis V. Gutierrez	0.017857684
Donald Manzullo	-0.011188858
Congressman Roscoe G. Bartlett	-0.012124382
Pete Hoekstra	-0.019684917
Congressman Pete Hoekstra	-0.003223151
Congressman Bennie G. Thompson	0.006210405
Senator Menendez	0.017148998
Peter King	-0.024741318
Congressman Jerry Nadler	0.0113077
Nydia Velazquez	0.016444315
Maurice Hinchey	0.013147131
Frank Lucas	-0.010516561
John Shadegg	-0.010407595
Representative Zoe Lofgren	0.01005254
George Radanovich	-0.003753277
Congressman Brian Bilbray	-0.017155894

Tom Latham	-0.00194237
Senator Roger Wicker	-0.004057531
Frank LoBiondo	-0.006472879
Senator Richard Burr	-0.006440576
Richard Burr	-0.016722069
Sue Myrick	-0.015505817
Congressman Steve LaTourette	-0.00579762
U.S. Representative Mike Doyle	0.006272431
U.S. Senator Lindsey Graham	-0.013443538
Zach Wamp	-0.008181225
Lloyd Doggett	0.010906447
Rep. Lloyd Doggett	0.00770686
Mac Thornberry	-0.011400271
Doc Hastings	-0.011249116
Jesse Jackson Jr.	0.02049041
Representative Elijah E. Cummings	0.01273026
Robert Aderholt	-0.01327477
Congressman Brad Sherman	0.008494435
Representative Loretta Sanchez (CA-47)	0.013234726
Diana DeGette	0.008936562
Congresswoman Diana DeGette	0.015185868
Congressman John Shimkus	-0.015828341
Jerry Moran	-0.010966179
Congressman Jim McGovern	0.007393672
Congressman John Tierney	0.006114975
John Tierney for Congress	0.00887478
Roy Blunt	-0.027150767
Congressman Bill Pascrell, Jr.	0.008761858
Bob Etheridge	0.006970639
U.S. Rep. Bob Etheridge (NC-02)	0.004031148
Congressman Mike McIntyre	0.006693797
Dennis Kucinich	0.057699307
Congressman Joe Pitts	-0.010379538
John Thune	-0.082269136
Pete Sessions	-0.012081106
Pete Sessions	-0.01444266
Kevin Brady	-0.014500281
Kay Granger	-0.010792405
Congressman RubŐn Hinojosa (TX-15)	0.006620023
Congressman Silvestre Reyes	0.009964259
Adam Smith for Congress	0.008025242
Rep. Ron Kind	0.009149311
Lois Capps	0.010074863
Congressman Robert Brady	0.00850357
Barbara Lee	0.028860847
Rep. Gary Miller	-0.006657792
Senator Tom Udall	0.012875182
Mark Udall	0.019088247

John Larson	0.013165792
Mike Simpson	-0.007570434
Jan Schakowsky	0.017829099
Judy Biggert	-0.009844264
Baron Hill	0.009727773
Senator David Vitter	-0.023511199
Representative Michael E. Capuano	0.006282045
Lee Terry	-0.010142209
Shelley Berkley	0.010377635
Anthony Weiner	0.050540941
Greg Walden	-0.005411074
Jim DeMint	-0.150060942
Jay Inslee	0.018200345
Paul Ryan	-0.108888732
Joe Baca	0.016219541
Charles A. Gonzalez	0.008757798
James E. Clyburn	0.012236731
Gene Green	0.007688184
Eddie Bernice Johnson	0.014496283
Congressman Bobby Scott	0.005804094
Ben Nelson	0.009213763
Lisa Murkowski	-0.008588448
Mark Pryor	0.007682357
Senator Lamar Alexander	-0.012264963
John Cornyn	-0.038871082
Amy Klobuchar	0.029189982
Senator Claire McCaskill	0.019431569
Claire McCaskill	0.018502751
Senator Jon Tester	0.00928578
Senator Robert P. Casey, Jr.	0.00839907
Bob Casey	0.016259856
Senator Sheldon Whitehouse	0.005759525
Sheldon Whitehouse	0.015909407
Bob Corker	-0.010610667
Senator Bob Corker	-0.010055188
Jim Webb	0.030824434
John Barrasso	-0.015081537
Mark Begich	0.009494987
Al Franken	0.09165131
U.S. Senator Mike Johanns	-0.012296695
Senator Jeanne Shaheen	0.010007997
Kay Hagan	0.017805849
Senator Jeff Merkley	0.014694235
Mark Warner	0.031253043
Michael F. Bennet	0.009884546
George LeMieux	-0.009440865
Senator Mark Kirk	-0.010216151
Senator Scott Brown	-0.016480855

Scott Brown	-0.113562618
U.S. Senator Chris Coons	0.004592105
Senator Dianne Feinstein	0.006961792
Dianne Feinstein	0.030966792
U.S. Senator Kay Bailey Hutchison	-0.012830265
Kay Bailey Hutchison	-0.017301857
Patty Murray	0.02777798
Jeff Sessions	-0.025060323
Mary Landrieu	0.006775177
Mitch Landrieu	0.00680152
Susan Collins	-0.008244962
Mike Enzi	-0.011571293
Evan Bayh	0.012448227
Rodney Alexander	-0.010239883
Richard Shelby	-0.016888749
Ralph Hall	-0.012296877
Barack Obama	0.382248246
Charles Kennedy	0.004027292
Alison McGovern MP	0.003519396
Beau Biden	0.01835444
Howard Dean	0.067759701
Vaughn L. Reid III	0.002871713
Mario de Marco	0.003458068
Ondej Lika	0.003448781
Lt. Governor Bill Halter	0.006732781
Chad Causey for Congress	0.005213111
Bill Hedrick for Congress	0.007329727
Ted Kennedy	0.057668133
Richard C. Nash	0.001792168
Obama Action Wire	0.030526242
Essam Sharaf	0.003776882
Rob Miller	0.013223734
Shimon Peres	0.002847865
John A. PŐrez	0.011555709
Joe Biden	0.11784103
Daniela SantanchŔ	0.003525538
Women for Obama	0.050633984
Commander Naval Meteorology and Oceanography Command Rear Admiral Jonathan White	0.001910017
Tammy Baldwin	0.025237914
Bronisaw Komorowski	0.003723088
Obama Pride	0.024668643
Bundeskanzler Werner Faymann	0.005197861
Chris Cummins	0.003378655
Tim Kaine	0.029994108
Dan Seals	0.007086148
SONDAGGI & SFIDE TRA PERSONAGGI TV	0.002878179
Health Care Reform	0.03177099
Martin Schulz	0.008145557

FrŐdŐric Recrosio	0.002867919
Brad Smallwood - CIA?	0.002874591
Tarryl Clark	0.021982743
General Martin E. Dempsey	0.001479207
Eva Joly	0.004866637
Bera for Congress	0.00859253
Brad Wall	0.001696296
Chris Clark	0.004252594
Pierluigi Bersani | Pagina Ufficiale	0.007745891
Patrick Murphy	0.019342682
Asian Americans & Pacific Islanders for Obama	0.013127804
Chris Borgia	0.001722676
Tim Geithner	0.002833919
HC Strache	0.004174913
Selahattin Demirta	0.002822318
Mayor Steve Pougnet	0.007284593
Jennifer "Jenna" Austin Wadsworth for Wake County Soil & Water Supervisor	0.005165826
Mark Menzies MP	0.003512506
Kelly Case for Judge of 9th District Court	0.001591643
Begum Khaleda Zia	0.004949145
Chris Larson	0.00833859
Simon Coveney	0.003286624
Margarita Stolbizer	0.003159832
U.S. Senate Democrats	0.013728955
Robin Carnahan	0.029945087
James Cargas for Congress	0.003461711
Karen Bass	0.023149216
Don Bordwell 2012	0.001503873
Supervisor Nate Miley	0.003337901
Heidi Heitkamp	0.005859168
Roger Goodman	0.003814178
Yuli Edelstein	0.000780199
Master Chief Petty Officer of the Navy (MCPON)(SS/SW) Rick D. West	-0.001086686
Eleanor Holmes Norton	0.017317489
Debora Serracchiani	0.006309577
Charlie Boyd IV for Commissioner of Kendall County Precinct No.1	0.001446518
Gavin Newsom	0.063568756
Fernando Pino Solanas	0.002794531
Students for Barack Obama	0.051504895
Suzan DelBene	0.007791146
Secretary Hilda Solis	0.015424147
Pedro Nava	0.007342123
House Budget Committee Democrats	0.008531385
Ambassador Bleich	0.002897719
Latinos for Obama	0.021695578
Paula Brooks	0.003929137
Mark Shurtleff	-0.000856206
Naheed Nenshi	0.003372566

Ways and Means Committee Democrats	0.011617833
Kamala D. Harris	0.041373671
No on Prop 8 | Don't Eliminate Marriage for Anyone	0.046192682
Mike Gravel	0.008616522
Denny Heck for Congress	0.006277672
We vote Mahinda Rajapaksa	0.004140015
Fajardo	0.005934323
Luke Messer	-0.003061597
Disclose, Reform, Amend	0.008748158
Deval Patrick	0.034481289
Brian Schweitzer	0.010673207
Martin Sonneborn	0.002954864
Ray LaHood	0.007729027
Grier Raggio	0.004506001
John Callahan	0.006164426
Robert F. Kennedy	0.038057691
Nils Schmid	0.003095877
Alexi Giannoulias	0.022258625
Dr. Rodolfo Torre CantIJ	0.00352972
Bill English	0.001100259
Gerard Kennedy	0.003052363
Speaker Beth Harwell	0.000264058
Michele Emiliano	0.006142773
Reflexiones de Fidel Castro	0.007085697
Martha Coakley	0.022595925
JosŐ Calzada	0.002890352
Congresswoman Donna M. Christensen	0.00931571
Raj Goyle	0.008429248
Bill Halter for U.S. Senate	0.016848625
(Sergey Mironov)	0.002867636
Friends of Carolyn McCarthy	0.007593101
Davide Faraone	0.003261384
Geert Wilders - PVV	-0.004579983
Marcelo Ebrard	0.003739338
Manan Trivedi	0.005933703
NEY GONZçLEZ SçNCHEZ	0.005551538
Mehdi Karroubi	0.006472176
Roxana Baldetti	0.003501236
Congressmember Karen Bass	0.009439409
Das Williams	0.005680183
Rahm Emanuel	0.040214195
Claudia Roth	0.005409761
Michelle Obama	0.275784983
Student Aid and Fiscal Responsibility Act	0.006499087
Jerry Brown	0.075741183
David Chiu	0.007713957
Manuela Schwesig	0.005373545
Joe Hackney	0.005402436

MALOVA MARIO LOPEZ VALDEZ	0.005139343
Richard J. Codey	0.004762182
Enrique PeŰa Nieto	0.013692725
Jon Corzine - Loretta Weinberg ’09	0.007855812
Julien Dray	0.003421998
Nallari Kiran Kumar Reddy	0.002711515
Nick Clegg	0.011316581
Leonidas Donskis	0.002786481
Tsai Ing-wen	0.00656745
Gary Latanich	0.003518543
Scott Wiener	0.006160301
Mir Hossein Mousavi	0.018668198
Svenja Schulze	0.002935316
Bevan Dufty	0.010035931
Mike Schmier for California Attorney General 2010 - www.VoteMikeAG.com	0.002749479
Ruby Dhalla	0.005842418
ALEJANDRO TOLEDO	0.003602126
Edu Manzano	0.003472351
Oversight Dems	0.004596649
Jean-Paul Huchon	0.00395524
Protect Maine Equality	0.012217651
Fran■ois Bayrou	0.008260676
Michelle Lujan Grisham for Congress	0.003881686
Frank-Walter Steinmeier	0.011061065
John Lynch	0.007226611
Mahinda Rajapaksa	0.005065825
Bobby Shriver	0.00381824
Sleiman Frangieh	0.005321244
Rep. Gus Bilirakis	-0.012376481
Creigh Deeds	0.019125412
Crin Antonescu	0.003664031
U.S. Senator Dean Heller	-0.009369161
Tom White	0.004889885
Rodrigo Medina	0.005897548
Colleen Hanabusa for Congress	0.008280771
House Committee on Natural Resources: Democrats	0.003471624
Jim Mitchell	0.003199687
The Adam Conner	0.00281709
Jens Bullerjahn	0.003313039
Audun Lysbakken	0.003130942
Eruviel çvila	0.004123123
Mayor Edwin M. Lee	0.005483118
George A. Papandreou	0.00687875
Mayor Syed Mustafa Kamal	0.003688612
GermĞn Vargas Lleras	0.005230411
Frank Jensen	0.003170447
Antonio Villaraigosa	0.016012081
Michael J. Rubio	0.005587924

Dalia Grybauskait	0.004391281
Rory Reid	0.009347269
Jean Charest	0.003284631
R.T. Rybak	0.009907234
Gary McDowell	0.005688159
Akbaruddin Owaisi - Youth Icon	0.002608074
Ed Miliband	0.008309547
Francisco De NarvĞez	0.004689696
David Miliband	0.007021295
David Hastings for City Council Ward 1 - Gulfport Florida	0.002603113
Ivo Josipovi	0.004391231
Cedric Richmond	0.007407595
David Luna	0.004589864
Al McAffrey	0.004379785
Laurent Wauquiez	0.003416775
Cem Ězdemir	0.006722646
Jonathan Bannon Maher	0.004424573
Antanas Mockus	0.009165854
Walter Dalton	0.006199827
Volker Bouffier	0.002777625
Bruno GILLES	0.002803953
Monika Hohlmeier	0.002660182
Alberto Torrico	0.013154157
Jane Lubchenco	0.005306725
Chris Kelly	0.019969936
SoumaŢla CissŐ	0.003226395
Brigid Shea	0.003028934
Ted Strickland	0.021009363
Craig Cates for Mayor	0.002565873
Vermont Governor Peter Shumlin	0.006205388
Joel Burns	0.008548523
Miguel Del Sel	0.002641182
Alex Phung for School Board	0.0030781
Rody Duterte	0.004063406
"Alle b£rnene betalte skat, undtagen Helle, hun kunne ikke t¿lle"	0.002661344
Bayani Fernando	0.004727502
StŐphane Dion	0.007136546
Oliver Pocher	0.00383904
Alvaro ArzIJ Irigoyen	0.005191811
California State Assemblymember Mary Hayashi	0.003162097
Nassib Lahoud	0.002699163
Leopoldo Lopez	0.004601198
John Carney	0.006349389
Ambassador Susan Rice	0.009588615
Miguel MĞrquez MĞrquez	0.002643356
Peer Steinbr§ck	0.006587475
Mariastella Gelmini	0.005284223
Premier Anna Bligh	0.003004784

Carla Antonelli	0.003254091
Cory Booker	0.032914295
Peter Kenneth	0.004418938
Congresswoman Frederica Wilson	0.006923405
His Highness Sheikh Mohammed bin Rashid Al Maktoum	0.010947784
Behgjet Pacolli	0.004862198
Fran■ois Hollande	0.009726057
Joe Garcia	0.007492015
Henrique Capriles Radonski	0.004173192
Scott Maddox	0.006732313
Thomas Mulcair	0.003778632
Mayor Sly James	0.00311358
Anthony Foxx	0.005760264
Christos Papoutsis	0.003109785
BINAY 2010	0.004226157
Congresswoman Colleen Hanabusa	0.006043167
Andrew Cuomo	0.032410267
Adam Putnam	-0.012480444
John Kitzhaber	0.012278161
Dominique de Villepin	0.007288652
didier reynders	0.00502003
Joseph Ejercito Estrada	0.005121236
Hiram M. Chittenden Locks	0.002537329
Darrell McGraw, West Virginia Attorney General	0.002880173
Eric Johnson (Texas)	0.003809936
Donald Bourque	0.002476604
Straw for Congress	0.003247396
Congressman Lou Barletta	-0.00718266
Congresswoman Sheila Jackson Lee	0.017232889
Senior Enlisted Leader of the National Guard Bureau	0.001737919
Alcalde de Jun	0.004446501
Gilchrist Olympio	0.003222185
Mike Thompson	0.010620917
Alisha Thomas Morgan	0.005233865
Pedro Passos Coelho	0.00348426
Vincenzo De Luca	0.00458591
Brendan Kelly	0.002641698
Julio Borges	0.002475259
Gordon-Bayani 2010	0.00408133
Congressman David Cicilline	0.006223221
Jens Stoltenberg	0.00766752
Jack Layton	0.012052979
Nana Addo Dankwa Akufo-Addo	0.004727946
Ledama Olekina	0.004390349
FŐlix Lavilla MartŠnez	0.002941481
Michel Daerden	0.005051261
Charlie Justice for Congress	0.004467894
John Key	0.002787981

Martin O'Malley	0.023784092
Juan Carlos Varela	0.002504211
Gustavo Petro	0.006858884
Ted Lieu	0.012279
Rama Yade	0.007941984
Tarek Al-Wazir	0.002915625
People against Sheriff Joe Arpaio	0.020743049
Marie-Anne Montchamp	0.003250316
Jean-Fran■ois CopŐ	0.007781314
Noynoy Aquino (P-Noy)	0.011946089
Gregorio Kilili Camacho Sablan	0.00753411
Lisa P. Jackson	0.01413343
Xavier Trias	0.003005993
Martha Karua	0.007421345
Steven Chu	0.016686359
Katerina Batzeli	0.00253147
Fran■ois Fillon	0.007464932
Victorin LUREL	0.003295108
Stephanie D. Neely	0.004593044
Ken Salazar	0.009777981
Michael B. Hancock	0.00458795
AnŠbal Acevedo VilĞ	0.006281964
Kemal Kõlõ■darolu	0.006296309
Juan Gutierrez	0.003065601
Madeleine Z. Bordallo	0.002882818
MartŠn Sabbatella	0.004087223
Renata Polverini	0.002705683
Eva HŽgl	0.003126858
JORGE RIVAS	0.003829101
SŐrgio Cabral Filho	0.002338498
Mike Bloomberg	0.01576167
Madeira Wine	0.002451217
Rick Perry	-0.10234951
Congressman Robert Hurt	-0.008221386
Fran■ois Rebsamen	0.003217736
Janusz Korwin-Mikke	0.002509299
Pat Quinn	0.011959002
Dora Bakoyannis	0.004809807
Julia Gillard	0.008413735
Ambassador Michael Oren	-0.000614673
Matt Gray for Assembly 2010	0.003961846
Ralph Gonsalves	0.003090193
MARIELLE DE SARNEZ	0.002673942
Helen Zille	0.004958276
Suat Kõlõ■	0.002325137
Richard Blumenthal	0.019353278
Moncef Belkhayat	0.002978862
Benito Mussolini	0.002312575

Cornilles for Congress	-0.004004773
Humberto Lay	0.002360758
Jim Flaherty	0.001023388
Renaud Muselier	0.003137343
Aung San Suu Kyi	0.026841779
Samy Gemayel	0.002596576
Olivier Chastel	0.002973375
Supervisor Josie Gonzales	0.002249353
Idrissa SECK (Officiel)	0.00347298
Senator Ron Johnson	-0.011299808
Sauli NiinistŽ	0.002325526
Joseph Daul	0.004447813
SŐgolŔne Royal	0.010562115
Janez Potonik	0.00230937
Elizabeth May	0.005544001
Pier Ferdinando Casini	0.004262328
Andrea Nahles	0.003763801
Giuliano Pisapia Sindaco X Milano	0.006267065
Justin Trudeau	0.008301493
Pablo Bruera	0.002935004
Laurence Parisot (MEDEF)	0.002468209
Susan Owens For Gilchrist County Clerk of Circuit Court	0.000766105
Franklin Drilon	0.00411715
Manufacturing Matters: Let's Make it in America	0.006121564
Hansen Clarke	0.005759255
Christian Lindner	0.003435839
Guido Westerwelle	0.005765376
State Senator Michael Frerichs	0.003165349
Wanda Hamidah	0.00441246
H§seyin nan	0.003880602
Mikheil Saakashvili	0.003983343
Chris Koster	0.003525426
Nicolas Sarkozy	0.016297451
Enrico Letta	0.003376785
Minister-president	0.002884905
Greg Kerr	0.001052943
Pat Utomi	0.007761144
Bachir Gemayel	0.004430922
CLAMO (Center for Law and Military Operations)	0.002287818
Goodluck Jonathan	0.008885072
Roy Herron	0.004366449
Roberto Maldonado	0.003043983
Esperanza Aguirre	0.005719091
Sal DiDomenico	0.003956065
David Weprin	0.007190928
NC Adjutant General	0.00199159
Hermann GrŽhe	0.003598937
General Colin L. Powell	0.016063211

francesco storace	0.003178653
Chief of Supply Corps	0.001958788
Rep. Terri A. Sewell	0.004590834
Rosy Bindi	0.006401846
Brig. Gen. Glenn H. Curtis	0.001935543
Gregg Harper for Congress	-0.001886095
ValŐrie Hoffenberg	0.002660722
Nasir El-Rufai	0.006798903
Senator John Hoeven	-0.007270412
Walter Veltroni	0.006244999
Department of State - Bureau of Democracy, Human Rights and Labor	0.006176742
Bill Flores	-0.011993506
Jordi Pujol Soley	0.001871656
Roy Moore	-0.010175335
Joining Politics-Why not you and I?	0.002954167
Stephen Fincher	-0.0083348
Rep. Steve Stivers	-0.008038484
Andrew Rohan, Liberal for Smithfield	0.001844668
Charlotte Britz	0.00292026
Rep. Grace F. Napolitano	0.00526528
Kevin Johnson	0.012177628
Re-elect Russell Watts County Treasurer 2014	0.001821942
Matteo Renzi	0.005897498
Trinidad JimŐnez	0.003202687
Citizens for Brackmann	0.001798189
Mohamed ElBaradei	0.008109596
Governor Chris Christie	-0.054940612
Mick Cornett	0.000812464
Chris Steineger	0.004391247
Bill Brady	-0.011430537
Brig. Gen. N. Lee S. Price (P)	0.001765474
Germano Rigotto	0.001763206
Dr. Dietmar Bartsch (MdB)	0.002824484
Air National Guard Director	0.001340692
Sarah Palin	-0.306562007
Air Force Network Integration Center Commander	0.001706459
Dino Rossi	-0.035518905
Sean Ryan	0.001591847
Colonel Scott Peel	0.001635639
GravesforDA	0.001632635
Gary Johnson	-0.022425965
Congressman Joe Walsh	-0.0218389
Jon Huntsman	-0.015039511
Muhyiddin Yassin	0.002866205
Victor Hugo Romo Guerra	0.002816198
Siv Jensen	0.003774405
Jack Bailey for Congress	0.00046384
Gideon Saar	0.001271305

Nichi Vendola	0.012365117
Marco Pannella	0.004351963
Kevin Rudd and Labor	0.005870126
Carlos Torres	0.003893808
Fred Thompson	-0.070312436
Richard Mourdock	-0.017801652
Congressman Mo Brooks	-0.009652246
Nitzan Horowitz	0.004838162
Leung Kwok Hung	0.003951868
Joe Arpaio	-0.095160503
Chris Cox For Congress	-0.002637364
Ray Odierno	-0.002584677
The Adjutant General of Florida	0.001429996
New Hampshire For Jon Huntsman	-0.000147048
Scott Walker for Governor of Wisconsin	-0.021976919
Ujjal Dosanjh	0.003250771
Mike Sanders	0.000562257
Troup County Sheriff's Office	0.00109153
Jean-Marie LE PEN	0.003290614
Marco Rubio	-0.183236033
Orrin Hatch	-0.036437285
Boris Johnson	0.002562131
Juan Soler	0.002786871
Steve Walsh	0.003221046
Antonio Di Pietro	0.009327876
Najib Razak	0.00741667
Citizens for Kirk Dillard	0.001170537
Mike Fitzpatrick for Congress	-0.009265448
Meglena Kuneva	0.002601044
Pedro Zerolo	0.005997704
Mitch Daniels	-0.060259567
Arnaud Montebourg	0.005466646
Congressman Patrick Meehan	-0.003852686
Tom Corbett	-0.012534653
Kurt Burneo	0.002811673
Pauline Marois	0.003306265
The Bill and Ben Party	0.003092751
2008 Republican National Convention	-0.007525584
Abdullah G§l	0.0039515
Governor Tom Corbett	-0.002610329
Adam Robinson	0.000827983
Sam Caligiuri for Congress	-0.005220662
Stephen Harper	-0.000877306
Rob Woodall	-0.001517021
George Allen	-0.041336836
Rob McKenna	-0.008748268
Manny Villar	0.010730547
Pedro Pierluisi - in Congress	0.007516738

Carlos Navarrete Ruiz	0.003862433
Botschafter Philip Murphy	0.005492431
AWANG FERDIAN HIDAYAT	0.003712886
Representative Rick Crawford	-0.006819717
Avigdor Liberman	0.000998873
Josefina VĞzquez Mota	0.007482674
Anders Fogh Rasmussen	0.007438954
Keith Rothfus	-0.003990142
CFKArgentina	0.005986289
Julio Cobos	0.005385668
Josep Anglada	0.002411204
Kay Ivey	-0.003254333
Mohsen Sazegara	0.006477651
Zahra Rahnavard	0.009974157
CŐsar Duarte JĞquez	0.005464014
James Lankford	-0.005155594
Sanford D. Bishop, Jr.	0.005408688
Apirak Kosayodhin	0.004831236
John Loughlin	-0.010086394
Doug Hoffman	-0.026263032
Mark Neumann	-0.011190547
Rep. David Schweikert	-0.011305011
Don Stenberg for U.S. Senate	-0.010409949
Mimoza Kusari	0.004515112
Senator Rand Paul	-0.026464651
Lim Kit Siang	0.00476979
John C. Henderson - Dennis-Yarmouth Regional School Committee Member	0.000988234
Congressman Francisco "Quico" Canseco	-0.01259816
Alev Korun	0.00260149
Mike Clark	-4.87E-05
Derek Schmidt	-0.00114629
George Galloway MP	0.008523406
Patricia Northey, Volusia County Council Member District 5	0.002798179
Congressman Robert J. Dold	-0.005356836
Congressman Cory Gardner	-0.008961245
Jorge Enrique Robledo Castillo	0.004642096
Jim Gibbons	-0.001400447
Vilma Ripoll	0.001885346
Michelle Litjens	-0.001332627
Matt Schultz	-0.002145051
Renato Soru	0.004782879
Governor of Nebraska	-0.002367928
Beatriz Zavala	0.005053925
Debra Medina	-0.011303255
Rand Paul	-0.094939782
Valeriu Zgonea	0.001887593
Paulo Portas	0.003766995
Kristina SchrŽder	0.003663516

Congressman Bob Turner	-0.000837458
Patxi LŮpez	0.004465593
Michela Vittoria Brambilla	0.002928427
McKinley for Congress	-0.00414886
Royal Thai Embassy, Washington D.C.	0.002327095
Recep Tayyip Erdoan	0.007962414
Congressman Tom Cole	-0.015010851
Felip Puig Godes	0.002711009
Ed Scanlan for Governor - scanlanforgov.com	0.002449746
Hishammuddin Hussein	0.004202606
Marco Enriquez-Ominami Los JŮvenes al Poder, Presidente 2010	0.00477488
Rep. Charles Bass	-0.008029682
AngŐlica Araujo Lara	0.006186956
Hugh Jidette	-0.009865272
Pierre Moscovici	0.003712938
RAFFAELE FITTO	0.00249229
Gegen die Jagd auf Karl-Theodor zu Guttenberg	0.003543399
Congressman Mike Fitzpatrick	-0.007198851
Ralph Nader for President 2008	0.005628584
Jaime Nebot	0.004229324
Matt Salmon	-0.002759291
Brice Hortefeux	0.00280461
Pilar del Castillo	0.002615927
Representative Todd Rokita	-0.007223182
Juan Manuel Santos - Presidente	0.00693401
Felipe CalderŮn Hinojosa	0.011511886
Rep. Bill Huizenga	-0.009593032
Carl Person For President 2012	0.001339151
Mike Shirkey - State Rep 65th District	-0.000882763
Greg Abbott	-0.021853822
Congressman Ed Whitfield	-0.007104276
John Neely Kennedy	-0.000430954
Eric BESSON	0.005221394
Congressman Reid Ribble	-0.007738018
Kelly Vincent	0.002527421
Othmar Karas	0.002689252
Owen Hill	0.000719418
Nan Hayworth	-0.011345953
Tim Scott	-0.01354655
Reid Ribble	-0.004932764
Congressman Tim Griffin	-0.010119427
Rep. Larry Bucshon	-0.006819412
Kamla Persad Bissessar	0.005338072
Walter Duke for Dania Beach City Commission	0.002347904
Cayo Lara	0.00352497
Rep. Steve Chabot	-0.004827584
Jaime Herrera	-0.008516228
Benet MaimŠ i Pou	0.002151577

Diane Black	-0.008239153
Rep Chip Cravaack	-0.008514398
Luis Salvador	0.003488556
Congressman Morgan Griffith	-0.007941637
Thorsten SchŁfer-G§mbel	0.004563294
Congressman Steven Palazzo	-0.00684115
Eric Klingemann for US Congress	0.000260789
Lou Barletta	-0.010541148
Lt. Governor Ron Ramsey	-0.006625243
Greg Brower	0.000239925
Jennifer Carroll	-0.010659081
Ann Marie Buerkle for Congress	-0.009040308
Linda Lingle	-0.002588667
Rafael Moreno Valle	0.002803809
Prabowo Subianto	0.004236423
Leonardo Farkas , El candidato 2.0 para una nueva politica en Chile	0.004718531
Angelino Alfano	0.003570867
Congresswoman Vicky Hartzler	-0.005869312
Thomas Tancredo	-0.010001748
Randy Altschuler	-0.006211582
Karl-Theodor zu Guttenberg	0.007873179
Dorothee BŁr, MdB	0.004828671
Dr. Mahathir bin Mohamad	0.010336115
Danny Harder	0.002429661
David Cameron	-0.003988761
Richard Hanna for Congress	-0.005390499
Dona Helena	0.003476911
Congressman Kevin Yoder	-0.007080255
Duane Sand for U.S. Senate	7.37E-05
Juan Ponce Enrile	0.00595482
U.S. Embassy Baghdad	0.004958192
Benjamin Netanyahu	-0.071195771
Michael Grimm	-0.014656664
Senator Loren Legarda	0.00407294
Tony Passalacqua for US Congress	0.000736579
Morten L£kkegaard	0.004168661
Pat Meehan	-0.008986449
Nikola Gruevski	0.002766199
Steve Pearce	-0.009841212
Tzipi Livni	0.005628049
Rep. Frank Guinta	-0.005899657
Christian Wulff	0.004725509
Ray McKinney	-0.005218506
Ted Cruz	-0.035089559
Scott DesJarlais	-0.01008461
Rep. Adam Kinzinger	-0.01135834
A son fiero de esar Veneto	0.003526463
Giorgia Meloni	0.007219528

Karen Diebel	-0.004491268
Steve Kuykendall for Congress, District 47	0.00017712
A Pledge to America	-0.031350514
Xavier Bertrand	0.003634474
Egemen Baõ	0.002552633
MartŠn Sabbatella	0.004673446
General Craig R. McKinley	0.001451729
MaratGuelman	0.002779631
Terry Branstad	-0.008988169
Rik Torfs	0.002648674
Stand with Jackie	-0.006280589
David Shafer	-0.008795379
Glastonbury Police Department	0.001849253
CSM Troy Tyler, The Regimental Command Sergeant Major U.S. Army JAGC	0.002409128
Philipp RŽsler	0.003762039
Congressman Jim Renacci	-0.007520219
Sandy Adams	-0.008045986
Bobby Jindal	-0.096785975
David McKinley	-0.005644067
Mike Haridopolos	-0.006559583
Jeremy Pritchett for Ward 2	0.000447894
Bert Mizusawa for Congress	5.01E-05
Senator Pat Toomey	-0.011209903
Tim Hudak	0.000577312
Mitt Romney	-0.237737945
Cynthia Lummis	-0.006175715
Rep. Bill Flores	-0.007831682
Norm Coleman	-0.011457465
Steve Chabot	-0.009160458
U.S. Army Garrison Hohenfels	0.001611492
Walter Jones	-0.01294173
Phil Bryant	-0.00540846
Koster for Congress	-0.005130296
Angelika Niebler	0.002042827
Heather Wilson	-0.006058302
D.O.Rogozin	0.002398998
Riga	0.002751934
Beth Anne Rankin for US Congress	-0.005086006
House Committee on Homeland Security	-0.005418516
Martha McSally	-0.000603359
SEBASTIAN PIĎERA EL PRESIDENTE DEL BICENTENARIO	0.002507489
Suat Kõlõ■	0.003194199
Vita Quench Water Police	0.001578298
Kevin Brooks	0.00071615
Yvonne Hughes for Warrick County Recorder	0.001542214
Paul LePage, Maine’s Governor	-0.007169596
Carly Fiorina	-0.027746831
Anders for Congress	0.000857371

Demos for Congress	-0.002286749
Chris Widener for U.S. Senate	-0.007374982
Rick Snyder For Michigan	-0.012824468
Dan Branch	-0.000885409
Peter Ramsauer	0.002782083
Gary Herbert	-0.004779644
Ways and Means Committee	-0.017555084
Rob Portman	-0.036732986
Fran■ois SAUVADET	0.00264981
ValŐrie PŐcresse	0.004723641
Tim Pawlenty	-0.106950806
GOP Doctors Caucus	-0.00375945
Bob Turner	-0.016750663
Bobby Schilling	-0.007824203
Scott Rigell	-0.01160835
AndrŐ Rouvoet	-0.000389199
Congressman Jeff Duncan	-0.006039122
Angela Merkel	0.01112347
Michael Pryce MD for U.S. Senate 2012	-0.000180592
Xavier Jaglin	0.002297083
Mel Pennington	0.001057428
Rachel Adato	0.002121283
Congressman Dan Benishek	-0.016356337
Pat Herrity	-0.005934844
Mike Huckabee	-0.165889593
Horst Seehofer	0.002468323
Neil Christensen for City Council	0.001142429
Jon Bruning	-0.027480686
Abhisit Vejjajiva	0.006009815
Silvio Berlusconi	0.005893563
Tzipi Hotovely	0.000818298
Rep. James Lankford	-0.007322983
Bob McDonnell	-0.058846516
wahl.de	0.002386225
Congressman Steve Womack	-0.006173258
John Brunner	-0.006857869
JŐrŹme Peyrat	0.002259203
Quico Canseco for Congress	-0.008820645
Senator Kelly Ayotte	-0.009602083
Earl Sholley	-0.005979499
Allen West	-0.094251276
Justin Amash	-0.032361958
Ariel Sharon	0.001263817
Myers Mermel	-0.01367333
Travis Rose for Public Service Commission	0.000525006
Bob Goodlatte	-0.016634467
Steve Stivers	-0.007090269
BG Patrick Finnegan	0.00139063

LTG Dana Chipman, The Judge Advocate General of the U.S. Army	0.002000582
Matt Doheny	-0.001597378
Renee Ellmers	-0.010544889
Nikki Haley	-0.035403692
Scott Bruun for Congress	-0.003540408
Danny Ayalon	-0.002793933
Tommy G. Thompson	-0.003927567
Roger Williams	-0.014734516
Nate Chesebro	0.000255211
Brian Sandoval	-0.011672735
U.S. House Judiciary Committee	-0.006050625
Congressman Blake Farenthold	-0.006904971
Wilfried Martens	0.004076602
Bill McCollum	-0.012896668
Eric Gregory for US Senate	-0.000297581
Keith Fimian	-0.016691439
Governor Jan Brewer	-0.180274468
Jeff Miller	-0.011894214
Rick Scott	-0.046335733
Kristi Noem	-0.009173585
John Dennis for Congress	-0.035134103
Mariano Rajoy Brey	0.006248905
Representative Martha Roby	-0.010660108
Cathy McMorris Rodgers	-0.020973471
Congressman Sean Duffy	-0.010515003
Randy Brogdon	-0.009653527
Senator Dan Coats	-0.008803442
Congressman Tim Huelskamp (KS-1)	-0.011795764
Mike Fedele	-0.000530162
Bill Johnson	-0.004091337
Richard Hudson for Congress	-0.002295578
House Armed Services Committee Republicans	-0.007941198
Jamie Estrada	0.000499459
Alex Salmond for First Minister	0.001936106
Tom Cotton	-0.002848079
Rick Lazio	-0.010406189
Morgan Griffith for Congress	-0.008781118
Rick Santorum	-0.091976042
Congressman Alan Nunnelee	-0.006741907
Dave Spence	-0.0013232
Troy King for Attorney General	-2.69E-05
Senator Marco Rubio	-0.02944785

Table C.2: Ideology scores for all pages.

Bibliography

Abramowitz, Alan and Kyle Saunders. 2008. "Is Polarization a Myth?" *Journal of Politics* 70(2):542–555.

Achen, Christopher. 1975. "Mass Political Attitudes and the Survey Response." *American Political Science Review* 69(4):1218–1231.

Aleman, Eduardo, Ernesto Calvo, Mark P. Jones and Noah Kaplan. 2009. "Comparing Cosponsorship and Roll-Call Ideal Points." *Legislative Studies Quarterly* 34(1):87–116.

Andolabehere, Stephen and Eitan Hersh. 2012. "Validation: What Big Data Reveal About Survey Misreporting and the Real Electorate." *Political Analysis* 20:43–459.

Aral, Sinan and Dylan Walker. 2012. "Identifying Influential and Susceptible Members of Social Networks." *Science* pp. 337–41.

Aral, Sinan, Lev Muchnik and Arun Sundararajan. 2009. "Distinguishing Influence Based Contagion from Homophily Driven Diffusion in Dynamic Networks." *Proceedings of the National Academy of Sciences* .

Bafumi, Joseph and Michael Herron. 2010. "Leapfrog Representation and Extremism: A Study of American Voters and their Members in Congress." *The American Political Science Review* 104:519–542.

Bailey, Michael. 2007. "Comparable Preference Estimates Across Time and Institutions for the Court, Congress, and Presidency." *American Journal of Political Science* 51:433–448.

Bartels, Larry. 2000. "Partisanship and Voting Behavior, 1952-1996." *American Journal of Political Science* 44(1):117–150.

Bennion, Elizabeth and David Nickerson. 2010. "The Cost of Convenience: An Experiment Showing E-Mail Outreach Decreases Voter Registration." *Political Research Quarterly* .

Berelson, Bernard R., Paul F. Lazarsfeld and William N. McPhee. 1954. *Voting*. Chicago: University of Chicago Press.

Berent, Matthew K., Jon A. Krosnick and Arthur Lupia. 2011s. "The Quality of Government Records and "OverâĂŘestimation" of Registration and Turnout in Surveys: Lessons from the 2008 ANES Panel Study's Registration and Turnout Validation Exercises." Working Paper no. nes012554. Ann Arbor, MI, and Palo Alto, CA: American National Election Studies. Available at http://www.electionstudies.org/resources/papers/nes012554.pdf.

Boase, Jeffrey, John B. Horrigan, Barry Wellman and Lee Rainie. 2006. "The Strength of Internet Ties.".
URL: *http://www.pewinternet.org/Reports/2006/The-Strength-of-Internet-Ties.aspx*

Bond, Robert, Christopher Fariss, Jason Jones, Adam Kramer, Cameron Marlow, Jaime Settle and James Fowler. 2012. "A 61-Million-Person Experiment in Social Influence and Political Mobilization." *Nature* 489:295–298.

Bonica, Adam. 2013. "Ideology and Interests in the Political Marketplace." *American Journal of Political Science* 57(2):294–311.

Brehm, Sharon S. and Jack W. Brehm. 1981. *Psychological Reactance: A Theory of Freedom and Control*. Academic Press.

Brewer, Mark. 2005. "The Rise of Partisanship and the Expansion of Partisan Conflict within the American Electorate." *Political Research Quarterly* 58(2):219–229.

Campbell, Angus, Philip E. Converse, Warren E. Miller and Donald E. Stokes. 1960. *The American Voter*. Univerity of Michigan.

Carrington, Peter J., John Scott and Stanley Wasserman. 2005. *Models and Methods in Social Network Analysis*. Cambridge University Press.

Christakis, Nicholas and James Fowler. 2007. "The Spread of Obesity in a Large Social Network Over 32 Years." *New England Journal of Medicine* 357(4):370–379.

Christakis, Nicholas and James Fowler. 2008. "The Collective Dynamics of Smoking in a Large Social Network." *New England Journal of Medicine* 358(21):2249–58.

Christakis, Nicholas and James Fowler. 2009. "Social Network Visualization in Epidemiology." *Norwegian Journal of Epidemiology* 19(1):5–16.

Cialdini, Robert B. and Melanie R. Trost. 1998. *The handbook of social psychology*. McGraw-Hill chapter Social influence: Social norms, conformity and compliance.

Cialdini, Robert and Noah Goldstein. 2004. "Social Influence: Compliance and Conformity." *Annual Review of Psychology* 55:591–621.

Clinton, Joshua, Simon Jackman and Douglas Rivers. 2004. "The Statistical Analysis of Roll Call Data." *The American Political Science Review* 98(2):355–370.

Converse, Philip. 2006. “The nature of belief systems in mass publics.” *Critical Review* 18(1-3):1–74.

Cornelis, Ilse, Alain Van Hiel, Arne Roets and Malgorzata Kossowska. 2009. “Age Differences in Conservatism: Evidence on the Mediating Effects of Personality and Cognitive Style.” *Journal of Personality* 77(1):51–88.

Enos, Ryan, Anthony Fowler and Lynn Vavreck. 2012. “GOTV and the Participation Gap.” Paper presented at The Midwest Political Science Association annual meeting.

Fiorina, Morris, Samuel Abrams and Jeremy Pope. 2006. *Culture War? The Myth of a Polarized America*. New York: Pearson Longman.

Fowler, James H., Michael T. Heaney, David W. Nickerson, John F. Padgett and Betsy Sinclair. 2011. “Causality in Political Networks.” *American Politics Research* 39(2):437–480.

Fowler, James H. and Nicholas Christakis. 2008. “Estimating Peer Effects on Health in Social Networks.” *Journal of Health Economics* 27(5):1400–1405.

Gerber, Alan and Donald Green. 1999. “Does canvassing increase voter turnout? A field experiment.” *Proceedings of the National Academy of Sciences* 96(19):10939–10942.

Gerber, Alan and Donald Green. 2000. “The effects of canvassing, telephone calls, and direct mail on voter turnout: A field experiment.” *American Political Science Review* 94(3):653–663.

Gerber, Alan and Donald Green. 2001. “Do phone calls increase voter turnout? A field experiment.” *Public Opinion Quarterly* 65(1):75–85.

Gerber, Alan, Donald Green and Christopher Larimer. 2008. “Social Pressure and Voter Turnout: Evidence from a Large-Scale Field Experiment.” *American Political Science Review* 102(1):33–48.

Gerber, Alan, Donald Green and Matthew Green. 2003. “The Effects of Partisan Direct Mail on Voter Turnout.” *Electoral Studies* 22:563–579.

Gerber, Alan, Tiffany Davenport, Christopher Larimer, Christopher Mann and Costas Panagopoulos. 2010. “The Enduring Effects of Social Pressure: Tracking Campaign Experiments Over a Series of Elections.” *Political Behavior* 32(3):423–30.

Gerber, Elisabeth R. and Jeffrey B. Lewis. 2004. “Beyond the Median: Voter Preferences, District Heterogeneity, and Political Representation.” *Journal of Political Economy* 112(6):1364–1383.

Glenn, Norval D. 1974. “Aging and Conservatism.” *The ANNALS of the American Academy of Political and Social Science* 415(1):176–86.

Green, Donald, Alan Gerber and David Nickerson. 2003. "Getting Out the Vote in Local Elections: Results from Six Door-to-Door Canvassing Experiments." *Journal of Politics* 65(4):1083–96.

Green, Donald and Holger Kern. 2010. "Detecting heterogenous treatment effects in large-scale experiments using Bayesian additive regression trees." Paper presented at The Annual Summer Meeting of the Society of Political Methodology.

Hatemi, Peter, John Hibbing, Sarah Medland, Matthew Keller, John Alford, Kevin Smith, Nicholas Martin and Lindon Eaves. 2010. "Not by twins alone: Using the extended family design to investigate genetic influence on political beliefs." *American Journal of Political Science* 54(3):798–814.

Hatemi, Peter, Nathan Gillespie, Lindon Eaves, Brion Maher, Bradley Webb, Andrew Heath, Sarah Medland, David Smyth, Harry Beeby, Scott Gordon, Grant Mongomery, Ghu Zhu, Enda Byrne and Nicholas Martin. 2011. "A Genome-Wide Analysis of Liberal and Conservative Political Attitudes." *Journal of Politics* 73(1):1–15.

Hetherington, Marc. 2001. "Resurgent Mass Partisanship: The Role of Elite Polarization." *American Political Science Review* 95(3):619–631.

Ho, Daniel, Kosuke Imai, Gary King and Elizabeth Stuart. 2007. "Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference." *Political Analysis* 15:199–236.

Hobbs, William R., Nicholas Christakis and James H. Fowler. forthcoming. "Widower Effects in Voter Participation." *American Journal of Political Science* .

Huckfeldt, Robert and John Sprague. 1995. *Citizens, Politics, and Social Communication: Information and Influence in an Election Campaign*. New York: Cambridge University Press.

Huckfeldt, Robert, Paul Johnson and John Sprague. 2004. *Political Disagreement: The Survival of Diverse Opinions Within Communication Networks*. New York: Cambridge University Press.

Imai, Kosuke and Aaron Strauss. 2011. "Estimation of heterogeneous treatment effects from randomized experiments, with application to the optimal planning of the get-out-the-vote campaign." *Political Analysis* 19(1):1–19.

Ithel de Sola Pool, Robert P. Abelson and Samuel Popkin. 1956. *Candidates, Issues and Strategies*. Cambirdge: MIT Press.

Jennings, M. Kent and Richard G. Niemi. 1978. "The Persistence of Political Orientations: An Over-Time Analysis of Two Generations." *British Journal of Political Science* 8(3):333–63.

Jones, Jason J., Jaime E. Settle, Robert M. Bond, Christopher J. Fariss, Cameron Marlow and James H. Fowler. 2012. "Inferring Tie Strength from Online Directed Behavior." *PLoS ONE* 8(1):e52168.

Jones, Jason J., Robert M. Bond, Christopher J. Fariss, Jaime E. Settle, Adam D. I. Kramer, Cameron Marlow and James H. Fowler. 2012. "Yahtzee: An Anonymized Group Level Matching Procedure." *PLoS ONE* 8(2):55760. doi:10.1371/journal.pone.0055760.

Jr, William P Eveland and Myiah Hutchens Hively. 2009. "Political Discussion Frequency, Network Size, and "Heterogeneity" of Discussion as Predictors of Political Knowledge and Participation." *Journal of Communication* 59(2):204–24.

Kenny, Christopher B. 1992. "Political Participation and Effects from the Social Environment." *American Journal of Political Science* 36(1):259–267.

Klofstad, Casey, Rose McDermott and Peter Hatemi. 2012. "The Dating Preferences of Liberals and Conservatives." *Political Behavior* Forthcoming.

Kohut, Andrew, Kim Parker, Scott Keeter, Carroll Doherty and Michael Dimock. 2007. "How Young People View Their Lives, Futures and Politics: A Portrait of "Generation Next".".
URL: *http://www.people-press.org/files/legacy-pdf/300.pdf*

Krasno, Jonathan and Donald Green. 2008. "Do Televised Presidential Ads Increase Voter Turnout? Evidence from a Natural Experiment." *Journal of Politics* 70(1):245–261.

Krosnick, Jon A. and Duane F. Alwin. 1989. "Aging and Susceptibility to Attitude Change." *Journal of Personality and Social Psychology* 57:416–25.

Laver, Michael, Kenneth Benoit and John Garry. 2003. "Extracting Policy Positions from Political Texts Using Words as Data." *The American Political Science Review* 97(2):311–332.

Lazarsfeld, Paul, Bernard Berelson and Hazel Gaudet. 1944. *The People's Choice: How the Voter Makes Up His Mind in a Presidential Campaign*. New York: Columbia University Press.

Lazer, David, Alex Pentland, Lada Adamic, Sinan Aral, Albert Laszlo Barabasi, Devon Brewer, Nicholas Christakis, Noshir Contractor, James Fowler, Myron Gutmann, Tony Jebara, Gary King, Michael Macy, Deb Roy and Marshall Van Alstyne. 2009. "Computational Social Science." *Science* 323:721–723.

Levendusky, Matthew. 2009. *The Partisan Sort: How Liberals Became Democrats and Conservatives Became Republicans*. Chicago: The University of Chicago Press.

Liang, Kug-Yee and Scott L. Zeger. 1986. "Longitudinal Data Analysis Using Generalized Linear Models." *Biometrika* 73(1):13–22.

Malbin, Michael J. 2009. "Small Donors, Large Donors and the Internet: The Case for Public Financing after Obama." Unpublished manuscript.

McClurg, Scott. 2003. "Social Networks and Political Participation: The Role of Social Interaction in Explaining Political Participation." *Political Research Quarterly* 56(4):449–64.

McClurg, Scott. 2004. "Indirect Mobilization: The Social Consequences of PartyContacts in an Election Campaign." *American Politics Research* 32(4):406–43.

McClurg, Scott. 2006. "The Electoral Relevance of Political Talk: Examining Disagreement and Expertise Effects in Social Networks on Political Participation." *American Journal of Political Science* 50(3):737–54.

Monroe, Burt L. and Ko Maeda. 2004. Rhetorical Ideal Point Estimation: Mapping legislative speech. Palo Alto: Stanford University.

Monroe, Burt L., Michael Colaresi and Kevin Quinn. 2009. "Fightin' Words: Lexical Feature Selection and Evaluation for Identifying the Content of Political Conflict." *Political Analysis* 15(4):372–403.

Mutz, Diana. 2002. "The Consequences of Cross-Cutting Networks for Political Participation." *American Journal of Political Science* 46(4):838–55.

Mutz, Diana and Paul S. Martin. 2001. "Facilitating Communication Across Lines of Political Difference: The Role of Mass Media." *American Political Science Review* 95(97-114).

Nickerson, David. 2006. "Volunteer Phone Calls Can Increase Turnout." *American Politics Research* 34(3):271–292.

Nickerson, David. 2007. "The Ineffectiveness of E-vites to Democracy: Field Experiments Testing the Role of E-mail on Voter Turnout." *Social Science Computer Review* 25(4).

Nickerson, David. 2008. "Is Voting Contagious? Evidence from Two Field Experiments." *American Political Science Review* 102:49–57.

Panagopoulos, Costas. 2010. "Affect, social pressure and prosocial motivation: Field experimental evidence of the mobilizing effects of pride, shame and publicizing voting behavior." *Political Behavior* 32(3):369–386.

Poole, Keith. 2005. *Spatial Models of Parliamentary Voting*. New York: Cambridge University Press.

Poole, Keith and Howard Rosenthal. 1997. *Ideology and Congress*. New Brunswick, NJ: Transaction Publishers.

Popkin, Samuel. 1994. *The Reasoning Voter*. Chicago: University of Chicago Press.

Ray, John J. 1985. "What Old People Believe: Age, Sex, and Conservatism." *Political Psychology* 6(3):525–28.

Ringold, Debra Jones. 2002. "Boomerang Effects in Response to Public Health Interventions: Some Unintended Consequences in the Alcoholic Beverage Market." *Journal of Consumer Policy* 25:27–63.

Robert Huckfeldt, Jeanette Morehouse Mendez, Tracy Osborn. 2004. "Disagreement, Ambivalence, and Engagement: The Political Consequences of Heterogeneous Networks." *Political Psychology* 25(1):65–95.

Robinson, John P. 1976. "Interpersonal Influence in Election Campaigns: Two Step-Flow Hypotheses." *Public Opinion Quarterly* 40:304–19.

Rosenstone, Steven and John Mark Hansen. 1993. *Mobilization, Participation, and Democracy in America*. Longman.

Ross, Michael. 1989. "Relation of Implicit Theories to the Construction of Personal Histories." *Psychological Review* 96(2):341–57.

Scheufele, Dietram A., Matthew C. Nisbet, Dominique Brossard and Erik C. Nisbet. 2004. "Social Structure and Citizenship: Examining the Impacts of Social Setting, Network Heterogeneity, and Informational Variables on Political Participation." *Political Communication* 21(3):315–338.

Schildcrout, Jonathan S. and Patrick J Heagerty. 2005. "Regression Analysis of Longitudinal Binary Data with Time-Dependent Environmental Covariates: Bias and Eciency." *Biostatistics* 6(4):633–652.

Shalizi, Cosma and Andrew Thomas. 2011. ""Homophily and Contagion Are Generically Confounded in Observational Social Network Studies." *Sociological Methods and Research* 40:211–39.

Tausanovitch, Chris and Christopher Warshaw. 2012. "Representation in Congress, State Legislatures and Cities." Unpublished manuscript.

Vavreck, Lynn. 2007. "The Exaggerated Effects of Advertising on Turnout: The Dangers of Self-Reports." *Quarterly Journal of Political Science* 2:287–305.

Wei, Pan. 2002. "Goodness-of-fit Tests for GEE with Correlated Binary Data." *Scandinavian Journal of Statistics* 29(1).

Wolfinger, Raymond and Steven Rosenstone. 1980. *Who Votes?* Yale University Press.