doi:10.1038/nature11421

A 61-million-person experiment in social influence and political mobilization

Robert M. Bond[1], Christopher J. Fariss[1], Jason J. Jones[2], Adam D. I. Kramer[3], Cameron Marlow[3], Jaime E. Settle[1] & James H. Fowler[1,4]

Human behaviour is thought to spread through face-to-face social networks, but it is difficult to identify social influence effects in observational studies[9–13], and it is unknown whether online social networks operate in the same way[14–19]. Here we report results from a randomized controlled trial of political mobilization messages delivered to 61 million Facebook users during the 2010 US congressional elections. The results show that the messages directly influenced political self-expression, information seeking and real-world voting behaviour of millions of people. Furthermore, the messages not only influenced the users who received them but also the users' friends, and friends of friends. The effect of social transmission on real-world voting was greater than the direct effect of the messages themselves, and nearly all the transmission occurred between 'close friends' who were more likely to have a face-to-face relationship. These results suggest that strong ties are instrumental for spreading both online and real-world behaviour in human social networks.

Recent experimental studies[6,14–16] have attempted to measure the causal effect of social influence online. At the same time, there is increasing interest in the ability to use online social networks to study and influence real-world behaviour[17–19]. However, online social networks are also made up of many 'weak-tie' relationships[20] that may not facilitate social influence[21], and some studies suggest that online communication may not be an effective medium for influence[22]. An open question is whether online networks, which harness social information from face-to-face networks, can be used effectively to increase the likelihood of behaviour change and social contagion.

One behaviour that has been proposed to spread through networks is the act of voting in national elections. Voter turnout is significantly correlated among friends, family members and co-workers in observational studies[23,24]. Voter mobilization efforts are effective at increasing turnout[25], particularly those conducted face-to-face and those that appeal to social pressure[26] and social identity[27]. There is also evidence from one face-to-face field experiment that voting is 'contagious', in the sense that mobilization can spread from person to person within two-person households[28]. Although anecdotal accounts suggest that online mobilization has made a big difference in recent elections[21], a meta-analysis of email experiments suggests that online appeals to vote are ineffective[24].

Voter mobilization experiments[26–28] have shown that most methods of contacting potential voters have small effects (if any) on turnout rates, ranging from 1% to 10%. However, the ability to reach large populations online means that even small effects could yield behaviour changes for millions of people. Furthermore, as many elections are competitive, these changes could affect electoral outcomes. For example, in the 2000 US presidential election, George Bush beat Al Gore in Florida by 537 votes (less than 0.01% of votes cast in Florida). Had Gore won Florida, he would have won the election.

To test the hypothesis that political behaviour can spread through an online social network, we conducted a randomized controlled trial with all users of at least 18 years of age in the United States who accessed the Facebook website on 2 November 2010, the day of the US 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'. This message encouraged the user to vote, provided a link to find local polling places, showed a clickable button reading 'I Voted', showed a counter indicating how many other Facebook users had previously reported voting, and displayed up to six small randomly selected 'profile pictures' of the user's Facebook friends who had already clicked the I Voted button (Fig. 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.

The design of the experiment allowed us to assess the impact that the treatments had on three user actions; clicking the I Voted button, clicking the polling-place link and voting in the election. Clicking the I Voted button is similar to traditional measures of self-reported voting, but here users reported their vote to their social community rather than to a researcher. We therefore use this action to measure political self-expression, as it is likely to be affected by the extent to which a user desires to be seen as a voter by others. In contrast, social desirability should not affect other user actions in the same way. Clicking the polling-place link took users to a separate website that helped them to find a polling location, and this action was not reported to the user's social community. We therefore use this action to measure a user's desire to seek information about the election. Finally, we used a group-level process to study the validated voting behaviour of 6.3 million users matched to publicly available voter records (see Supplementary Information).

We first analyse direct effects. We cannot compare the treatment groups with the control group to assess the effect of the treatment on self-expression and information seeking, because the control group did not have the option to click an I Voted button or click on a polling-place link. However, we can compare the proportion of users between the two treatment groups to estimate the causal effect of seeing the faces of friends who have identified themselves as voters (Fig. 1). Users who received the social message were 2.08% (s.e.m., 0.05%; *t*-test, $P < 0.01$) more likely to click on the I Voted button than those who received the informational message (20.04% in the social message group versus 17.96% in the informational message group). Users who received the social message were also 0.26% (s.e.m., 0.02%; $P < 0.01$) more likely to click the polling-place information link than users who received the informational message (Fig. 1).

Although acts of political self-expression and information seeking are important in their own right, they do not necessarily guarantee that a particular user will actually vote. As such, we also measured the effect that the experimental treatment had on validated voting, through examination of public voting records. The results show that users

[1]Political Science Department, University of California, San Diego, La Jolla, California 92093, USA. [2]Psychology Department, University of California, San Diego, La Jolla, California 92093, USA. [3]Data Science, Facebook, Inc., Menlo Park, California 94025, USA. [4]Medical Genetics Division, University of California, San Diego, La Jolla, California 92093, USA.

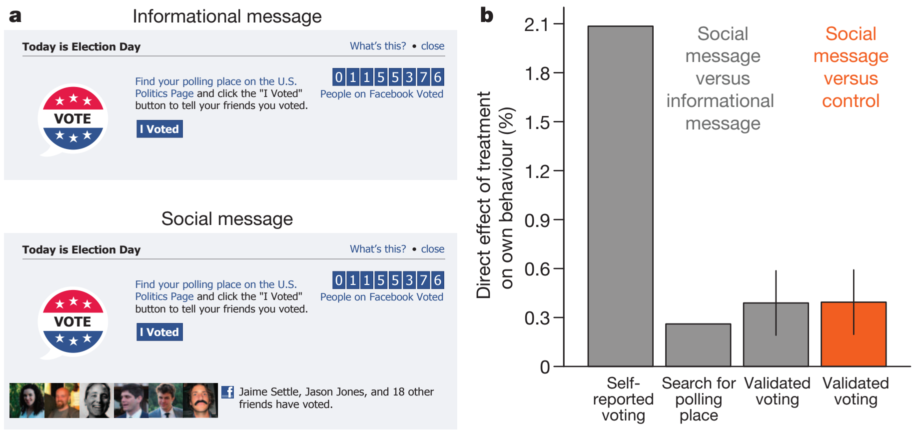

Figure 1 | The experiment and direct effects. a, b, Examples of the informational message and social message Facebook treatments (**a**) and their direct effect on voting behaviour (**b**). Vertical lines indicate s.e.m. (they are too small to be seen for the first two bars).

who received the social message were 0.39% (s.e.m., 0.17%; 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 informational message was 0.39% (s.e.m., 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%, s.e.m., 0.28%; $P = 0.98$), which raises doubts about the effectiveness of information-only appeals to vote in this context.

These results show that online political mobilization can have a direct effect on political self-expression, information seeking and real-world voting behaviour, and that messages including cues from an individual's social network are more effective than information-only appeals. But what about indirect effects that spread from person to person in the social network? Users in our sample had on average 149 Facebook friends, with whom they share social information, although many of these relationships constitute 'weak ties'. Past research indicates that close friends have a stronger behavioural effect on each other than do acquaintances or strangers[9,11,13,21]. We therefore expected mobilization to spread more effectively online through 'strong ties'.

To distinguish users who are likely to have close relationships, we used the degree to which Facebook friends interacted with each other on the site (see Supplementary Information for more detail). Higher levels of interaction indicate that friends are more likely to be physically proximate and suggest a higher level of commitment to the friendship, more positive affect between the friends, and a desire for the friendship to be socially recognized[29]. We counted the number of interactions between each pair of friends and categorized them by decile, ranking them from the lowest to highest percentage of interactions. A validation study (see Supplementary Information) shows that friends in the highest decile are those most likely to be close friends in real life (Fig. 2a).

We then used these categories to estimate the effect of the mobilization message on a user's friends. Random assignment means that any relationship between the message a user receives and a friend's behaviour is not due to shared attributes, as these attributes are not correlated with the treatment (see Supplementary Information). To measure a per-friend treatment effect, we compared behaviour in the friends connected to a user who received the social message to behaviour in the friends connected to a user in the control group. To account for dependencies in the network, we simulate the null distribution using a network permutation method (see the Supplementary Information). Monte Carlo simulations suggest that this method minimizes the risk of false positives and recovers true causal effects without bias (see Supplementary Information).

Figure 2 shows that the observed per-friend treatment effects increase as tie-strength increases. All of the observed treatment effects fall outside the null distribution for expressed vote (Fig. 2b), suggesting that they are significantly different from chance outcomes. For validated vote (Fig. 2c), the observed treatment effect is near zero for weak ties, but it spikes upwards and falls outside the null distribution for the top two deciles. This suggests that strong ties are important for the spread of real-world voting behaviour. Finally, the treatment effect for polling place search gradually increases (Fig. 2d), with several of the effects falling outside the 95% confidence interval of the null distribution.

To simplify the analysis and reporting of results, we arbitrarily define 'close friends' as people who were in the eightieth percentile or higher (decile 9) of frequency of interaction among all friendships in the sample (see the Supplementary Information). 'Friends' are all other Facebook friends who had less interaction. A total of 60,491,898 (98%) users in our sample had at least 1 close friend, with the average user having about 10 close friends (compared with an average of 139 friends who were not close).

The results suggest that users were about 0.011% (95% confidence interval (CI) of null distribution −0.009% to 0.010%) more likely to engage in an act of political self-expression by clicking on the I Voted button than they would have been had their friend seen no message. Similarly, for each close friend who received the social message, an individual was on average 0.099% (null 95% CI –0.042% to 0.048%) more likely to express voting.

We also found an effect in the validated vote sample. For each close friend who received the social message, a user was 0.224% (null 95% CI –0.181% to 0.174%) more likely to vote than they would have been had their close friend received no message. Similarly, for information-seeking behaviour we found that for each close friend who received the social message, a user was 0.012% (null 95% CI –0.012% to 0.012%) more likely to click the link to find their polling place than they would have been had their close friends received no message. In both cases there was no evidence that other friends had an effect (see Supplementary Information). Thus, ordinary Facebook friends may affect online expressive behaviour, but they do not seem to affect private or real-world political behaviours. In contrast, close friends seem to have influenced all three.

The magnitude of these contagion effects are small per friend, but it is important to remember that they result from a single message, and in many cases it was not possible to change the target's behaviour. For example, users may have already voted by absentee ballot before Election Day, or they may have logged in to Facebook too late to vote or to influence other users' voting behaviour. In other words, all effects measured here are intent-to-treat effects rather than treatment-on-treated effects, which would be greater if we had better information about who was eligible to receive the treatment.

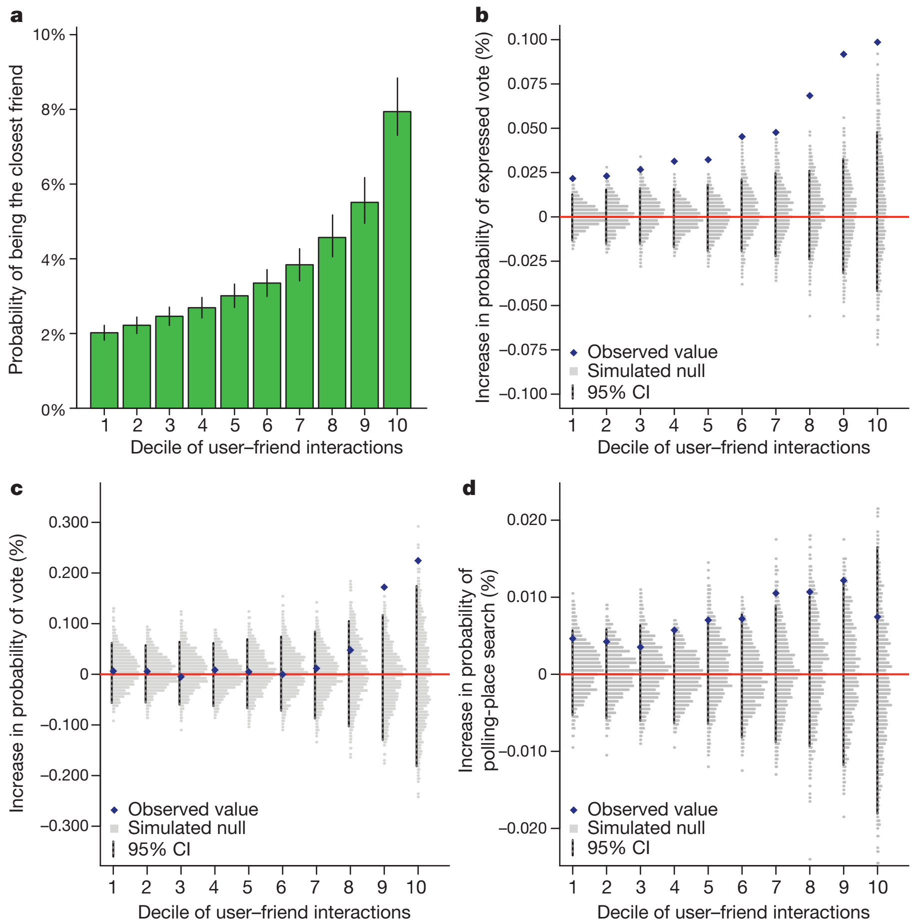

Figure 2 | The effect of mobilization treatment that a friend received on a user's behaviour. a–d, A validation study shows that at increasing levels of interaction, Facebook friends are more likely to have a close real-world relationship (**a**; see also the Supplementary Information). As the interaction increases, so does the observed per-friend effect of friend's treatment on a user's expressed voting (**b**), validated voting (**c**) and polling-place search (**d**). Blue diamonds indicate the observed treatment effect. Horizontal grey bars show the null distribution derived from simulations of identical networks in which the topology and incidence of the behaviour and treatment are the same but the assignments of treatment are randomly reassigned.

Moreover, the scale of the number of users, their friendship connections and the potential voters in a given election is very large. We estimated the per-user effect (the per-friend effect multiplied by the average number of friends per user) and the total effect (the per-user effect multiplied by the total number of users) on the behaviour of everyone in the sample (see Supplementary Information). The results suggest that friends generated an additional 886,000 expressed votes (+1.4%, null 95% CI −1.1% to 1.1%), and close friends generated a further 559,000 votes (+0.9%, null 95% CI –0.3% to 0.3%). In the Supplementary Information we also show that close friends of close friends (2 degrees of separation) generated an additional 1 million expressed votes (+1.7%, null 95% CI –0.8% to 0.9%). Thus, the treatment clearly had a significant impact on political self-expression and how it spread through the network, and even weak ties seem to be relevant to its spread.

However, the effect of the social message on real-world validated vote behaviour and polling-place search was more focused. The results suggest that close friends generated an additional 282,000 validated votes (+1.8%, null 95% CI –1.3% to 1.2%) and an additional 74,000 polling-place searches (+0.1%, null 95% CI –0.1% to 0.1%), but there is no evidence that ordinary friends had any effect on either of these two behaviours. In other words, close friendships accounted for all of the significant contagion of these behaviours, in spite of the fact that they make up only 7% of all friendships on Facebook.

To put these results in context, it is important to note that turnout has been steadily increasing in recent US midterm elections, from 36.3% of the voting age population in 2002 to 37.2% in 2006, and to 37.8% in 2010. Our results suggest that the Facebook social message increased turnout directly by about 60,000 voters and indirectly through social contagion by another 280,000 voters, for a total of 340,000 additional votes. That represents about 0.14% of the voting age population of about 236 million in 2010. However, this estimate does not include the effect of the treatment on Facebook users who were registered to vote but who we could not match because of nicknames, typographical errors, and so on. It would be complex to estimate the number of users on Facebook who are in the voter record but unmatchable, and it is not clear whether treatment effects would be of the same magnitude for these individuals, so we restrict our estimate to the matched group that we were able to sample and observe. This means it is possible that more of the 0.60% growth in turnout between 2006 and 2010 might have been caused by a single message on Facebook.

The results of this study have many implications. First and foremost, online political mobilization works. It induces political self-expression, but it also induces information gathering and real, validated voter turnout. Although previous research suggested that online messages do not work[19], it is possible that conventional sample sizes may not be large enough to detect the modest effect sizes shown here. We also show that social mobilization in online networks is significantly more effective than informational mobilization alone. Showing familiar faces to users can dramatically improve the effectiveness of a mobilization message.

Beyond the direct effects of online mobilization, we show the importance of social influence for effecting behaviour change. Our validation study shows that close friends exerted about four times more influence on the total number of validated voters mobilized than the message itself. These results are similar to those from a prior network simulation study based on observational data that suggested each act of voting on average generates an additional three votes as this behaviour spreads through the network[30]. Thus, efforts to influence behaviour should pay close attention not only to the effect a message will have on those who receive it but also to the likelihood that the message and the behaviour it spurs will spread from person to person through the social network. And, in contrast to the results for close friends, we find that Facebook friends have less effect. Online mobilization works because it primarily spreads through strong-tie networks that probably exist offline but have an online representation. In fact, it is plausible that unobserved face-to-face interactions account for at least some of the social influence that we observed in this experiment.

More broadly, the results suggest that online messages might influence a variety of offline behaviours, and this has implications for our understanding of the role of online social media in society. Experiments are expensive and have limited external validity, but the growing availability of cheap and large-scale online social network data[17] means that these experiments can be easily conducted in the field. If we want to truly understand—and improve—our society, wellbeing and the world around us, it will be important to use these methods to identify which real world behaviours are amenable to online interventions.

Received 17 April; accepted 18 July 2012.

1. Berger, J. & Le Mens, G. How adoption speed affects the abandonment of cultural tastes. *Proc. Natl Acad. Sci. USA* **106,** 8146–8150 (2009).
2. Borgatti, S. P. *et al.* Network analysis in the social sciences. *Science* **323,** 892–895 (2009).
3. Ohtsuki, H., Hauert, C., Lieberman, E. & Nowak, M. A. A simple rule for the evolution of cooperation on graphs and social networks. *Nature* **441,** 502–505 (2006).
4. Onnela, J. P. & Reed-Tsochas, F. Spontaneous emergence of social influence in online systems. *Proc. Natl Acad. Sci. USA* **107,** 18375–18380 (2010).
5. Apicella, C. L., Marlowe, F. W., Fowler, J. H. & Christakis, N. A. Social networks and cooperation in hunter-gatherers. *Nature* **481,** 497–501 (2012).
6. Aral, S. & Walker, D. Identifying influential and susceptible members of social networks. *Science* **337,** 337–341 (2012).
7. Centola, D. An experimental study of homophily in the adoption of health behavior. *Science* **334,** 1269–1272 (2011).
8. Gonzalez-Bailon, S., Borge-Holthofer, J., Rivero, A. & Moreno, Y. The dynamics of protest recruitment through an online network. *Sci. Rep.* **1,** 197 (2011).
9. Christakis, N. A. & Fowler, J. H. Social contagion theory: examining dynamic social networks and human behavior. Preprint at http://arXiv.org/abs/1109.5235v2 (2011).
10. Aral, S., Muchnik, L. & Sundararajan, A. Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *Proc. Natl Acad. Sci. USA* **106,** 21544–21549 (2009).
11. Christakis, N. A. & Fowler, J. H. The spread of obesity in a large social network over 32 years. *N. Engl. J. Med.* **357,** 370–379 (2007).
12. Fowler, J. H. & Christakis, N. A. The dynamic spread of happiness in a large social network. *Br. Med. J.* **337,** a2338 (2008).
13. Christakis, N. A. & Fowler, J. H. The collective dynamics of smoking in a large social network. *N. Engl. J. Med.* **358,** 2249–2258 (2008).
14. Salganik, M. J., Dodds, P. S. & Watts, D. J. Experimental study of inequality and unpredictability in an artificial cultural market. *Science* **311,** 854–856 (2006).
15. Centola, D. The spread of behavior in an online social network experiment. *Science* **329,** 1194–1197 (2010).
16. Fowler, J. H. & Christakis, N. A. Cooperative behavior cascades in human social networks. *Proc. Natl Acad. Sci. USA* **107,** 5334–5338 (2010).
17. Lazer, D. *et al.* Computational social science. *Science* **323,** 721–723 (2009).
18. Ellison, N. B., Steinfield, C. & Lampe, C. The benefits of facebook "friends:" Social capital and college students' use of online social network sites. *J. Comput. Mediat. Commun.* **12,** 1143–1168 (2007).
19. Traud, A. L., Kelsic, E. D., Mucha, P. J. & Porter, M. A. Comparing community structure to characteristics in online collegiate social networks. *SIAM Rev.* **53,** 526–543 (2011).
20. Granovetter, M. The strength of weak ties. *Am. J. Sociol.* **78,** 1360–1380 (1973).
21. Christakis, N. A. & Fowler, J. H. *Connected: The Surprising Power of Our Social Networks and How They Shape Our Lives* (Little, Brown, and Company, 2009).
22. Nickerson, D. W. Does email boost turnout? *Q. J. Polit. Sci.* **2,** 369–379 (2007).
23. Kenny, C. B. Political participation and effects from the social environment. *Am. J. Pol. Sci.* **36,** 259–267 (1992).
24. Huckfeldt, R. & Sprague, J. *Citizens, Politics, and Social Communication: Information and Influence in an Election Campaign* (Cambridge Univ. Press, 1995).
25. Gerber, A. S. & Green, D. P. Does canvassing increase voter turnout? A field experiment. *Proc. Natl Acad. Sci. USA* **96,** 10939–10942 (1999).
26. Gerber, A. S., Green, D. P. & Larimer, C. W. Social pressure and voter turnout: evidence from a large-scale field experiment. *Am. Polit. Sci. Rev.* **102,** 33–48 (2008).
27. Bryan, C. J., Walton, G. M., Rogers, T. & Dwecka, C. S. Motivating voter turnout by invoking the self. *Proc. Natl Acad. Sci. USA* **108,** 12653–12656 (2011).
28. Nickerson, D. W. Is voting contagious? Evidence from two field experiments. *Am. Polit. Sci. Rev.* **102,** 49–57 (2008).
29. Vitak, J. *et al.* It's complicated: Facebook users' political participation in the 2008 election. *Cyberpsychol. Behav. Soc. Netw.* **14,** 107–114 (2011).
30. Fowler, J. H. in *The Social Logic of Politics: Personal Networks as Contexts for Political Behavior* (ed. Zuckerman, A.) 269–287 (Temple Univ. Press, 2005).

Supplementary Information is available in the online version of the paper.

Acknowledgements We are grateful to S. Aral, J. Berger, M. Cebrian, D. Centola, N. Christakis, C. Dawes, L. Gee, D. Green, C. Kam, P. Loewen, P. Mucha, J. P. Onnela, M. Porter, O. Smirnov and C. Volden for comments on early drafts. This work was supported in part by the James S. McDonnell Foundation, and the University of Notre Dame and the John Templeton Foundation as part of the Science of Generosity Initiative.

Author Contributions All authors contributed to study design, data collection, analysis and preparation of the manuscript. J.H.F. secured funding.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.H.F. (jhfowler@ucsd.edu).

doi:10.1038/nature11421

TABLE OF CONTENTS

Research Design

The research design for this study was reviewed and approved by the University of California, San Diego Human Research Protections Program (protocol #101273).

All registered Facebook users over the age of 18 who logged in to their Facebook account on November 2, 2010, were automatically included in the experiment. Random assignment to one of the three treatment groups was done using a random number generator. In total, 61,279,316 Facebook users participated in the study. Most participants (98%) were exposed to the "social message" condition (60,055,176). Half of the remaining participants were exposed to the "informational message" condition (613,096) and the rest were in the control ("no message") group (611,044).

Ideally, we would have designed the experiment with equal sized treatment and control groups to maximize power. However, Facebook wanted to encourage *all* users to participate in the 2010 US Congressional Election, and they therefore asked us to limit the size of the groups that did not receive the standard "get out the vote" (GOTV) message. As a result, 98% of users were exposed to the social message, while only 1% received the alternative informational message and another 1% received no message. Fortunately, the large number of users means there were still over 600,000 people in each of the 1% groups.

When Facebook users log into their account, they are normally greeted with their "News Feed," which includes informational content posted about or by themselves and the friends to whom they are connected. This standard view constituted our control condition. The two treatment conditions included a message from Facebook placed at the beginning of the "News Feed" (see Fig. 1 in the main text). This message provided users a link to information about how to find their polling place. The cumulative total number of Facebook users who had reported voting was shown in the upper right corner. In the middle of the box, users could press a button that read "I Voted" to post a message on their profile that indicated they had voted.

The two treatment conditions were differentiated based on the presence of information about voting behaviour among the user's social network on Facebook. In the "social message" condition, the bottom of the box had small pictures of up to six pictures of the user's friends who had already reported voting. The names of two friends and the total number of the users' friends who had reported voting were included with the pictures. Participants who had six or fewer friends that had voted saw the photos of all of their voting friends; participants who had more than six friends who had voted saw a randomly selected six (the number six was arbitrarily chosen due to space constraints).

Distribution of Key Variables and Balance Testing

Table S1 shows summary statistics for age, sex, and the following variables:

- *Identity as a Partisan*. Respondents can choose to identify their partisanship. Particular party variables (Democrat and Republican) were coded as a 1 when the name of the party appeared in the user's political views and 0 otherwise.

- *Ideology*. Facebook users can write in their political ideology in an open-ended response box. Particular ideology variables (Liberal and Conservative) were coded as a 1 when the ideological label appeared in the user's political views and 0 otherwise.

- *Expressed Voting*. For those respondents in the two treatment conditions, the site recorded when the respondent clicked the “I Voted” button.

- *Polling Place Search*. For those respondents in the two treatment conditions, the site recorded when the respondent clicked the “Find Your Polling Place” link.

- *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).

Table S2 shows balance tests for the demographic variables. There were no significant differences (all pairwise two-tailed t tests indicated $p > 0.05$) between the treatment and control groups on any of these variables, suggesting that random assignment was successful.

Tables S3-S5 show additional balance tests for the demographic variables of friends, close friends, and close friends of close friends. These results show that the user treatment is uncorrelated with the attributes of the people the user is connected to, suggesting that any difference we find between friends of those who received the treatment and friends of those who were in the control group is either due to sampling variation or due to a causal effect of the user treatments on the friends. Figure S1 shows a few of several possibilities for how the user treatment might generate a change in a friend’s or friend’s friend’s behaviour.

This lack of correlation is important because it means that even if people who have more friends who vote are more likely to see a message with many friends, the number of friends shown is not correlated with the treatment either of the user or their friends. Variation in the number of friends shown therefore cannot drive a spurious relationship between treatment and behaviour.

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 birthdates in publicly available voting records. From these, we chose a set that minimized cost per population, but allowed us to detect a 0.5% effect with 80% power given a treatment rate of 98% and a turnout rate of 40% based on rough estimates. 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 behaviour 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 their records 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 GOTV studies, in which more than 50% of users are matched[1]. 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. In Figure S2 we show the distribution of users by age in the 13 states used for matching voters and the match rate by age in those states. Additionally, as in other studies in which individuals self-enter data[2], 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[3]. 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.

For the validated vote results, we assume that the states in which we matched users are a representative sample of all states. Since these states represent about 40% of the U.S. population, we divide the total number of matched users by 0.40 to estimate the total we would have matched if we had acquired voter records in all states, and we use this value for estimating total effects. Note that even with this adjustment, our estimate is probably conservative because it assumes that the treatment effect on unmatched users who are actually in the voter record (as in the Lucille/Lucy example above) is zero.

Overreporting and Underreporting of Voting Behaviour

Since we collected information about both online self-reported voting and real world voting validated by government records, we can compare these two measures to learn more about truthful reporting and the effect the experiment had on it. A comparison of the two measures shows that 3.8% of those in the matched sample self-reported voting when the validated record shows they abstained (an “overreport”), while 50.1% declined to report voting when they actually voted (an “underreport”). The Pearson’s ϕ correlation between the two measures is 0.46 (SE 0.03, $p < 0.01$), which is somewhat lower than the correlation found in most survey research because our self-reported voting measure is not forced-response.

In addition to measuring the effect of the treatment on validated voting (described in the main text), we also analysed the effect of the treatment on overreport and underreport. The results show that users who received the social message were 0.99% (SE 0.14%, $p < 0.01$) more likely to overreport voting and 4.19% (SE 0.27%, $p < 0.01$) less likely to underreport voting than those who received the informational message. Thus, the social message appears to have affected both the desire to vote and the desire to be perceived as a voter.

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.[4] 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.

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 (see Table S3). This is the same procedure we used to create the deciles of interaction for users in the political mobilization experiment.

In Table S6 and Figure 2a of the main text we show the probability that one of the friendships in each decile is the closest friendship identified by the survey-taker. The results show that as the decile of interaction increases, the probability that a friendship is the user's 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.

Tables S7, S8, and S9, and Figures 2b, c, and d in the main text also show that as the number of interactions increases, so does the effect of the user's treatment on his friend's behaviour.

To simplify analysis, we arbitrarily labelled any friend in the 9th decile (80th percentile of above) a "close friend." All other friends are labelled "friends." To measure the person-to-person-to-person effect of a treatment, we labelled people who were not friends or close friends but who shared a close friend in common as "close friends of close friends" (2 degrees of separation). Figure S2 shows the cumulative distribution of the number of users who have a certain number of each of these types of relationships. We therefore studied three mutually exclusive sets of relationships: 5.9 billion "friends," 3 billion "close friends", and 4.6 billion "close friends of close friends" (see Figure S3).

Correlation in Behaviour Between Friends

The town of Abilene, Texas, was selected for illustrative purposes in Figure S4. It features 868 users in the largest connected component of the close friend network who list Abilene, Texas, as their current city in their profile, and shows those who clicked on the "I Voted" button. The graph was generated using the Kamada-Kawai algorithm[5], which is implemented in the igraph library[6] in R[7] and visualized using Pajek[8].

In the whole network, expressed voting is correlated between friends (Pearson's ϕ =0.05) and even more correlated between close friends (ϕ =0.12, see Fig. 3), consistent with other observational studies[9] that found somewhat higher correlations in a smaller set of closer friends. However, these observational associations might result from homophily (the tendency to choose friends who are similar) or exposure to shared environments rather than a process of social influence. Our experimental design allows us to isolate the effect of influence from these alternative explanations since the treatment is uncorrelated with any attribute of the users, their friends, or their friends' friends, and we measure the relationship between *user treatment* and *friend's behaviour*.

Analysis of Direct and Indirect Effects

For direct effects, we used *t* tests to compare the percentage of individuals who exhibited a certain political behaviour (clicking on the "I Voted" button, clicking on the polling place link, or validated vote) in the treatment and control conditions or between the two treatment conditions.

For indirect effects, we want to estimate the relationship between a user's political behaviour (expressed vote, validated vote, and polling place search) and the experimental condition to which their friend was exposed. To estimate the influence that the friend has on the user whose behaviour is being studied, we must be sure that any relationship between the user's behaviour and the friend's experimental condition is not due to chance. Standard techniques like ordinary least squares regression assume independence of observations, which is not the case here due to the complex interdependencies in the network.

To take the network into account, we measure the empirical probability of observing a behaviour by a friend, conditional on a user's treatment (see Figure S1 for examples of how the user treatment might cause changes in a friend's or friend's friend's behaviour). A single user will be connected to many friends, so we conduct this analysis on a *per-friend basis*. For example, looking across all friendships, we may find that 6 of 10 users connected to a friend in the treatment group vote (for a rate of 60%) while only 5 of 10 of those connected to a friend in the control group vote (for a rate of 50%), suggesting a per-friend average treatment effect of 10%.

To compare this observed value to what is possible due to chance, we keep the network topology fixed but randomly permute the assignment to treatment for each user and once again measure the per-friend treatment effect. We repeat this procedure 1,000 times. The simulated values generate a theoretical null distribution we would expect due to chance when there is no treatment effect. We then compare the observed value to the simulated null distribution to evaluate significance. We obtain confidence intervals for the null distribution by sorting the results and taking the appropriate percentiles (in our case, we are interested in the 95% confidence interval, so we use the 25th and 975th values). The random permutation method overcomes the problem of non-independent observations by taking the specific network structure into account when the null distribution is generated.

The results shown in Tables S7, S8, and S9 are for the effect of Social Message vs. Control and they include estimates of the null distribution as described here. Results are also summarized in Figures 2b, c, and d and Figure 3 of the main text.

For each of the three behaviours we studied (expressed vote in Tables S7 and S10, validated vote in Tables S8 and S11, and polling place search in Tables S9 and S12) we used the same procedure. We analysed each friendship in the sample, first calculating the mean rate of behaviour for each user conditional on their friend's experimental condition (see Figure S1 for some examples of how the treatment might affect a user and spread through the network). We then subtracted the rate of behaviour of the users whose friends were in the control condition from the rate of behaviour of the users whose friends were in the treatment condition to calculate the per-friend treatment effect.

We used the permutation procedure to calculate the 95% confidence interval of the null distribution of treatment effects that we would expect due to chance. When the observed effect size falls outside of the confidence interval we consider that result to be statistically significant. We also calculate the average treatment effect *per user* by multiplying the per-friend treatment effect times the average number of friends per user. To calculate the null distribution of the per-user effects we repeat this calculation on each of the simulated networks generated by the permutation procedure.

And finally, for significant treatment effects that fall outside the 95% confidence interval of the null distribution, we calculate the *total effect* by multiplying the per-user effect times the number of users. To calculate the null distribution of the per-user effects, we repeat this calculation on each of the simulated networks generated by the permutation procedure.

Notice that for expressed voting, the treatment effects were strong enough to be detectable at two degrees of separation. For each *close friend of a close friend* who saw the social message, an individual was 0.022% (null 95% CI –0.011% to 0.012%) more likely to express voting. And given the large number of such connections, the number of people affected was also large. We estimate that the per-user effect was +1.7% (null 95% CI –0.8% to 0.9%), which means the treatment caused 1,025,000 close friends of close friends (2 degrees of separation) to express voting.

For validated voting and information seeking we did not find significant effects for close friends of close friends, but it is important to note that these results may be due to limited power since validated voting was measured in a sample one tenth the size, and the direct effects on information seeking were also approximately one-tenth the size as those for expressed voting.

Using the Permutation Method to Calculate Direct Effects

For consistency, we also re-calculated the direct effects using the permutation procedure. For clicking the "I Voted" button and for clicking the polling place link, we calculated the observed difference in means by comparing the social message group and the informational message group. We then randomly re-assigned treatment status to the subjects and re-calculated the difference in means 1,000 times. We then calculated the 95% confidence interval of the null distribution by taking the 25th and 975th values from this simulation. For validated voting, we used the same procedure to simulate the null distribution of the difference in means between the social message group and the informational message group as well as the difference in means between the social message group and the control group.

This procedure provided similar evidence of the statistical significance of the direct effects. For the "I Voted" comparison, we observed a difference of means of 2.08% (NULL CI –0.10% to 0.10%). For clicking on the polling place link we observed a difference of means of 0.26% (NULL CI –0.04% to 0.04%). For validated voting we observed a difference of means between the social message group and the control group of 0.39% (NULL CI –0.39% to 0.37%) and a difference of means between the social message group and the informational message group of 0.39% (NULL CI –0.41% to 0.37%).

Variation in the Treatment Effect by Number of Friends Shown

We were interested in the possibility that the treatment effect varied with the number of friends shown. For the social message group, up to six friends who had previously clicked the "I voted" button were shown in the message as friends who had voted. For users who had less than six friends who had previously reported voting, all previously-voting friends were shown. We recorded the number of friends shown for 1% of the social message group. We then calculated the difference in means of those exposed to the social message with a specified number of friends shown on the initial login to the informational message group for expressed voting and polling place search (see Table S13), and to both the informational message group and the control group for validated voting (see Table S14). The results show no variation in treatment effect for polling place search or validated voting. However, for expressed voting, the treatment effect increases as more faces are shown on the initial login. Note that this pattern shows that the treatment varies in strength depending on how many faces are shown for expressed voting, but does not affect the interpretation of the overall average treatment effect.

Recency of Contact

For indirect effects, we were curious if the *recency* of contact (rather than the *frequency* of contact) could also predict the extent to which friends are influential. We took the same set of interacting friends, but instead of dividing them into deciles by the number of interactions they had in the three months prior to the election, we divided them into deciles based on the number of days prior to the election that the friends had most recently interacted on Facebook. The correlation between the number of interactions and the number of days prior to the election that friends had interacted was 0.25.

Using this measure of the closeness of friendship, Tables S15-S17 show that no groups (other than the full set of interacting friends for expressed vote) had significant indirect effects on any of the three dependent variables (expressed voting, polling place search and validated voting). Users who interacted more recently did not show any signs of larger treatment effects. This suggests that the *quantity* of interaction rather than its *recency* is important for influence.

Average Per-Friend Treatment Effect vs. Number of Friends

As with any experiment that estimates an average treatment effect, our experimental design may obscure important differences in the marginal effect. In our case, the first friend who received the treatment may have a much different effect on the user than the 100th friend. One possibility is that the effect declines with the number of friends, as people pay less attention to a message or behaviour the more often they see it. Another possibility is that the effect increases with the number of friends receiving the treatment, as reinforcement can sometimes induce "complex contagion"[10].

In order to examine differences in the effect of treatment of friends on behaviour we divided the sample of users into deciles by the number of close friends (from lowest to highest) and measured the treatment effect for each decile (see Figure S5). The null distributions are not shown since the observed values in each decile all fall well within the 95% confidence interval (due to dividing the sample by 10). Note that the per-friend effect sizes for vote and expressed vote do tilt slightly downward, but the rate of decrease is probably too small relative to the sampling error in the treatment effect estimates to claim that the difference is significant.

Monte Carlo Tests of the Network Permutation Method

The network permutation method described here has been used in several other publications[11-18]. However, in those applications the goal was to measure the likelihood that a correlation in observed behaviour between connected individuals in the network was due to chance. Here we use the network permutation method to evaluate an observed correlation between a *treatment variable* and a *resulting behaviour* in the treated individual, the treated individual's friends, and the treated individual's friends of friends.

To evaluate whether this procedure yields accurate estimates of causal treatment effects, we have written a computer program in R that 1) generates a network, 2) endows individuals within the network with an initial likelihood of a behaviour, 3) randomly assigns them to treatment and control groups, 4) updates their likelihood of the behaviour according to treatment effects that we can assign (the "true" effects), and then 5) uses these probabilities to determine which individuals exhibit the behaviour. Specifically, we assume

$$y = \alpha + \beta_0 x_0 + \beta_1 x_1 + \beta_2 x_2$$

where y is the total probability a user engages in a specific behaviour (e.g. expressed vote, validated vote, poll search), α is the baseline probability of the behaviour, x_0 is a variable taking the value 1 if the user was in the treatment group and 0 otherwise, and β_0 is the direct effect of the treatment. In addition, x_1 and x_2 are the number of friends at one and two degrees of separation who are in the treatment group, and β_1 and β_2 are the per-friend and per-friend-of-friend effects, respectively.

We can then test our permutation procedure to see whether or not there is bias in the estimated treatment effects and the rate at which our estimation procedure produces false positives. The computation resources necessary to run these Monte Carlo tests necessitated the use of the Gordon super computer[19] at the San Diego Super Computer Center because we simulate a random, 5 million node network 1000 times for each of the scenarios described below.

We tested a variety of scenarios with various parameter combinations:

> **All Scenarios**: We created a network with `n` = 5 million users, with an average of 10 `friends_per_person`, and assumed the control group made up 1% of this population (`control_proportion`). We repeated each scenario 1,000 times, each time setting the "true" effect sizes to be similar to those observed in the real data. Specifically, we let the `direct_effect` of the message on the behaviour be drawn from a uniform distribution with range $\beta_0 \in [0.0\%, 0.8\%]$, and we let the `per_friend_effect` be drawn from a uniform distribution with range $\beta_1 \in [0.00\%, 0.34\%]$ to test effect sizes similar to those

we estimated for validated vote. We also let the `per_friend_of_friend_effect` be drawn from a uniform distribution with range $\beta_2 \in [0.000\%, 0.044\%]$ to test effect sizes similar to those we estimated for expressed vote at two degrees of separation.

Scenario A (baseline): We assumed a Watts–Strogatz (small world) network[20] (`network_type` = 1) and set the rewiring parameter in this model to yield network transitivity of about 0.2 (similar to transitivity in the Facebook close friend network). The initial probability of voting was set to 0.5 for everyone ($\alpha =$ `initial_behaviour_type` = 1). We also assumed the treatment effect was linear in the number of friends ($k =$ `update_behaviour_type` = 1).

Scenario B: Same as scenario A except the initial probability of the behaviour was drawn from a uniform distribution for each user ($\alpha =$ `initial_behaviour_type` = 2). This allows us to test whether heterogeneity in initial behaviour interferes with estimates of treatment effects.

Scenario C: Same as scenario A except the initial probability of the behaviour was assigned to each user based on his or her index id ($\alpha =$ `initial_behaviour_type` = 3). Since users are initially placed on a lattice in order, this causes the initial probability of the behaviour to be very highly correlated between connected users (Pearson's ϕ =0.20 in our simulations, higher than in the observed data). This allows us to test whether homophily on initial behaviour interferes with estimates of treatment effects.

Scenario D: Same as scenario A except we assumed an Erdos-Renyi random network[21] (`network_type` = 2). This allows us to test whether assumptions about the network structure interfere with estimates of treatment effects.

Scenario E: Same as scenario B except we assumed an Erdos-Renyi random network (`network_type` = 2).

Scenario F: Same as scenario C except we assumed an Erdos-Renyi random network (`network_type` = 2).

Scenario G: Same as scenario A except we assumed a Barabasi-Albert "scale free" network[22] (`network_type` = 3). In particular, this kind of network generates extreme skewness in the degree distribution (most users have only a few friends but a small number have very many friends) similar to that observed in the Facebook data. This allows us to test whether assumptions about the network structure interfere with estimates of treatment effects.

Scenario H: Same as scenario B except we assumed a Barabasi-Albert "scale free" network (`network_type` = 3).

Scenario I: Same as scenario C except we assumed a Barabasi-Albert "scale free" network (`network_type` = 3).

The results from all scenarios are presented in Figure S6. There are 9 panels for each treatment effect (the direct effect, the effect on friends, and the effect on friends of friends). Each panel shows results from one of the scenarios described above (labelled by the letter of the scenario), and each point in a plot is one simulation. The dotted line is the theoretical relationship between the "true" values we set and the values estimated by our method one would expect if there were no bias in the procedure, and the solid line is the actual relationship estimated by ordinary linear regression. Notice that in all cases the solid line lies very close to the dotted line.

In Table S18 we report the intercept and slope for each of these lines, and note that all intercepts are near zero (no bias) and all slopes are near one (bias does not emerge as effect sizes increase). These results suggest that our estimates of direct, per-friend, and per-friend-of-friend treatment effects are not overstated.

In Table S19 we report the results of conducting the same analysis for each scenario, but setting the "true" effect sizes to 0 for all 1,000 simulations. In each simulation, we also sample the null distribution 1,000 times and calculate its 95% confidence interval. We then count the number of times our estimate of the treatment effect falls outside this interval, suggesting there is a treatment effect when one does not exist (this is the false positive rate). Notice that all scenarios generate false positive rates of about 5%, consistent with what one would expect at this level of confidence.

Finally, note that in more limited tests with larger networks, it appears that the standard deviation of the estimates decreases with the square root of the number of users in the network. This means that the dispersion in estimates will be lower in the real data than that shown in Figure S6. For our estimates based on 6 million users (validated vote) the variance decreases by a factor of 1.2 and for our estimates based on 60 million users (expressed vote, poll search) it decreases by a factor of 12, and in all cases power to detect effects of the same size as those we observe in the real data exceeds 0.8.

Monte Carlo Code in R

```
# load a network library
library(igraph)

# population size
n <- 1000000

# proportion of population in control group
control_proportion <- 0.01

# initial likelihood of behaviour
# 1 = everyone has 50-50 chance
# 2 = everyone has random uniform chance
# 3 = uniform chance that it is highly correlated between neighbours
initial_behaviour_type <- 1

# network type (1 = watts-strogatz, 2 = erdos-renyi, 3 = barabasi)
network_type <- 1

# friends per person
friends_per_person <- 10

# number of simulations
sims <- 1000

# change in likelihood of behaviour conditional on directly receiving treatment
direct_effect <- 0.008 * runif(1)

# change in likelihood of behaviour conditional on friend receiving treatment
per_friend_effect <- 0.0034 * runif(1)

# change in likelihood of behaviour conditional on friend of friend receiving treatment
per_friend_of_friend_effect <- 0.00044 * runif(1)

# generate initial behaviour
if(initial_behaviour_type == 1) initial_behaviour_probability <- 0.5
if(initial_behaviour_type == 2) initial_behaviour_probability <- runif(n)
if(initial_behaviour_type == 3) initial_behaviour_probability <- 1:n/n

# generate random network ties
if(network_type == 1) g <- simplify(watts.strogatz.game(dim = 1, size = n,
                              nei = round(friends_per_person / 2), p=0.17))
if(network_type == 2) g <- simplify(erdos.renyi.game(n = n, p.or.m = friends_per_person / n))
if(network_type == 3) g <- simplify(barabasi.game(n = n, m = friends_per_person / 2,
                              directed=F, power=0.5))

# create list of friends
friendlist <- sapply(0:(n-1), function(x) neighbors(g,x))
friends <- matrix(NA, nrow = ecount(g) * 2, ncol=2)
friendcount <- sapply(friendlist,length)
friends <- cbind(rep(c(0:(n-1)),friendcount), unlist(friendlist))
friends <- friends[which(!is.na(friends[,1])),]+1

# create list of friends' friends
friendsfriendslist <- sapply(0:(n-1), function(x) neighbors(g, neighbors(g,x)))
friendsfriends <- matrix(NA, nrow = dim(friends)[1] * friends_per_person * 2, ncol=2)
friendsfriendscount <- sapply(friendsfriendslist,length)
friendsfriends <- cbind(rep(c(0:(n-1)), friendsfriendscount), unlist(friendsfriendslist))
friendsfriends <- friendsfriends[which(!is.na(friendsfriends[,1])),]
friendsfriends <- friendsfriends[which(friendsfriends[,1]!=friendsfriends[,2]),]+1

# randomly assign treatment (1) and control (0)
treatment <- rep(1,n)
treatment[sample(n,round(n * control_proportion))] <- 0
```

```
# count friend treatments
friends_treated <- table(factor(friends[,1][which(treatment[friends[,2]] == 1)],
                      levels=1:n))
friendsfriends_treated <-
       table(factor(friendsfriends[,1][which(treatment[friendsfriends[,2]] == 1)], levels=1:n))

# update behaviour probabilities conditional on treatment
behaviour_probability <- initial_behaviour_probability +
       direct_effect * treatment +
       per_friend_effect * friends_treated +
       per_friend_of_friend_effect * friendsfriends_treated

# generate behaviour
behaviour <- as.numeric(runif(n) < behaviour_probability)

# measure observed treatment effects
obs_self_effect <- mean(behaviour[which(treatment == 1)])-
                    mean(behaviour[which(treatment == 0)])
obs_per_friend_effect <- mean(behaviour[friends[,1]][which(treatment[friends[,2]] == 1)])-
                              mean(behaviour[friends[,1]][which(treatment[friends[,2]] == 0)])
obs_per_friend_of_friend_effect <-
       mean(behaviour[friendsfriends[,1]][which(treatment[friendsfriends[,2]] == 1)])-
       mean(behaviour[friendsfriends[,1]][which(treatment[friendsfriends[,2]] == 0)])

# sample the null
random_self_effect <- random_per_friend_effect <- random_per_friend_of_friend_effect <-
       rep(NA,sims)
for(j in 1:sims) {
       treatment <- sample(treatment)
       random_self_effect[j] <- mean(behaviour[which(treatment == 1)])-
                                mean(behaviour[which(treatment == 0)])
       random_per_friend_effect[j] <-
              mean(behaviour[friends[,1]][which(treatment[friends[,2]] == 1)])-
               mean(behaviour[friends[,1]][which(treatment[friends[,2]] == 0)])
       random_per_friend_of_friend_effect[j] <-
              mean(behaviour[friendsfriends[,1]][which(treatment[friendsfriends[,2]] == 1)])-
               mean(behaviour[friendsfriends[,1]][which(treatment[friendsfriends[,2]] == 0)])
}
```

Tables

	Mean	Min	Max
Age	34.7 (SD 14.8)	18	110
Male	41.3%	0	1
Partisan	0.2%	0	1
Ideologue	0.8%	0	1
Liberal	0.4%	0	1
Conservative	0.5%	0	1
Democrat	0.1%	0	1
Republican	0.1%	0	1
Self-Reported Vote	20.0%	0	1
Polling Place Search	2.4%	0	1
Validated Vote	50.8%	0	1

Table S1. Summary statistics for 61 million Facebook users who logged in on Election Day.

	Social Message		Message		No Message	
Age	34.894	(0.003)	34.907	(0.032)	34.904	(0.032)
Female	58.145%	(0.011%)	58.187 %	(0.106%)	58.255%	(0.106%)
Partisan	0.198%	(0.001%)	0.193%	(0.009%)	0.197%	(0.009%)
Ideologue	0.730%	(0.002%)	0.714%	(0.018%)	0.764%	(0.019%)
Liberal	0.381%	(0.001%)	0.355%	(0.013%)	0.410%	(0.014%)
Conservative	0.397%	(0.001%)	0.410%	(0.014%)	0.413%	(0.014%)
Democrat	0.122%	(0.001%)	0.108%	(0.007%)	0.122%	(0.008%)
Republican	0.088%	(0.001%)	0.099%	(0.007%)	0.088%	(0.006%)

Table S2. Comparison of means across the two message types and the control. Here we show the mean and the standard error. It is important to note that people rarely self-report political characteristics on their Facebook profile (less than 2%, as shown).

	Social Message		Message		No Message	
Age	34.811	(14.513)	34.788	(14.506)	34.820	(14.510)
Female	61.565%	(51.494%)	61.571%	(51.480%)	61.563%	(51.501%)
Partisan	0.221%	(3.319%)	0.220%	(3.313%)	0.219%	(3.302%)
Ideologue	0.898%	(6.677%)	0.893%	(6.659%)	0.897%	(6.671%)
Liberal	0.399%	(6.305%)	0.397%	(6.292%)	0.397%	(6.290%)
Conservative	0.499%	(7.049%)	0.496%	(7.026%)	0.500%	(7.051%)
Democrat	0.120%	(3.464%)	0.119%	(3.451%)	0.119%	(3.445%)
Republican	0.101%	(3.172%)	0.101%	(3.174%)	0.100%	(3.155%)
Number of dyads	599,832,198		6,119,130		6,123,174	
Number of users	60,055,176		611,044		613,096	

Table S3. Comparison of means of close friend attributes across the ego's two message types and the control group. Here we show the mean and the standard deviation. It is important to note that people rarely self-report political characteristics on their Facebook profile (less than 2%, as shown).

	Social Message		Message		No Message	
Age	29.829	(12.752)	29.815	(12.748)	29.829	(12.752)
Female	59.414%	(52.133%)	59.374%	(52.135%)	59.382%	(52.139%)
Partisan	0.207%	(3.153%)	0.206%	(3.150%)	0.205%	(3.149%)
Ideologue	0.866%	(4.819%)	0.863%	(4.811%)	0.865%	(4.811%)
Liberal	0.410%	(6.388%)	0.408%	(6.376%)	0.408%	(6.377%)
Conservative	0.456%	(6.741%)	0.455%	(6.730%)	0.457%	(6.744%)
Democrat	0.106%	(3.464%)	0.105%	(3.451%)	0.105%	(3.445%)
Republican	0.101%	(3.250%)	0.101%	(3.247%)	0.100%	(3.246%)
Number of dyads	8,890,938,491		90,886,141		91,017,926	
Number of users	60,055,176		611,044		613,096	

Table S4. Comparison of means of friend attributes across the ego's two message types and the control group. Here we show the mean and the standard deviation. It is important to note that people rarely self-report political characteristics on their Facebook profile (less than 2%, as shown).

	Social Message		Message		No Message	
Age	35.907	(14.502)	35.890	(14.511)	35.920	(14.514)
Female	61.962%	(51.496%)	61.968%	(51.425%)	61.911%	(51.449%)
Partisan	0.219%	(3.303%)	0.217%	(3.292%)	0.220%	(3.311%)
Ideologue	0.917%	(6.743%)	0.940%	(6.829%)	0.941%	(6.829%)
Liberal	0.404%	(6.343%)	0.414%	(6.420%)	0.414%	(6.423%)
Conservative	0.513%	(7.142%)	0.527%	(7.237%)	0.526%	(7.236%)
Democrat	0.120%	(3.460%)	0.113%	(3.362%)	0.117%	(3.418%)
Republican	0.099%	(3.146%)	0.104%	(3.222%)	0.103%	(3.205%)
Number of dyads	4,604,163,753		46,861,898		46,997,144	
Number of users	60,055,176		611,044		613,096	

Table S5. Comparison of means of close friends of close friends' attributes across the ego's two message types and the control group. Here we show the mean and the standard deviation. It is important to note that people rarely self-report political characteristics on their Facebook profile (less than 1%, as shown).

Decile	*Interaction Threshold*	*Number of Friendships*	*Probability Friend is Close Friend in Real Life*	*95% CI Low*	*95% CI High*
1	>0.0%	19235	2.02%	1.82%	2.22%
2	≥0.2	17313	2.22	2.00	2.44
3	≥0.4	15393	2.46	2.21	2.70
4	≥0.7	13470	2.69	2.42	2.97
5	≥1.0	11546	3.01	2.70	3.33
6	≥1.5	9647	3.35	2.99	3.71
7	≥2.1	7715	3.84	3.41	4.27
8	≥3.1	5771	4.57	4.04	5.11
9	≥4.8	3938	5.51	4.80	6.22
10	≥8.9	1927	7.94	6.73	9.15

Table S6. Rates of interaction on Facebook and the probability that a friend is a close friend in real life, by decile. Each decile encompasses an increasingly restrictive subset of all friends who each account for at least XX% of the interactions (as shown) with the user (so decile 10 is a subset of decile 9, 9 is a subset of 8, and so on).

Expressed Vote							
			Observed Values			*Simulated Null Distribution*	
Decile of Interaction	*Interaction Threshold*	*Number of Friendships*	*Social Message Mean*	*Control Mean*	*Per-User Treatment Effect*	*95% CI low*	*95% CI high*
1	**>0.00%**	**3037246623**	**22.727%**	**22.705%**	**0.022%**	**−0.013%**	**0.013%**
2	**≥0.14**	**2732742699**	**22.508**	**22.485**	**0.023**	**−0.015**	**0.015**
3	**≥0.24**	**2430351234**	**22.385**	**22.358**	**0.027**	**−0.015**	**0.016**
4	**≥0.36**	**2127084596**	**22.261**	**22.229**	**0.031**	**−0.017**	**0.016**
5	**≥0.51**	**1823447544**	**22.127**	**22.095**	**0.032**	**−0.019**	**0.018**
6	**≥0.72**	**1516279553**	**21.961**	**21.916**	**0.045**	**−0.019**	**0.021**
7	**≥1.02**	**1216735474**	**21.750**	**21.703**	**0.048**	**−0.022**	**0.025**
8	**≥1.50**	**915502537**	**21.463**	**21.395**	**0.068**	**−0.024**	**0.026**
9	**≥2.37**	**606735353**	**20.997**	**20.905**	**0.092**	**−0.032**	**0.033**
10	**≥4.51**	**303480053**	**20.098**	**19.999**	**0.099**	**−0.042**	**0.048**

Table S7. The observed effect each friend in the treatment group has on a user's **expressed vote**, ordered by tie strength. Each decile encompasses an increasingly restrictive subset of all friends who each account for at least XX% of the interactions with the user (so decile 10 is a subset of decile 9, 9 is a subset of 8, and so on). The per-friend treatment effect is the mean behaviour among users connected to a friend who received the social message minus the mean behaviour among users connected to a friend in the control group (those who received no message). Since friends are connected to many users and users have many friends, we simulate the null distribution of chance outcomes. We do this by randomly reassigning treatment and control to each user while keeping the network and incidence of the treatments fixed. We repeat this procedure 1000 times, each time measuring the simulated treatment effect, and we report the 95% confidence interval (CI) as indicated by the 25th and 975th values in a sorted list. Significant observed treatment effects (those falling outside the 95% confidence interval of the null) are shown in bold.

Validated Vote							
			Observed Values			*Simulated Null Distribution*	
Decile of Interaction	*Interaction Threshold*	*Number of Friendships*	*Social Message Mean*	*Control Mean*	*Per-User Treatment Effect*	*95% CI low*	*95% CI high*
1	>0.00%	319938668	47.962%	47.955%	0.006%	−0.057%	0.062%
2	≥0.14	291199144	48.213	48.208	0.006	−0.056	0.057
3	≥0.24	260756650	48.588	48.592	−0.005	−0.060	0.064
4	≥0.36	229305366	48.962	48.954	0.008	−0.063	0.062
5	≥0.51	197338921	49.357	49.353	0.005	−0.067	0.069
6	≥0.72	164559444	49.762	49.762	0.000	−0.072	0.074
7	≥1.02	132275117	50.154	50.142	0.012	−0.087	0.084
8	≥1.50	99557388	50.569	50.521	0.048	−0.103	0.105
9	**≥2.37**	**65777012**	**51.034**	**50.862**	**0.172**	**−0.130**	**0.117**
10	**≥4.51**	**32453453**	**51.606**	**51.382**	**0.224**	**−0.181**	**0.174**

Table S8. The observed effect each friend in the treatment group has on a user's **validated vote**, ordered by tie strength. Each decile encompasses an increasingly restrictive subset of all friends who each account for at least XX% of the interactions with the user (so decile 10 is a subset of decile 9, 9 is a subset of 8, and so on). The per-friend treatment effect is the mean behaviour among users connected to a friend who received the social message minus the mean behaviour among users connected to a friend in the control group (those who received no message). Since friends are connected to many users and users have many friends, we simulate the null distribution of chance outcomes. We do this by randomly reassigning treatment and control to each user while keeping the network and incidence of the treatments fixed. We repeat this procedure 1000 times, each time measuring the simulated treatment effect, and we report the 95% confidence interval (CI) as indicated by the 25th and 975th values in a sorted list. Significant observed treatment effects (those falling outside the 95% confidence interval of the null) are shown in bold.

Polling Place Search							
				Observed Values		*Simulated Null Distribution*	
Decile of Interaction	*Interaction Threshold*	*Number of Friendships*	*Social Message Mean*	*Control Mean*	*Per-User Treatment Effect*	*95% CI low*	*95% CI high*
1	>0.00%	3037246623	3.007	3.002	0.005	−0.005	0.006
2	≥0.14	2732742699	2.891	2.886	0.004	−0.006	0.006
3	≥0.24	2430351234	2.817	2.814	0.004	−0.006	0.006
4	≥0.36	2127084596	2.756	2.750	0.006	−0.006	0.006
5	**≥0.51**	**1823447544**	**2.701**	**2.694**	**0.007**	**−0.006**	**0.007**
6	≥0.72	1516279553	2.648	2.640	0.007	−0.008	0.008
7	**≥1.02**	**1216735474**	**2.593**	**2.582**	**0.011**	**−0.009**	**0.009**
8	**≥1.50**	**915502537**	**2.533**	**2.523**	**0.011**	**−0.009**	**0.010**
9	**≥2.37**	**606735353**	**2.458**	**2.445**	**0.012**	**−0.012**	**0.012**
10	≥4.51	303480053	2.340	2.332	0.007	−0.018	0.016

Table S9. The observed effect each friend in the treatment group has on a user's **polling place search**, ordered by tie strength. Each decile encompasses an increasingly restrictive subset of all friends who each account for at least XX% of the interactions with the user (so decile 10 is a subset of decile 9, 9 is a subset of 8, and so on). The per-friend treatment effect is the mean behaviour among users connected to a friend who received the social message minus the mean behaviour among users connected to a friend in the control group (those who received no message). Since friends are connected to many users and users have many friends, we simulate the null distribution of chance outcomes. We do this by randomly reassigning treatment and control to each user while keeping the network and incidence of the treatments fixed. We repeat this procedure 1000 times, each time measuring the simulated treatment effect, and we report the 95% confidence interval (CI) as indicated by the 25th and 975th values in a sorted list. Significant observed treatment effects (those falling outside the 95% confidence interval of the null) are shown in bold.

Expressed Vote		*Observed Values*			*Simulated Null Distribution*		*Observed Values*		*Simulated Null Distribution*			*Simulated Null Distribution*	
Relationship	*Number of Friendships*	*Social Message Mean*	*Control Mean*	*Per-Friend Treatment Effect*	*95% CI Low*	*95% CI High*	*Ave. Friends Per User*	*Per-User Treatment Effect*	*95% CI Low*	*95% CI High*	*Total Effect*	*95% CI low*	*95% CI high*
Friends	**8387533272**	**19.716%**	**19.706%**	**0.011%**	**−0.008%**	**.008%**	**136.4**	**1.4%**	**−1.1%**	**1.1%**	**886,000**	**−696,000**	**687,000**
Close Friends	**606735353**	**20.997**	**20.905**	**0.092**	**−0.032**	**0.033**	**9.9**	**0.9**	**−0.3**	**0.3**	**559,000**	**−194,000**	**200,000**
Close Friends of Close Friends	**4655525113**	**21.474**	**21.452**	**0.022**	**−0.011**	**0.012**	**76.0**	**1.7**	**−0.8**	**0.9**	**1,025,000**	**−512,000**	**559,000**

Table S10. The observed effect each friend in the treatment group has on a user's **expressed vote**, by type of friend. The per-friend treatment effect is the mean behaviour among users connected to a friend who received the social message minus the mean behaviour among users connected to a friend in the control group (those who received no message). The per-user treatment effect is the per-user treatment effect times the number of friends per user. The total effect is the per-user effect times the number of users. Since friends are connected to many users and users have many friends, we simulate the null distribution of chance outcomes. We do this by randomly reassigning treatment and control to each user while keeping the network and incidence of the treatments fixed. We repeat this procedure 1000 times, each time measuring the simulated null treatment effect, and we report the 95% confidence interval (CI) as indicated by the 25th and 975th values in a sorted list. Significant observed treatment effects (those falling outside the 95% confidence interval of the null) are shown in bold. n.s. = not significant.

Validated Vote		*Observed Values*			*Simulated Null Distribution*		*Observed Values*		*Simulated Null Distribution*			*Simulated Null Distribution*	
Relationship	*Number of Friendships*	*Social Message Mean*	*Control Mean*	*Per-Friend Treatment Effect*	*95% CI Low*	*95% CI High*	*Ave. Friends Per User*	*Per-User Treatment Effect*	*95% CI Low*	*95% CI High*	*Total Effect*	*95% CI low*	*95% CI high*
Friends	885681823	44.070%	44.099%	−0.023%	−0.033%	0.030%	140.6	-4.0%	4.7%	4.2%	n.s.	n.s.	n.s.
Close Friends	**65777012**	**51.034**	**50.862**	**0.172**	**−0.130**	**0.117**	**10.4**	**1.8**	**−1.3**	**1.2**	**282,000**	**−215,000**	**193,000**
Close Friends of Close Friends	519588698	52.304	52.302	0.002	−0.046	0.045	82.0	0.2	-3.8	3.7	n.s.	n.s.	n.s.

Table S11. The observed effect each friend in the treatment group has on a user's **validated vote**, by type of friend. The per-friend treatment effect is the mean behaviour among users connected to a friend who received the social message minus the mean behaviour among users connected to a friend in the control group (those who received no message). The per-user treatment effect is the per-user treatment effect times the number of friends per user. The total effect is the per-user effect times the number of users. Since friends are connected to many users and users have many friends, we simulate the null distribution of chance outcomes. We do this by randomly reassigning treatment and control to each user while keeping the network and incidence of the treatments fixed. We repeat this procedure 1000 times, each time measuring the simulated null treatment effect, and we report the 95% confidence interval (CI) as indicated by the 25th and 975th values in a sorted list. Significant observed treatment effects (those falling outside the 95% confidence interval of the null) are shown in bold. (n.s. = not significant.)

Polling Place Search		*Observed Values*			*Simulated Null Distribution*		*Observed Values*		*Simulated Null Distribution*			*Simulated Null Distribution*	
Relationship	*Number of Friendships*	*Social Message Mean*	*Control Mean*	*Per-Friend Treatment Effect*	*95% CI Low*	*95% CI High*	*Ave. Friends Per User*	*Per-User Treatment Effect*	*95% CI Low*	*95% CI High*	*Total Effect*	*95% CI low*	*95% CI high*
Friends	8387533272	2.593%	2.592%	0.001%	−0.003%	0.003%	136.4	0.1%	0.4%	0.4%	n.s.	n.s.	n.s.
Close Friends	**606735353**	**2.458**	**2.445**	**0.012**	**−0.012**	**0.012**	**9.9**	**0.1**	−0.1	0.1	**74,000**	**-71,000**	**72,000**
Close Friends of Close Friends	4655525113	2.434	2.434	0.000	−0.004	0.005	76.0	0.0	−0.3	0.4	n.s.	n.s.	n.s.

Table S12. The observed effect each friend in the treatment group has on a user's **polling place search**, by type of friend. The per-friend treatment effect is the mean behaviour among users connected to a friend who received the social message minus the mean behaviour among users connected to a friend in the control group (those who received no message). The per-user treatment effect is the per-user treatment effect times the number of friends per user. The total effect is the per-user effect times the number of users. Since friends are connected to many users and users have many friends, we simulate the null distribution of chance outcomes. We do this by randomly reassigning treatment and control to each user while keeping the network and incidence of the treatments fixed. We repeat this procedure 1000 times, each time measuring the simulated null treatment effect, and we report the 95% confidence interval (CI) as indicated by the 25th and 975th values in a sorted list. Significant observed treatment effects (those falling outside the 95% confidence interval of the null) are shown in bold. (n.s. = not significant.)

	Expressed Vote Social vs. Informational Message			Polling Place Search Social vs. Informational Message				
Number of friends shown	*Difference of mean*	*95% CI Low of Difference of mean*	*95% CI High of Difference of mean*	*Difference of mean*	*95% CI Low of Difference of mean*	*95% CI High of Difference of mean*	*Social message N*	*Informational message N*
1	0.332%	0.033%	0.631%	0.486%	0.364%	0.609%	72311	615572
2	1.877%	1.521%	2.233%	0.520%	0.377%	0.663%	52432	615572
3	2.507%	2.107%	2.908%	0.445%	0.287%	0.602%	41663	615572
4	2.898%	2.454%	3.342%	0.555%	0.378%	0.731%	33941	615572
5	3.189%	2.703%	3.675%	0.408%	0.220%	0.595%	28407	615572
6	3.854%	3.694%	4.014%	0.196%	0.137%	0.255%	262309	615572

Table S13. The difference in means of the social message group with a specified number of friends shown in the message versus the informational message group for expressed vote and polling place search.

Number of friends shown	Validated Vote Social Message vs. Control: Difference of mean	95% CI Low of Difference of mean	95% CI High of Difference of mean	Validated Vote - Social vs. Informational Message: Difference of mean	95% CI Low of Difference of mean	95% CI High of Difference of mean	Social message N	Informational message N	Control N
1	–0.237%	–1.437%	0.963%	–0.243%	–1.443%	0.957%	7448	64060	64009
2	1.999%	0.600%	3.398%	1.994%	0.595%	3.393%	5304	64060	64009
3	–0.509%	-2.041%	1.022%	–0.514%	-2.046%	1.017%	4377	64060	64009
4	–0.574%	-2.232%	1.084%	–0.579%	-2.237%	1.079%	3699	64060	64009
5	–0.534%	-2.348%	1.280%	–0.539%	-2.353%	1.275%	3061	64060	64009
6	–0.080%	–0.689%	0.528%	–0.086%	–0.694%	0.522%	28881	64060	64009

Table S14. The difference in means of the social message group with a specified number of friends shown in the message versus the informational message group for validated voting and the difference in means of the social message with a specified number of friends shown and the control group.

Expressed Vote							
				Observed Values		*Simulated Null Distribution*	
Decile of Interaction	*Maximum Days Since Interacting*	*Number of Friendships*	*Social Message Mean*	*Control Mean*	*Per-User Treatment Effect*	*95% CI low*	*95% CI high*
1	**92**	**3037246623**	**22.727%**	**22.705%**	**0.022%**	**–0.015%**	**0.013%**
2	77	2724176579	22.210%	22.207%	0.003%	–0.014%	0.015%
3	62	2397704109	21.846%	21.844%	0.003%	–0.015%	0.017%
4	49	2142221054	21.627%	21.627%	0.000%	–0.017%	0.018%
5	37	1802053902	21.418%	21.422%	–0.004%	–0.019%	0.017%
6	28	1505095056	21.230%	21.230%	0.000%	–0.020%	0.019%
7	20	1213860733	21.142%	21.152%	–0.009%	–0.021%	0.022%
8	12	890547047	21.050%	21.060%	–0.010%	–0.027%	0.026%
9	7	588474228	20.895%	20.909%	–0.013%	–0.033%	0.031%
10	2	303467468	20.796%	20.837%	–0.041%	–0.043%	0.043%

Table S15. The observed effect each friend in the treatment group has on a user's **expressed vote**, ordered by recency of contact. Each decile encompasses an increasingly restrictive subset of all friends who each have had contact with the user at minimum X days prior to the election (so decile 10 is a subset of decile 9, 9 is a subset of 8, and so on). The per-friend treatment effect is the mean behaviour among users connected to a friend who received the social message minus the mean behaviour among users connected to a friend in the control group (those who received no message). Since friends are connected to many users and users have many friends, we simulate the null distribution of chance outcomes. We do this by randomly reassigning treatment and control to each user while keeping the network and incidence of the treatments fixed. We repeat this procedure 1000 times, each time measuring the simulated treatment effect, and we report the 95% confidence interval (CI) as indicated by the 25th and 975th values in a sorted list. Significant observed treatment effects (those falling outside the 95% confidence interval of the null) are shown in bold.

Validated Vote							
			Observed Values			*Simulated Null Distribution*	
Decile of Interaction	*Maximum Days Since Interacting*	*Number of Friendships*	*Social Message Mean*	*Control Mean*	*Per-User Treatment Effect*	*95% CI low*	*95% CI high*
1	92	319938668	47.962%	47.955%	0.006%	–0.056%	0.057%
2	77	284655747	47.453%	47.460%	–0.007%	–0.058%	0.060%
3	62	249391708	47.134%	47.148%	–0.014%	–0.066%	0.060%
4	49	222159698	46.974%	46.986%	–0.013%	–0.065%	0.063%
5	37	186203424	46.864%	46.879%	–0.014%	–0.072%	0.077%
6	28	155100870	46.795%	46.774%	0.021%	–0.082%	0.079%
7	20	124973168	46.836%	46.813%	0.023%	–0.087%	0.092%
8	12	91678647	46.899%	46.872%	0.027%	–0.101%	0.106%
9	7	60343083	46.898%	46.860%	0.038%	–0.118%	0.127%
10	2	30980090	46.994%	47.068%	–0.074%	–0.170%	0.174%

Table S16. The observed effect each friend in the treatment group has on a user's **validated vote**, ordered by recency of contact. Each decile encompasses an increasingly restrictive subset of all friends who each have had contact with the user at minimum X days prior to the election (so decile 10 is a subset of decile 9, 9 is a subset of 8, and so on). The per-friend treatment effect is the mean behaviour among users connected to a friend who received the social message minus the mean behaviour among users connected to a friend in the control group (those who received no message). Since friends are connected to many users and users have many friends, we simulate the null distribution of chance outcomes. We do this by randomly reassigning treatment and control to each user while keeping the network and incidence of the treatments fixed. We repeat this procedure 1000 times, each time measuring the simulated treatment effect, and we report the 95% confidence interval (CI) as indicated by the 25th and 975th values in a sorted list. Significant observed treatment effects (those falling outside the 95% confidence interval of the null) are shown in bold. No observed treatment effects are significant.

Polling Place Search							
			Observed Values			*Simulated Null Distribution*	
Decile of Interaction	*Maximum Days Since Interacting*	*Number of Friendships*	*Social Message Mean*	*Control Mean*	*Per-User Treatment Effect*	*95% CI low*	*95% CI high*
1	92	3037246623	3.007%	3.002%	0.005%	–0.005%	0.006%
2	77	2724176579	2.916%	2.914%	0.002%	–0.006%	0.006%
3	62	2397704109	2.854%	2.851%	0.003%	–0.006%	0.006%
4	49	2142221054	2.816%	2.814%	0.002%	–0.006%	0.007%
5	37	1802053902	2.773%	2.774%	0.000%	–0.007%	0.007%
6	28	1505095056	2.736%	2.737%	–0.001%	–0.008%	0.008%
7	20	1213860733	2.711%	2.710%	0.002%	–0.008%	0.008%
8	12	890547047	2.684%	2.683%	0.001%	–0.010%	0.010%
9	7	588474228	2.649%	2.645%	0.003%	–0.011%	0.014%
10	2	303467468	2.617%	2.616%	0.001%	–0.017%	0.016%

Table S17. The observed effect each friend in the treatment group has on a user's **polling place search**, ordered by recency of contact. Each decile encompasses an increasingly restrictive subset of all friends who each have had contact with the user at minimum X days prior to the election (so decile 10 is a subset of decile 9, 9 is a subset of 8, and so on). The per-friend treatment effect is the mean behaviour among users connected to a friend who received the social message minus the mean behaviour among users connected to a friend in the control group (those who received no message). Since friends are connected to many users and users have many friends, we simulate the null distribution of chance outcomes. We do this by randomly reassigning treatment and control to each user while keeping the network and incidence of the treatments fixed. We repeat this procedure 1000 times, each time measuring the simulated treatment effect, and we report the 95% confidence interval (CI) as indicated by the 25th and 975th values in a sorted list. Significant observed treatment effects (those falling outside the 95% confidence interval of the null) are shown in bold. No observed treatment effects are significant.

	Baseline	Heterogeneous Initial Behaviour	Homophily in Initial Behaviour
INTERCEPTS (Mean = 0.00)			
Watts-Strogatz Network	***A***	***B***	***C***
Direct Effect	0.02	0.00	0.01
Friends	0.01	0.00	0.00
Friends of Friends	0.01	0.01	0.00
Erdos-Renyi Network	***D***	***E***	***F***
Direct Effect	0.01	0.01	0.01
Friends	0.00	0.00	0.00
Friends of Friends	0.00	0.00	0.00
Barabasi-Albert Network	***G***	***H***	***I***
Direct Effect	0.00	0.01	–0.01
Friends	0.00	0.01	0.00
Friends of Friends	0.00	0.00	–0.01
SLOPES (Mean = 0.96)			
Watts-Strogatz Network	***A***	***B***	***C***
Direct Effect	0.97	1.00	0.94
Friends	0.97	0.96	0.95
Friends of Friends	0.80	1.02	1.05
Erdos-Renyi Network	***D***	***E***	***F***
Direct Effect	0.97	0.97	0.96
Friends	1.01	0.94	0.96
Friends of Friends	0.88	1.02	1.00
Barabasi-Albert Network	***G***	***H***	***I***
Direct Effect	0.98	0.96	0.97
Friends	1.02	0.90	0.97
Friends of Friends	0.95	0.84	1.07

Table S18. Estimates of the intercept and slope of the regression lines shown in Figure S6 from 1,000 Monte Carlo simulations for each Scenario.

	Baseline	*Heterogeneous Initial Behaviour*	*Homophily in Initial Behaviour*
Watts-Strogatz Network	***A***	***B***	***C***
Direct Effect	0.06	0.05	0.05
Friends	0.05	0.05	0.05
Friends of Friends	0.06	0.06	0.05
Erdos-Renyi Network	***D***	***E***	***F***
Direct Effect	0.05	0.05	0.05
Friends	0.05	0.05	0.05
Friends of Friends	0.06	0.05	0.06
Barabasi-Albert Network	***G***	***H***	***I***
Direct Effect	0.05	0.06	0.05
Friends	0.05	0.05	0.05
Friends of Friends	0.05	0.06	0.05

Table S19. Estimates of the false positive rate generated by the network permutation procedure from 1,000 Monte Carlo simulations for each Scenario, assuming the true effect sizes are all 0.

Figures

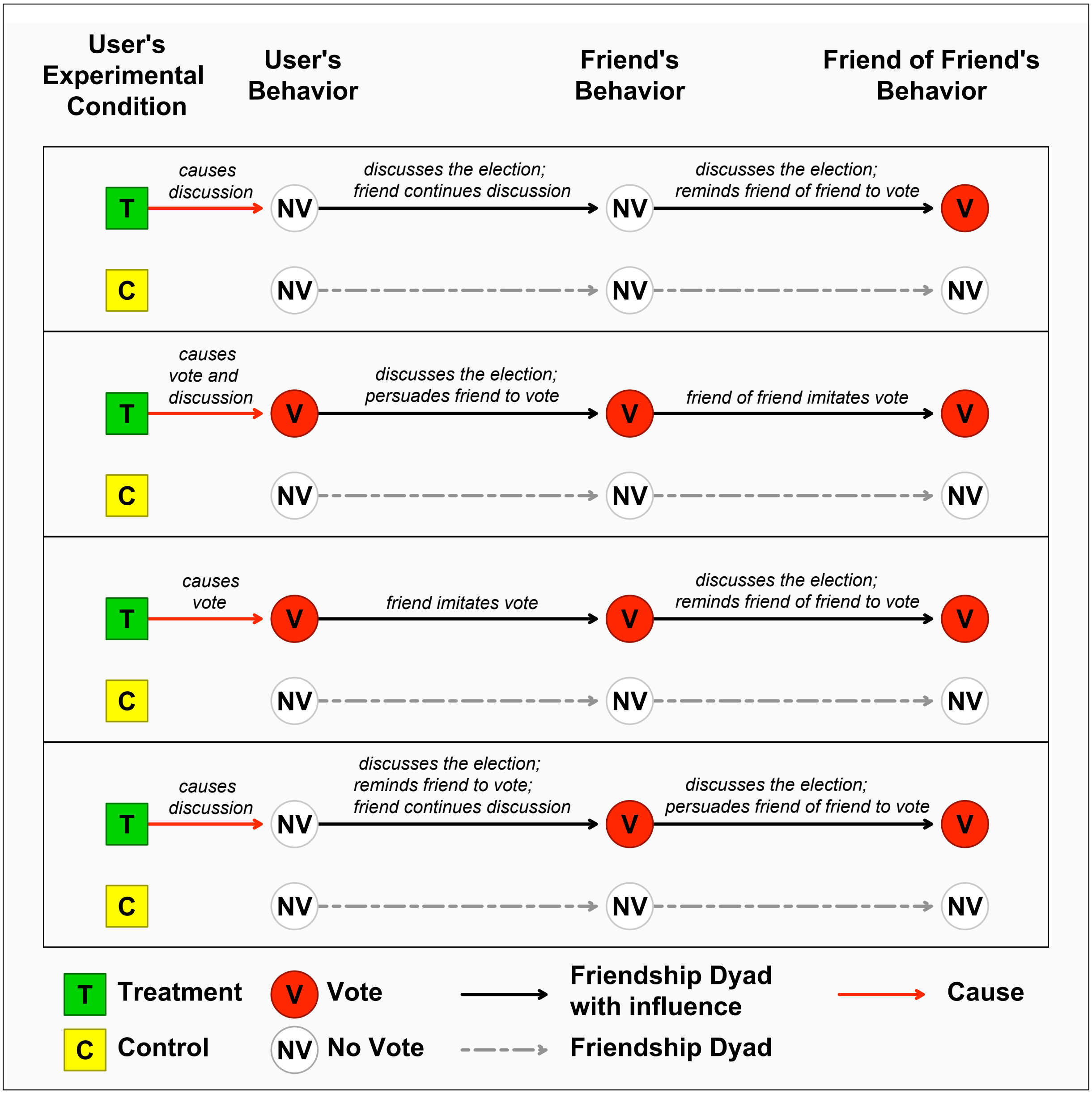

Figure S1. Examples showing how the user treatment may ultimately cause changes in friends' and friends' friends' behaviour. We measure these treatment effects by comparing behaviour of friends of users in the treatment group to the behaviour of friends of users in the control group. This method allows for accurate estimates of the average treatment effect (see Monte Carlo tests in Figure S6) that do not depend on which causal pathways generated the effect.

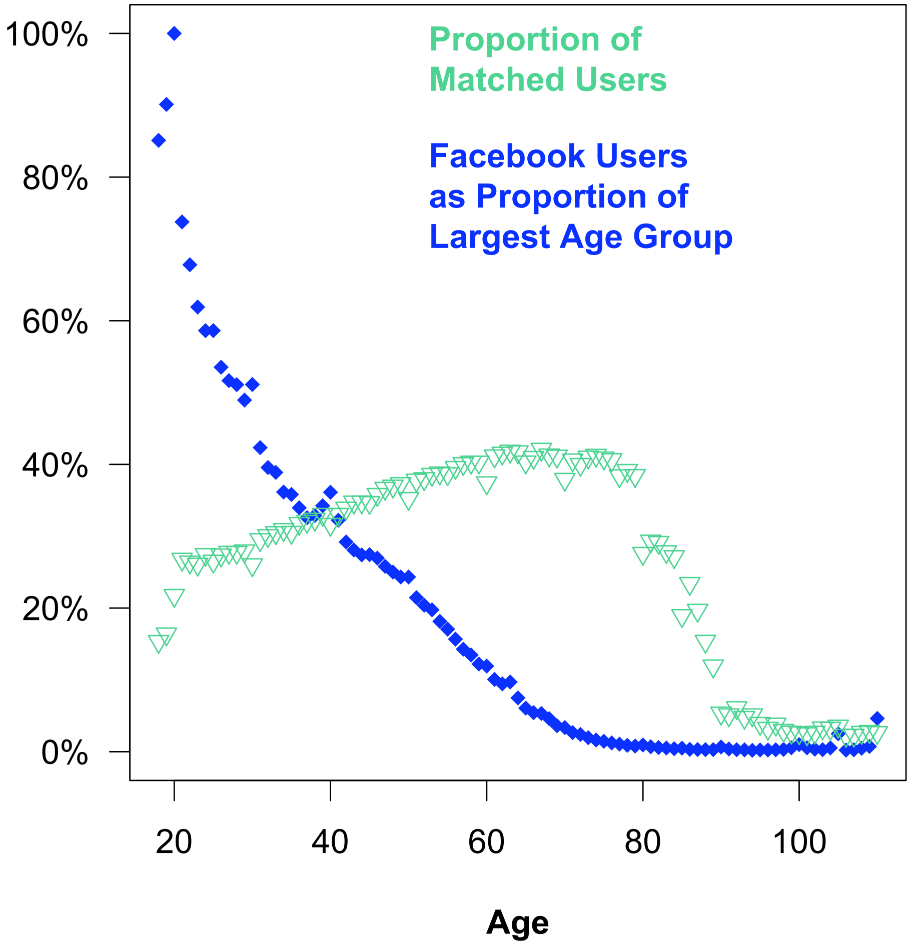

Figure S2. A comparison of Facebook users and match rates by age.

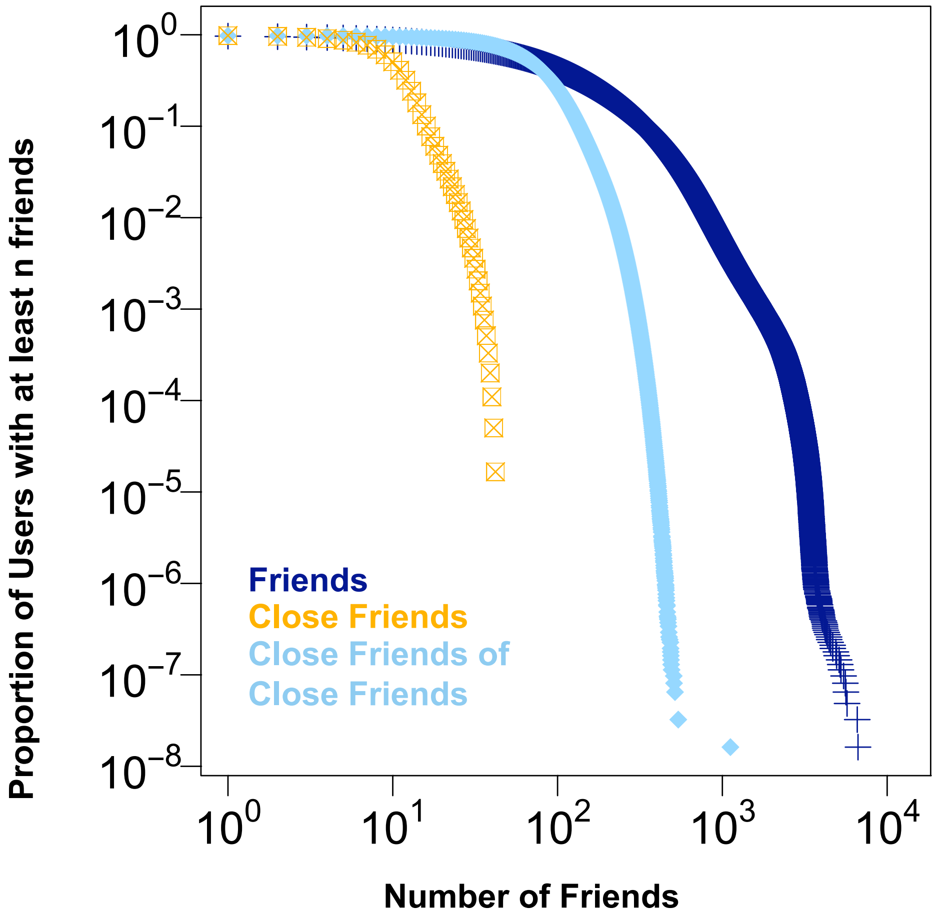

Figure S3. Log-log plot of cumulative distribution of friends, close friends, and close friends of close friends.

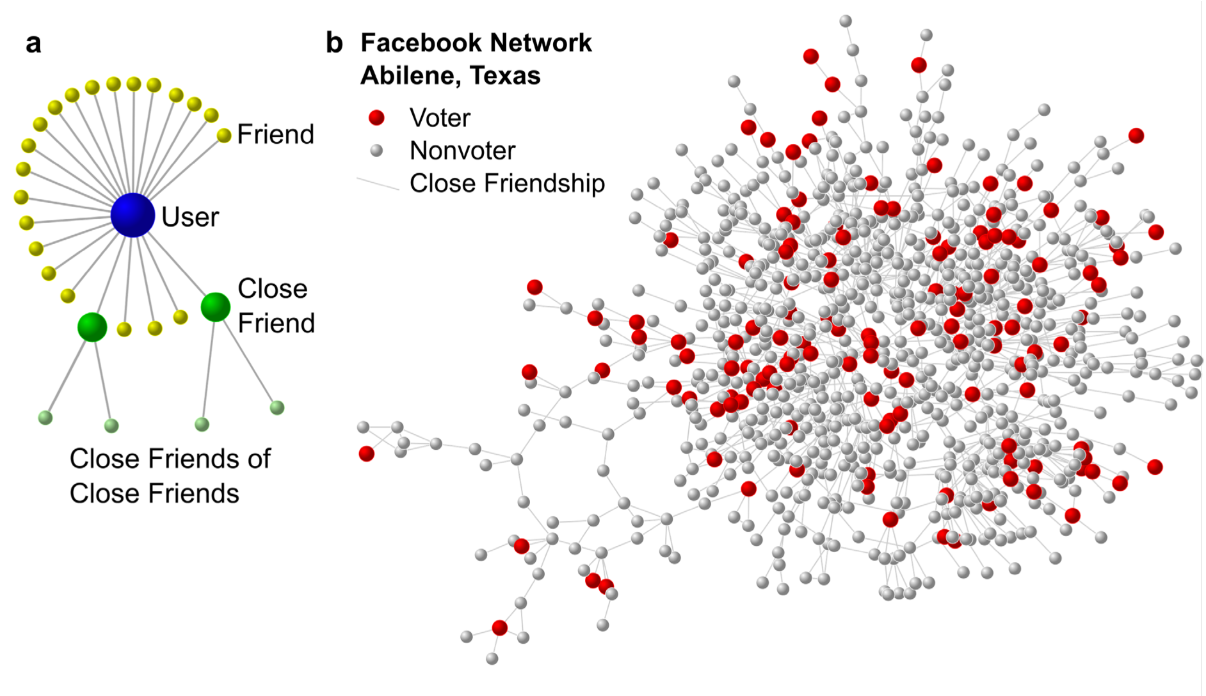

Figure S4. (a) Example egocentric network illustrating the three mutually exclusive types of relationships analysed. "Close Friends" are relationships in the upper two deciles of closeness as measured by the fraction of total interactions of one user with another. "Friends" are connected users who have not interacted or who have interacted at a rate below the 80th percentile threshold. "Close Friends of Close Friends" share a close friend in common but are not directly connected (not even as "Friends"). **(b)** Illustrative map of largest component of the social network of "Close Friends" from Abilene, Texas, who logged in on Election Day. Red nodes are those who clicked on clicked on the "I Voted" button.

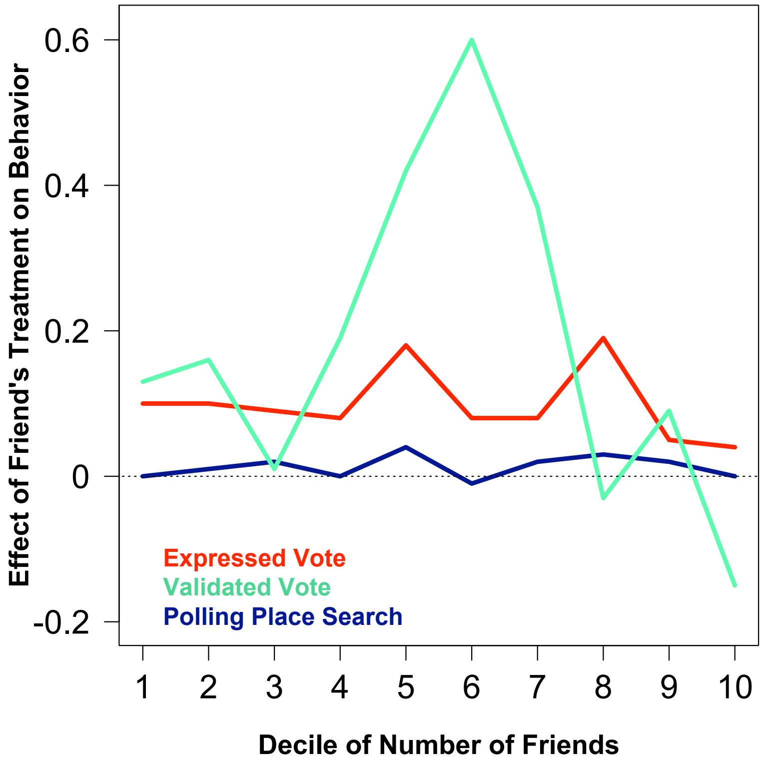

Figure S5. Effect on user behaviour (expressed vote, validated vote, polling place search) of a close friend receiving the social message (versus receiving no message), by decile of number of friends (Decile 1 = users with least friends, Decile 10 = users with most friends).

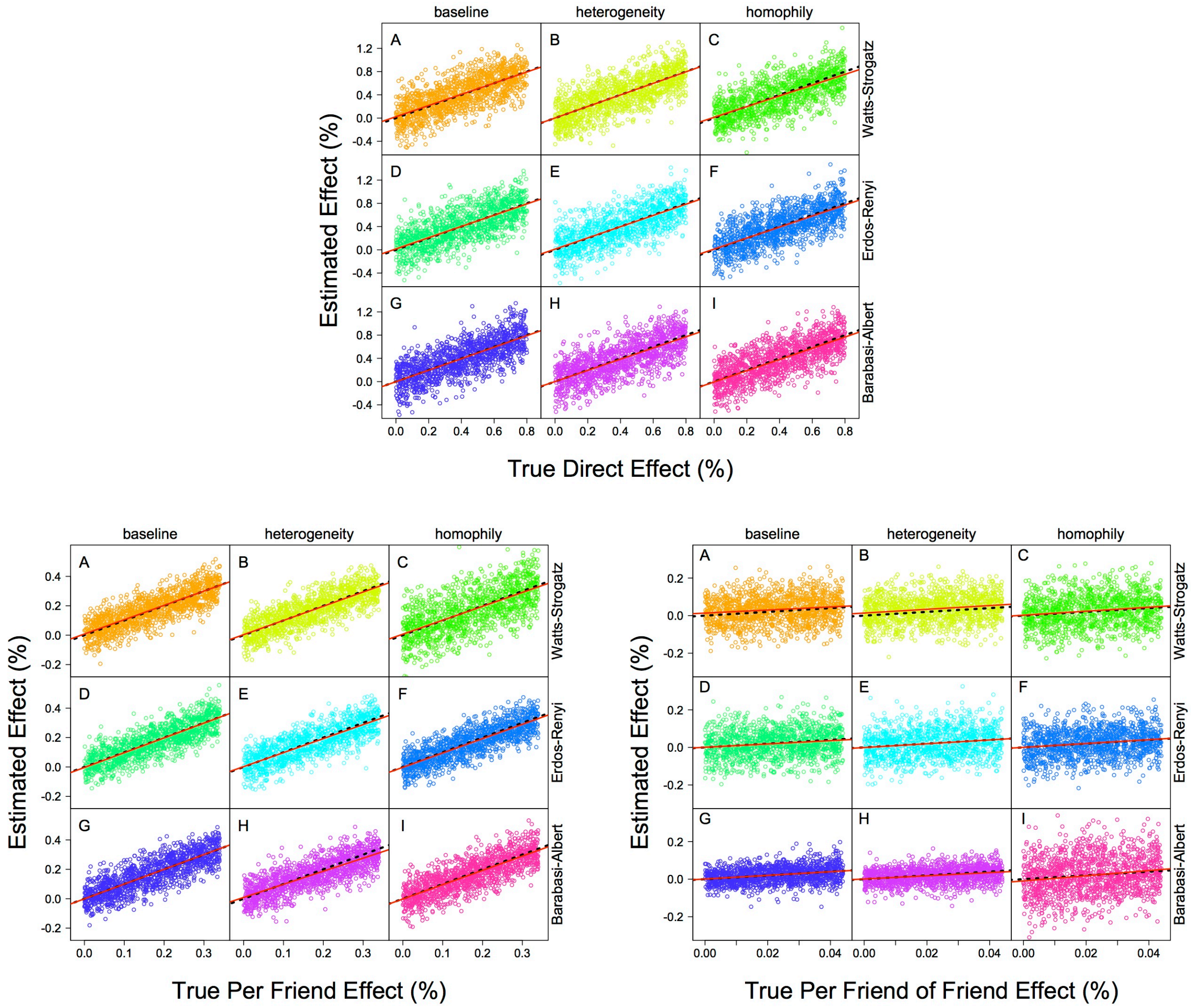

Figure S6. Monte Carlo simulations of the network permutation method used to estimate treatment effects on users, friends, and friends' friends. There are 9 panels for each treatment effect (the direct effect, the effect on friends, and the effect on friends of friends). Each panel shows results from one of the scenarios described above (labelled by the letter of the scenario), each point in a panel is one simulation, and there are 1,000 simulations per panel. The black dotted line is the theoretical relationship between the "true" values we set and the values estimated by our method one would expect if there were no bias in the procedure, and the solid red line is the actual relationship estimated by linear regression. Notice that in all cases the solid line lies very close to the dotted line. Slope and intercept terms for the linear regressions are reported in Table S18.

SI References

[1] Nickerson, D. W. Does email boost turnout? *Q J Polit Sci* **2**, 369–379 (2007).

[2] Haythornthwaite, C. Social networks and Internet connectivity effects. *Information, Communication & Society*, **8**, 125–147 (2005).

[3] Jones, J. J., Bond, R. M., Fariss, C. J., Settle, J. E., Kramer, A., Marlow, C., Fowler J. H. Yahtzee: An Anonymized Group Level Matching Procedure *arXiv*: 1112.1038 (2011).

[4] Boase, J., Horrigan, J. B., Wellman, B., Rainie, L. The Strength of Internet Ties *Pew Internet and American Life Project* Retreived May 15, 2012 from http://www.pewinternet.org/Reports/2006/The-Strength-of-Internet-Ties.aspx (2006).

[5] Kamada, T. & Kawai, S. An algorithm for drawing general undirected graphs. *Information Processing Letters* **31**, 7–15 (1989).

[6] Csardi, G. & Nepusz, T. The igraph software package for complex network research. *InterJournal Complex Systems* 1695. URL: http://igraph.sf.net (2006)

[7] R Development Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3–900051–07–0. URL: http://www.R-project.org (2009).

[8] Nooy, W., Mrvar, A., Batagelj, V. *Exploratory Social Network Analysis with Pajek.* (Cambridge University Press, New York, 2005).

[9] Fowler, J. H. Turnout in a Small World, in A. Zuckerman, ed., *The Social Logic of Politics: Personal Networks as Contexts for Political Behaviour*. Temple University Press, 269–287 (2005).

[10] Centola, D., Macy, M., Complex Contagions and the Weakness of Long Ties *American Journal of Sociology* **113**, 702–734 (2007).

[11] Cacioppo, J. T., Fowler, J. H. and Christakis, N. A. Alone in the crowd: The structure and spread of loneliness in a large social network. *Journal of Personality and Social Psychology* **97**, 977–991 (2009).

[12] Christakis, N. A. and Fowler, J. H. The spread of obesity in a large social network over 32 years. *New England Journal of Medicine* **357**, 370–379 (2007).

[13] Christakis, N. A. and Fowler, J. H. The collective dynamics of smoking in a large social network. *New England Journal of Medicine* **358**, 2249–2258 (2008).

[14] Fowler, J. H. and Christakis, N. A. Dynamic spread of happiness in a large social network: longitudinal analysis over 20 Years in the Framingham Heart Study. *British Medical Journal* **337**, a2338 (2008).

[15] Fowler, J. H., Settle, J. E. and Christakis, N. A. Correlated genotypes in frienship networks. *PNAS: Proceedings of the National Academy of Sciences* **106**, 1993–1997 (2011).

[16] Rosenquist, J. N., Fowler, J. H. and Christakis, N. A. Social network determinants of depression. *Molecular Psychiatry* **16**, 273–281 (2011).

[17] Rosenquist, J. N., Murabito J., Fowler, J. H. and Christakis, N. A. The spread of alcohol consumption behaviour in a large social network. *Annals of Internal Medicine* **152**, 426–433 (2010).

[18] Mednick, S. C., Christakis, N. A. and Fowler, J. H. The spread of sleep loss influences drug use in adolescent social networks. *PLoS One* **5**, e9775 (2010).

[19] Sinkovits, R. S., Cicotti, P., Strande, S., Tatineni, M., Rodriguez, P., Wolter, N., Balac, N. Data intensive analysis on the gordon high performance data and compute system. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining* (KDD '11). ACM, New York, NY, USA, 747–748 (2011).

[20] Watts, D. J. & Strogatz, S. H. Collective dynamics of 'small world' networks, *Nature* **393**, 440–442 (1998).

[21] Erdos, P. & Renyi, A. On random graphs, *Publicationes Mathematicae* **6**, 290–297 (1959).

[22] Barabasi, A.-L. & Albert, R. Emergence of scaling in random networks. *Science* **286**, 509–512 (1999).