Volunteer Phone Calls Can Increase Turnout

Evidence From Eight Field Experiments

David W. Nickerson
University of Notre Dame

Gerber and Green argue that get-out-the-vote phone calls do not increase turnout based upon field experiments testing nonpartisan professional phone banks. This article argues that the quality of the phone calls matter and that brief, nonpartisan phone calls can raise voter turnout if they are sufficiently personal. To test this hypothesis, a series of eight volunteer nonpartisan phone campaigns to mobilize voters were studied using randomized, controlled experiments. The campaigns targeted voters across six different cities in 2000 and 2001. Contra Gerber and Green, the phone calls are found to boost turnout 3.8 percentage points. Based on these estimates, volunteer phone calls produce one vote for every $26 per vote, which is cost competitive with door-to-door canvassing. Differences between the professional phone banks previously studied by Gerber and Green and the volunteer phone banks herein are also discussed.

Keywords: election campaigns; political participation; voter turnout; field experiments; phone banks; natural experiments; voting

In the 2004 presidential election, telephones played a crucial role in the get-out-the-vote (GOTV) strategy for both political parties. In Ohio alone, the Democrats claimed that 3 million phone calls had been made by volunteers prior to Halloween, whereas Republicans boasted 2.5 million volunteer phone calls (Polman, 2004). In the 96 hours leading up to election day, both parties claimed to make nearly 400,000 phone contacts per day (Balz & Edsall, 2004). In an election that came down to the wire, both parties were relying heavily upon telephone conversations to move supporters to the polls.

GOTV phone calls now constitute a large part of American electioneering; however, the efficacy of phone calls as a means of increasing voter turnout has been recently questioned. In 1998, Gerber and Green conducted two large-scale field experiments to test the utility of nonpartisan phone calls as a means of increasing voter turnout in New Haven and West Haven, Connecticut (Gerber & Green, 2000b, 2001). In contrast to the numerous smaller experiments (Adams & Smith, 1980; Eldersveld, 1956; Miller, Bositis, & Baer, 1981) and observational studies (e.g., Rosenstone & Hansen, 1993), Gerber and Green conclude that phone calls do not move voters to the polls. Contrary to expectations, both Gerber and Green experiments detected a noisy *decrease* in voter turnout from receiving a brief, nonpartisan GOTV call from a professional phone bank (–2.0 percentage points for New Haven and –1.3 percentage points for West Haven, each with a standard error of roughly 2).[1] Pooling the two Gerber and Green experiments, the mobilization effect of the phone calls is estimated to be –1.6 percentage points with a standard error of 1.4. The odds of the Gerber and Green calling campaigns conducted in 1998 being effective at mobilizing the vote (i.e., having a positive treatment effect) are roughly 13 in 100. If Gerber and Green (2001, p. 82) are correct when they describe phone calls as ineffective and increasing turnout "not at all" (Gerber & Green, 2000b, p. 653), then both political parties made a costly mistake, financially and politically.

The expectations of both political parties and the Gerber and Green null finding can be reconciled if the assumption that all phone calls are equally effective is relaxed. The random assignment to treatment and control groups assures the internal validity of the two 1998 experiments, however, the external validity of the finding remains an open question.[2] There are two primary reasons to believe that phone calls can be effective in mobilizing voters and that Gerber and Green detected the effect of a particular type of phone call.

First, phone calls are effective across a wide range of settings. Telemarketing would not have developed into an industry with $100 billion in annual sales if phone calls were not effective at motivating consumers ("A New Internet Democracy?", 2003). Phones are also used to encourage charitable donations and volunteer work in a wide range of settings. For instance, a randomized field experiment found phone calls using a "foot in the door" message to be effective in eliciting blood donations (Hayes, Dwyer, Greenwalt, & Coe, 1984). If a phone call can motivate monetary and blood donations, why should voter turnout differ?

Second, door-to-door canvassing and direct mailings are effective at mobilizing voters (Gerber & Green, 2000b; Gerber, Green, & Green, 2003; Green, Gerber, & Nickerson, 2003). The reason typically given for the increased efficacy of face-to-face efforts over mailings is the more personal

nature of the appeal. The theory is intuitive and holds for encouraging recycling as well (see Reams & Ray, 1993). Why would phone calls fail to mobilize voters where direct mail succeeds? Although contemporary direct mail is far more targeted and personalized than in the past, a letter still cannot adapt, respond, and react to a voter the way a live person can. Why should an otherwise sensible theory supported by all the other voter mobilization data not fit phone calls?

If it is true that the personal nature of the appeal is positively correlated with its effectiveness, then phone calls of a personal nature ought to boost voter turnout. This article argues that Gerber and Green detected a feature of the particular type of call studied rather than GOTV calls in general. Perhaps the *professional* phone calls employed by Gerber and Green were insufficiently persuasive. A series of eight voter turnout experiments using *volunteer* phone banks were conducted to test this hypothesis. The volunteer phone callers were chatty and informal; the overarching goal of the caller was to make a personal connection with the voter. In every one of the eight different experiments, the treatment group voted at a higher rate than the control group. The conclusion drawn is that nonpartisan volunteer phone banks can increase voter turnout. Moreover, volunteer phone campaigns appear cost competitive with door-to-door canvassing campaigns as a mobilization technology.

One implication of the article is that in the 2004 election the Republican Party executed a better endgame strategy than the Democratic Party. The Democrats boasted an impressive army of 250,000 volunteers, but the Republicans claimed to have 4 times as many volunteers working on the campaign (Balz & Edsall, 2004). To make up the difference, the Democrats used paid workers. If professional phone banks are ineffective and volunteer phone banks are highly effective, then the narrow defeat of the Democrats in key battleground states is partially attributable to inferior mobilization technology.

The next section describes the experimental protocols, experimental estimators, and the setting for each of the field experiments. The second section presents the results showing a substantively and statistically significant mobilization to vote. The article concludes by discussing directions for future research.

Experimental Design and Setting

Selection bias plagues studies where contact from political campaigns is endogenous and beyond the control of the researcher. With regard to phone

voter mobilization experiments, selection bias can result from behavior of the campaign (e.g., political campaigns often target active voters rather than habitual nonvoters) and individual voters (e.g., voters who cannot be reached via phone are less likely to vote than those who answer the phone). Observational studies will produce misleading results because of both levels of selection bias. To avoid such bias, the controlled experiments conducted in 1998 and for this article randomly assign voters to be called by the political campaign, thereby assuring that the people called possess the same underlying propensity to vote as the people the campaign did not contact. The 1998 Gerber and Green professional phone voter mobilization experiments are internally valid; the question is the degree to which the findings can be extended to other types of phone calls. Field experiments testing the efficacy of volunteer GOTV calls can provide a more complete picture of the role phone calls play in political campaigns.

Part of the beauty of controlled experiments is that random assignment of the treatment facilitates simplicity of design and analysis. All eight of the experiments in this article share the same straightforward design. Local campaigns were attempting to mobilize a target population and agreed to participate in the experiment. The respondents in the target population were randomly divided into treatment (i.e., respondents to receive a phone call) and control groups (i.e., respondents who would not receive a call). The random nature of the assignment ensures that on average the treatment and control groups are identical across both observed characteristics (e.g., age, party membership, neighborhood) and unobserved characteristics (e.g., interest in the election, health, availability for contact). Thus, to obtain an unbiased estimate of the effectiveness of the overall campaign, one can simply subtract the voter turnout rate of the treatment group, Y_T, from that of the control group, Y_C. This difference, $Y_T - Y_C$, is also called the intent-to-treat effect.

The chief advantage of field experiments compared to lab experiments is that inquiry takes place under real-world conditions, thereby ameliorating concerns about external validity. The downside to the real-world setting is that political parties cannot feasibly contact every targeted voter.[3] Because the prescribed treatment is not provided to each respondent, estimating the boost in turnout from direct contact, β, involves slightly more analysis than the intent-to-treat effect. Let γ equal the percentage of respondents who can be contacted by the campaign. Because the treatment and control groups are randomly determined, γ is the same for both the treatment and control groups.[4] We can therefore represent the voter turnout rates for the treatment and control groups as follows:

$$V_T = \gamma \, (V_R + \beta) + (1 - \gamma)V_{\sim R} \qquad\qquad (1)$$
$$V_C = \gamma V_R + (1 - \gamma)V_{\sim R},$$

where V_R is the baseline voting rate for those individuals who can be reached by the campaign and $V_{\sim R}$ equals the baseline voting rate for people who cannot be reached by the campaign. Solving for the boost in turnout from those contacted by the campaign, also called the treatment upon the treated effect, one can derive the following estimator:

$$\beta \frac{V_T - V_C}{\gamma}. \qquad\qquad (2)$$

In other words, the direct effect of contact from the campaign can be calculated by dividing the intent-to-treat effect by the contact rate. As long as information about the rate at which respondents assigned to the treatment group are actually reached by the campaign is available, the effect of contact can be accurately estimated.

The experiments were conducted in two waves with the cooperation of YouthVote, a coalition of nonpartisan, nonprofit groups committed to increasing voter turnout among citizens aged 18 to 30. The first set of six experiments (Sites 1-4) was conducted during the 2000 presidential election. The second set of experiments (Sites 5 and 6) was conducted a year later in local elections. In both years, phone banks were hastily put together in the week prior to the election and calls were made the weekend prior to the election. Table 1 summarizes many of the important facets of the calling campaigns.

Site 1: Albany, New York. The population used in this study consisted of voters registered by student organizations during September and October of 2000 on the State University of New York (SUNY)–Albany campus. Thus, the respondents in the experiment were new voters living on or near college campuses. Calling began the day before election day by students with varying degrees of phone campaign experience. No script was provided to the callers. The instructions to the callers were to simply introduce themselves, to ask the person to vote, to not offer an opinion on any political matter, and to work their way through the entire treatment group before calling numbers a second time. The actual calls were made in several offices across campus; oversight was limited.

Site 2: Stonybrook, New York. The population used in this study consisted of voters registered by student organizations during September and October 2000 on the SUNY–Stonybrook campus. Calls were made the day

Table 1
Description of the Volunteer Phone Campaigns

	Albany	Stonybrook	Eugene	Eugene	Boulder	Boulder	Boston	Seattle
Year	2000	2000	2000	2000	2000	2000	2001	2001
Target age	18-30	18-30	18-30	18-30	18-30	18-30	All	All
Target area	Campus	Campus	Campus	Around campus	Campus	Around campus	Entire city	Entire city
Source	Registered by group	Registered by group	Registered by group	Purchased from vendor	Registered by group	Purchased from vendor	Purchased from vendor	County clerk
Script	None	None	Yes	Yes	Appendix	Appendix	Yes	Yes
Training	Night of	Night of	Day before	Day before	Day before	Day before	Night of	Day before
Oversight	Little	None	Yes	Yes	Little	Little	Yes	Yes
Nights	Monday	Monday	Wednesday-Monday	Wednesday-Monday	Friday-Monday	Friday-Monday	Sunday-Monday	Tuesday-Monday
Treatment n	804	680	705	953	653	1,143	1,209	1,0757
Calls attempted	564	664	705	953	615	1,075	1,209	1,412
% spoke with target	39	27	49	25	30	7	29	34
% spoke with roommate	11	11	11	9	17	9	7	8
% left message	37	53	15	15	30	20	20	31
% line busy	3	8	13	13	2	2	4	4
% not home	NA	NA	NA	NA	NA	NA	16	12
% bad number	11	1	13	37	20	61	24	10

Note: NA = not available. "Calls attempted" means a caller dialed the number over the course of the phone campaign. "Spoke with target" means a caller spoke directly with the targeted respondent. "Spoke with roommate" means the caller spoke with the roommate of the intended target (and verified the respondent used the phone number). "Left message" means the caller left a message on the answering machine of the intended respondent. "Line busy" means the line was busy when the caller dialed the number. "Not home" means that the caller was told that the intended respondent was not home. "Bad number" means the line was disconnected or the target respondent no longer used the phone number.

before election day. The callers all had prior phone advocacy experience. No script was provided. The instructions to the callers consisted only of a firm reminder to remain strictly nonpartisan and a description of how to record call dispositions on the sheet (e.g., "contacted," "left message," "busy," etc.). After receiving instructions and a list of registered voters to call, the volunteers returned to dorm rooms to make the calls. Oversight of the callers was virtually nonexistent. The organizers visited the volunteers to pick up the sheets on the day after election day. The callers were extremely effective at reaching respondents in the treatment group, speaking with or leaving messages for 89% of them.

Sites 3 and 4: Boulder, Colorado, and Eugene, Oregon. The experiments conducted in Boulder and Eugene consisted of two populations: citizens registered by student groups on campus (as in Albany and Stonybrook) and names purchased from a commercial vendor. This distinction provides a test of whether mobilization campaigns are unusually effective when a group tries to mobilize people it had registered during the preceding months. An independent control group was randomly extracted from both the student-generated list and the purchased list. The treatment groups for both lists were mixed together so that callers were unable to determine the source of the names.

One hypothesis is that the respondents registered to vote by the callers would be easier to mobilize because a connection had already been established. On the other hand, it could be argued that the respondents from the vendor list would be more strongly influenced by a phone call because when compared to the sample of registered students, these young people have on average less day-to-day exposure to political activities in and around campus and fewer social contacts with other political active youth (Bennett, 1991).

In fact, the most striking difference between the student-generated and the vendor-supplied lists was the accuracy of the phone numbers. Because the groups had registered individuals on campus in the month prior to the registration deadline, the phone numbers for these respondents were relatively accurate. In contrast, information on the vendor-supplied list was at least a year old, resulting in many disconnected and wrong numbers and a low contact rate (see Table 1). The low contact rate does not bias estimates, because inaccurate phone numbers were equally likely to be assigned to the treatment and control groups. However, the low contact rates decrease the precision of the estimate for the direct treatment effect. Thus, young voters move frequently, and recent databases are necessary to run an efficient calling campaign.

Site 3: Eugene. Calling began 6 days prior to election day, so volunteers were able to make a second pass through the list in an effort to reach numbers

for which there was no answer during the first pass through the list. Most of the callers had past experience with phone campaigning. Callers were provided a detailed script but were urged to develop their own style. The calling took place in a large room with multiple phone lines, and a supervisor was present at all times to monitor and answer questions. The contact rate for the student-generated list was relatively high, with one in two respondents being contacted directly. The vendor-purchased list proved somewhat out of date, and 50% of the numbers were found to be disconnected or incorrect. However, callers did manage to speak directly with 25% of the treatment respondents from the vendor-supplied list.

Site 4: Boulder. Calling took place between the Friday and Monday prior to election day. The labor pool for the Boulder phone campaign was slightly different from the other sites in this article. In addition to typical college activists, the Boulder YouthVote Coalition utilized the manpower of individuals attempting to fulfill court-mandated community service hours.[5] The training of new volunteers occurred each night prior to calling. The phones were located in many different offices, and oversight of the actual calls was minimal. During the course of the conversation, callers reminded respondents of the location of their polling places. Much of the information on the vendor-supplied list proved to be outdated, and 60% of the phone numbers were incorrect or disconnected.

Sites 5 and 6: Boston, Massachusetts, and Seattle, Washington. Two experiments were conducted in Boston and Seattle during 2001 local elections as a follow-up to the phone experiments conducted during the 2000 election described above. Perhaps the mobilization effect uncovered in 2000 was unique to the population targeted (i.e., voters younger than 30) or to presidential elections. In 2001, respondents were drawn from the entire population of registered voters in a city to determine if the mobilization effect was age specific.

Site 5: Boston. The respondent population consisted of 7,055 randomly selected registered voters of all ages (purchased from Survey Sampling Inc.). Local volunteers began calling respondents the Thursday prior to election day. Callers were drawn from a pool of local activists. The callers were young, and roughly two thirds of the callers had prior experience in phone canvassing. Calling took place in a single centralized room, so organizers could oversee every call. Sample scripts were provided and callers quickly developed individual variations on the script. Mayoral elections in Boston are generally uncompetitive, with low salience among voters. The most competitive issue on the ballot was a "livable city initiative" aimed at curbing rising residential rents. Callers were instructed to not endorse any particular

candidate but were allowed to express an opinion on the livable-city initiative.

The respondent population in Boston was slightly idiosyncratic and deserves mention. After the 2000 election, the Boston city clerk conducted the purge of the voter rolls mandated by law by mailing nonvoters an address verification card. Nonvoting residents who did not return the card were removed from the list of active voters and labeled inactive. The list of registered voters purchased contained only active voters, so the sample is heavily overweighted with citizens who voted in the 2000 election (93% of the respondents in the sample voted in 2000). Such imbalance does not bias the results in any fashion but might limit the extent to which the results can be generalized to nonvoters.

The Boston experiment also used a slightly different experimental protocol than the one used in the experiments in 2000. Rather than dividing respondents into monolithic treatment and control groups, they are placed in a random order. Volunteers are told to begin calling at the top of the list and gradually work their way down. The names toward the end of the list who were never attempted constitute the control group because placement within the list is random. The protocol maintains statistical efficiency in the face of labor shortfalls (see Nickerson, in press-b).

Site 6: Seattle. The list of registered voters was obtained for King County, and residents of Seattle were phone matched by consumer research firm InfoUSA. The roughly 44,000 households for which a phone number was available were then placed in a random order. The volunteers began calling the week prior to election day and managed to dial only 1,420 numbers. The callers were members of the campus Public Interest Research Group (PIRG) chapter. Almost every one of the callers possessed considerable experience in phone canvassing. A script was provided, but every caller with a little experience ignored the script and used an individual style. The calls were strictly nonpartisan throughout the campaign. Because all of the calls took place in a single office, there was considerable oversight of the calls by organizers. Callers were extremely successful in contacting people at the numbers dialed (73% were completed), but very few phone numbers in the treatment group were attempted. Thus, the contact rate for the Seattle experiment is a very low 9.6%, and the results are therefore noisy and essentially uninformative ($SE = 5.5$). The ballot was headlined by a mayoral race that was not hotly contested.[6]

To summarize, in every one of the experiments the callers consisted of college-aged men and women. The volunteers tended to be active in campus politics but exhibited a broad range of experience levels with regard to phone

canvassing. The minimal training provided to the volunteers prior to making the calls focused on how to correctly fill out the form for the experiment rather than on the content of the conversation. At times, organizers suggested language that callers could use, but volunteers quickly developed their own spiel. When used, the scripts were broadly similar across the various sites, and the script from Boulder is provided in the appendix as an example. Consistent across every site, callers could use their own language as long as callers remained strictly nonpartisan.

In contrast, callers on behalf of professional phone banks, as used in Gerber and Green (2000b, 2001), can be characterized as rigidly adhering to a script and attempting to get through the conversation as quickly as possible. The incentive structure for the firm and the callers is to complete as many calls as possible each hour. For the eight experiments in this article, the volunteer phone calls were chatty and informal. Callers paused to wait for responses from the respondent and attempted to engage in conversation. The campaigns attempted to convey the characteristics of "a call from a concerned peer" rather than "a call from someone doing a job." Although the conversations may have been informative, no one would confuse the student-organized calling campaign with a professionally managed call center. In short, this article tests a different type of phone campaign from that tested by Gerber and Green.

Each one of the eight individual experiments is small and will not provide a definitive answer. However, by pooling all of the experiments, clear patterns emerge, and the reader can be confident in the results.[7] Conducting the experiment across such a wide variety of locations and during two different elections should reassure the reader that the voter mobilization from receiving a phone call detected by the experiments is not an idiosyncratic result of a particular election or geographic region. The calls made in each experiment were very similar; examining the experiments jointly allows the reader to know the average effect of the volunteer GOTV phone calls across a range of settings.

The next section will present and analyze the results of the experiments to establish that volunteer phone calls can be an effective means of mobilizing the vote. Because the 2000 YouthVote experiments share the same time period and target populations, results for 2000 experiments will be presented first. Once it has been established that the 2000 volunteer calling campaign mobilized voters, the Boulder and Eugene experiments will be quickly analyzed to determine whether respondents from the student-generated lists were mobilized at the same rate as those respondents drawn from vendor-supplied lists. The 2001 experiments will then be analyzed. Once all eight experiments have been used to estimate the mobilization effect of volunteer

phone calls, the cost-effectiveness of volunteer phone calls will be calculated and compared to other mobilization technologies.

Results

Each of the individual phone experiments during 2000 was small in scale (see Table 2, row 5), but pooled they can provide a reasonably precise estimate of the effectiveness of a particular type of phone campaign (i.e., volunteer phone banks targeting young voters). In all five of the experiments, the treatment group turned out to vote at higher rates than the control group (see Table 2, row 9).[8] The intent-to-treat effect for one site stands out. Voters in Stonybrook were mobilized considerably more than in other sites.[9] Of course, the estimated experimental effect is drawn from a normal distribution; thus, the high Stonybrook estimate is balanced by modest levels of mobilization witnessed among the voters registered by YouthVote in Eugene[10] and among the voters whose names were purchased from a vendor in Boulder. Pooled across all of the sites from the 2000 presidential election, the treatment group voted at a rate 2.3 percentage points higher than the control group with a standard error of 1.0 (see Table 2, column 8, row 9).

To estimate the mobilizing effect of volunteer phone calls on those actually contacted by the campaign, the intent-to-treat effect is divided by the contact rate. The contact rate tended to be higher in the sites where respondents were registered to vote by the same campus groups later doing the calling. The higher contact rate is primarily a result of the information being gathered more recently. The vendors were relying on information that was at least a year old, which, given the transitory nature of young voters, means that the phone numbers were more likely to be incorrect (see Nickerson, in press-a). The low contact rates will not bias the results but will considerably increase the standard errors.

The average boost in turnout from all nonpartisan phone contacts in the 2000 YouthVote campaign is estimated to be 4.5 percentage points with a standard error of 1.7 (see Table 2, row 10).[11] That is, a person contacted by the campaign is 4.5 percentage points more likely to vote in the election. If the likelihood of a person turning out to vote were 50%, after receiving a GOTV phone call the same person would be 54.5% likely to vote. One can also interpret the number from the campaign's perspective. If a nonpartisan volunteer campaign were to make 1,000 phone contacts, it would create 45 "new" votes.

In addition, some of the 2000 YouthVote experiments were designed to determine whether those registered by the campaign were easier to mobilize

Table 2
Results for Volunteer Phone Voter Mobilization Field Experiments

	Albany	Stonybrook	Eugene	Eugene	Boulder	Boulder	Pooled	Boston	Seattle	Pooled	Pooled
Year	2000	2000	2000	2000	2000	2000	2000	2001	2001	2001	Both
n	1,122	959	1,202	1,960	1,094	2,318		7,055	32,271		
% assigned treatment	71.7	70.9	58.7	48.6	59.7	49.3		17.1	24.6		
% control turnout	74.2	70.6	61.2	55.0	62.6	28.0		54.5	64.7		
% treatment turnout	78.0	78.8	61.7	57.6	64.9	28.2		56.1	64.7		
Intent-to-treat effect[a]	3.8 (2.9)	8.2 (3.1)	0.5 (2.9)	2.6 (2.2)	2.3 (3.0)	0.2 (1.9)	2.3 (1.0)	1.6 (1.6)	0.0 (0.6)	0.2 (0.5)	0.6 (0.5)
% contact rate	61.6	88.7	74.3	49.0	72.1	34.5		55.4	9.6		
Estimated effect on the treated[a]	6.1 (4.7)	9.3 (3.7)	0.7 (3.8)	5.3 (4.6)	3.3 (4.1)	0.5 (5.4)	4.5 (1.7)	2.9 (2.8)	0.1 (5.5)	2.3 (2.5)	3.8 (1.4)
p value	.10	.01	.43	.12	.22	.46	< .01	.16	.50	.18	.004

Note: Numbers in parentheses represent standard errors; pooled values are precision-weighted averages; p values calculated are single-tailed.
a. Percentage point estimates.

than citizens whose names were obtained from a list vendor. The experiments in Boulder and Eugene were split into samples registered by the group conducting the campaign and those names purchased from a vendor (see Table 2, columns 4-7). In Eugene, the respondents registered by the campaign were less mobilized than those respondents whose names were purchased, whereas in Boulder, the students registered by the campaign were more mobilized. These contradictory findings would suggest that there is little difference between mobilizing voters with whom the campaign has an established relationship and cold calling to encourage turnout.

The YouthVote experiments in 2001 attempted to determine whether the mobilization effect uncovered in 2000 was limited to presidential elections and/or the youth population. Volunteer phone banks were still the method of mobilization, but the population included all ages of registered voters, and the local elections were of extremely low salience. The results are somewhat ambiguous (see Table 2, columns 9 and 10). In Boston, voters appear to be mobilized by the phone calls from the volunteers. The rate of voter turnout in the treatment group was 1.6 percentage points higher than in the control group. Adjusting for the contact rate in Boston yields an estimated mobilization of 2.9 percentage points with a standard error of 2.8.

A low contact rate causes the Seattle experiment to be essentially uninformative, ruling out only very large mobilization effects. A large treatment group is unlikely to vote at higher rates because 10% of the respondents have been called and asked to vote. In Seattle, there appears to have been no mobilization to speak of as a result of the phone calls. Even after adjusting for the rate of contact, the treatment group barely votes at a higher rate than the control group (0.1 percentage points with a standard error of 5.5). The Seattle result is hardly surprising, given the small percentage of phone numbers in the treatment group attempted by the campaign. However, the experiment is an unbiased estimate of the mobilization effect from volunteer phone calls, albeit a very noisy estimate. Pooling the two experiments in 2001, volunteer phone calls are estimated to boost turnout by 2.3 percentage points (with a standard error of 2.5). The results are slightly lower than those found in 2000, but still positive and within a standard error of the 2000 estimate.

Taking together all of the experiments discussed in this article, one would conclude that a call from a volunteer phone bank boosts voter turnout by 3.8 percentage points with a standard error of 1.4 ($p < .01$).[12] That is, for every 1,000 completed calls, 38 votes are created.[13] Placing the 3.8 percentage point increase in turnout in perspective, volunteer phone calls are roughly half as effective as face-to-face meetings and 8 times as effective as a piece of nonpartisan direct mail (Gerber et al., 2003; Gerber & Green, 2000b; Green et al., 2003).

Possessing a relatively precise and unbiased estimate of the effectiveness of volunteer GOTV phone calls, it is now possible to compare the cost-effectiveness of volunteer phone calls to other mobilization technologies (see Table 3). If a volunteer caller can make 15 contacts per hour and is paid an hourly wage of $15, then volunteer phone banks require an investment of $26 to create one vote.[14] Compared to other technologies, volunteer nonpartisan phone calls are relatively cheap. Face-to-face canvassing generates roughly one vote for every $32 spent, leaflets cost between $23 and $47 per vote, and direct mail is much more expensive at $69 per vote. Furthermore, one cannot even calculate a dollar-per-vote estimate for professional nonpartisan phone calls because there appears to be no mobilization effect (Gerber & Green, 2000b, 2001). In short, volunteer phone calls are not only effective at mobilizing voters, but they are also cost-effective.

Discussion

The core finding of this article is that nonpartisan, volunteer GOTV phone calls are effective at mobilizing voters. The results of these experiments do not necessarily contradict the null findings in the 1998 New Haven and West Haven experiments (Gerber & Green, 2000b; Green & Gerber, 2001), because the types of phone banks considered are very different. It is entirely possible that professional phone banks are only marginally effective, whereas volunteer phone banks produce movement toward the polls. Thus, Gerber and Green overstate their case when they conclude that phone calls are not an effective means of boosting voter turnout. The correct conclusion is that the professional phone calls tested by Gerber and Green are ineffective, and the volunteer phone calls tested here boost turnout beyond any reasonable threshold for statistical significance.

A secondary finding of the eight experiments was that the voters registered by the students making the GOTV calls were no more motivated to vote than people who never had any previous contact from the campaign. The implication for political campaigns is that purchasing lists can be an effective means of gathering the names of supporters. That is, it is not necessary to organically grow lists of people to mobilize. However, campaigns need to ensure that the contact information contained on the lists is high quality. The student-generated lists were far more accurate than the vendor-supplied lists, allowing for more efficient calling campaigns.

As a side note, it should be noted that the results from the 2000 YouthVote experiments casts doubt on the hypothesis that phone calls can only be effec-

Table 3
Cost-Effectiveness of Experimentally
Tested Nonpartisan Voter Mobilization Technologies

Technology	Source	Boost per Contact (percentage points)	No. of Contacts per Hour	Wage Rate	$ per Vote
Volunteer phone	see Table 2	3.8	15	$15	26
Professional phone	Gerber and Green (2000b, 2001)	0.0	na	$1 per completion	na
Door-to-door canvassing	Gerber and Green (2000b); Gerber, Green, & Nickerson (2003)	7.8	6	$15	32
Leaflets	Gerber and Green (2000a)	0.8	40	$15	47
Direct mail	Gerber and Green (2000b)	0.6	na	$0.40 per mailing	69

Note: na = not applicable.

285

dential election was tightly contested and featured high levels of voter turn-out (by American standards). Indeed, the treatment and control groups in five of the six experiments conducted in 2000 featured rates of turnout above 50%. Not only was turnout high, many of the sites featured active campaigns and the accompanying media attention. Oregon featured one of the tightest presidential contests in the country. Although Colorado and New York were not tight in the presidential race, New York faced a high-profile senatorial race and Colorado was voting on a medicinal marijuana bill (which was salient among people younger than 30 in the Boulder area). With a standard error of 1.7, the 4.5 percentage point boost in turnout from a phone call in 2000 is not due to chance and demonstrates that volunteer phone campaigns can be effective in high-turnout, high-activity races.

The experiments in this article are slightly idiosyncratic in a couple of ways. First, the calling was strictly nonpartisan. Callers did not encourage respondents to vote for a particular candidate. The vast majority of voter mobilization calls in the United States are made on behalf of parties and candidates encouraging the citizen to vote in a particular manner. Such partisan phone calls may behave in a different manner than nonpartisan calls and deserve to be studied in their own right.

Second, calls were made the weekend prior to the election and were thus small in scope. Large phone campaigns typically begin several weeks prior to the election to facilitate contacting large numbers of voters. Intuition suggests that the mobilization tactics are less effective the further away from the election they are used. In support of this hypothesis, the two biggest mobilization effects detected in this article occurred at sites where calling occurred only the day prior to election day (in Stonybrook and Albany). However, it is impossible to disentangle the unique effects of the population being mobilized, the political context, quality of the callers, and the timing of the call. To determine the role timing plays in the effectiveness of telephone calls, an experiment where citizens are randomly selected to be called at different times would be informative.

Finally, the callers were all college students active in a student group. It is possible that the nature of the caller matters a great deal when mobilizing voters. In 2000, the target population was extremely similar to the callers (i.e., college students). In 2001, most of the respondents called were older than the callers and not students. The mobilization effect detected was smaller in 2001 than 2000, so perhaps the fit between caller and voter matters. The difference between the 2 years is not statistically significant, and the match between caller and target should be examined systematically in future experiments.

Another direction for future research is determining the reason for the difference in performance between professional and volunteer phone calls. A few different factors could account for the effectiveness of volunteer phone banks when professional phone banks have proven to be worthless. First, the speed with which the call is delivered may account for the difference. Professional callers typically have an incentive to complete as many calls as possible in an hour. In contrast, volunteers tend to put emphasis on making a personal connection with the person on the other end of the line. The more pleasant and calm tone of the conversation may be persuasive to voters. It should be noted that the volunteer calls were very short. However, well-placed pauses of even a second can make a call seem unhurried and relaxed.

Second, calls from professional phone banks are standardized, and callers are trained to adhere strictly to a script. In contrast, the volunteer phone banks considered in these experiments did not use scripts and callers developed their own dialogue. The flexibility allows the caller to adapt the tactic over the course of an evening and tailor the conversation to an individual person. It also allows the caller to find a pitch that works for her, which may be important because different callers will possess different strengths. Adaptability may account for the improved performance of volunteer phone calls over professional phone calls.

Third, the people called were able to ask questions of the callers. Because the callers resided in the same area as the voters, the callers were able to answer many basic questions about polling places and candidates. The information in itself might have motivated respondents to turn out, but the possibility for interaction probably also helped to make the phone call seem more personal.

Finally, the difference may be technological. Professional call centers use automatic dialers, where recipients hear a brief pause before being connected to the caller. The phone banks in this set of experiments did not use automatic dialers and volunteers dialed the numbers themselves. Many voters hang up the phone during the pause to avoid speaking to a telemarketer. For those who do stay on the line, the pause prior to being connected may trigger a negative response from voters who associate the sound with telemarketing calls. If the pause causes the ineffectiveness of professional phone banks, then political campaigns would be well advised to revert to the older and less efficient tactic of manually dialing numbers.[15]

None of these three reasons for the superiority of volunteer phone campaigns is intrinsic to the nature of the volunteer campaign. Professional phone banks do possess some advantages over volunteers and could perhaps be coached to improve results. Professional phone callers are well trained

and very practiced at calling and persuading strangers. Presumably, the talent pool in professional phone firms has been selective, so the worst callers have quit or been fired. Further work is needed to determine whether professional phone banks simply need to change tactics and mimic volunteer phone banks or whether volunteers are inherently better at mobilizing voters than professionals.

The cost-effectiveness of the volunteer phone calls as a mobilization strategy should not be taken to mean that political campaigns should rely upon them to the exclusion of more expensive tactics. First, phone numbers will not be available for every voter in a district. Second, contact rates for face-to-face canvassing and phone banks will vary across political contexts: A volunteer going door-to-door will typically contact 3-8 people an hour; a caller can reach 15-20 people per hour over the phone. In areas where a door-to-door canvasser will make very few contacts per hour, a volunteer phone bank is clearly a superior technology for mobilizing voters. Similarly, when few phone numbers are available to the campaign and/or houses are densely situated, then door-to-door canvassing deserves the emphasis from the campaign. Campaign managers need to adapt tactics to fit the reality on the ground.

In the 2004 presidential election, the Republicans realized the utility of volunteer voter mobilization calls and made extensive use of the tactic. For instance, in Nevada 78% of campaign contacts were from volunteers via phone (Bumiller, 2004). What made such extensive grassroots outreach possible was the army of 1 million Republican volunteers. The Democrats were at a competitive disadvantage because only 250,000 volunteers were recruited (Balz & Edsall, 2004). Facing a large deficit in labor power, the Democrats faced three options: (a) Ask volunteers to make a large number of phone calls, perhaps rushing the conversations and lowering the quality of each call, (b) pay professional phone banks to make phone calls, or (c) simply make fewer calls than the Republicans. The results from these eight experiments suggest that all three solutions placed the Democrats at a competitive disadvantage. Thus, the difference in volunteer labor power may account for a large part of the difference in turnout between the two parties in 2004.

Appendix
Instructions and Script for Boulder, Colorado 2000

Thanks for volunteering to help mobilize voters. Young people don't receive much attention from politicians and your call can help to make sure the youth vote is noticed. Making GOTV [get-out-the-vote] calls is easy and fun to do. We ask only two things:

1) Do not express any opinion about candidates or ballot propositions. Youth Vote is a non-profit organization (501c(3)) and we are not allowed to endorse political parties. If a person asks what you think, just say "I'm not allowed to tell you my view. We just want to make sure that you vote and have your unique voice heard." You can answer very basic questions, but if a person asks for information about the candidates direct them to the websites www.vote-smart.org (make sure you say vote-dash-smart.org) and www.youthvote.org.

2) Please keep careful track of how each call goes. Mark the box next to the name and number that best describes what happened during the call. For instance, if you spoke and left a message with a roommate, put an "X" in the box that says "Roommate."

The key to successful calling is to be as friendly and personal as possible. Try to engage the person you are calling in conversation. That means asking questions and waiting for answers. A sample script is provided below to help guide you through your first few calls, but you should develop your own rap. Find out what works for you and go with it (just be sure to be non-partisan).

Sample Script

Hi, can I speak to (full name)?

Hi (first name), my name is (your full name) and I'm calling on behalf of the Youth Vote Coalition. How are you tonight? (wait for reply)

I'm not calling to sell you anything or ask for money. I'm just calling you to encourage you to vote this Tuesday. We don't care who you vote for, but we think it is important that your voice is heard. The Presidential election is close and there are two interesting initiatives on the ballot that young people might care about. [Note: There is a "smart growth initiative" and a "medicinal marijuana" proposition.]

Do you know where to vote on Tuesday? (wait for reply)

[If polling information is available] We have your polling place listed as (read polling place).

The polls are open from 7am to 7pm, and I hope you help the youth vote be heard. Thanks for your time and have a good night.

(If they have questions, they can call the Boulder County Elections Office at 303-413-7740.)

Notes

1. The New Haven result is the corrected version reported on Donald Green's Web site (see Gerber & Green, 2002).

2. Imai (2002) catches a small merging error in Gerber and Green (2000b) that does not substantively change the null finding from the experiment (see Gerber & Green, 2002). There was no corresponding coding error in Gerber and Green (2001). Imai's substantive suggestion of replacing experimental analysis with a matching procedure leads to biased estimates, because he compares people contacted by the phone campaign to respondents in the control group who may or may not be the type of person who can be contacted by the campaign (see Arceneaux, Gerber, & Green, 2004).

3. There are many reasons a campaign may be unable to reach a voter via phone: disconnected phone numbers, call screening, hang-ups, persistently busy lines, movement among voters.

4. Because the control group is never called, γ can only be inferred for the control group. The law of large numbers assures that the ratio of reachable and nonreachable persons is the same in the treatment and control groups.

5. Usually nonviolent drug offenses.

6. The rolling control group used in Boston was also attempted in Seattle. Unfortunately, a correlation exists between past voting history and being near the top of the "to call" list. The source of the correlation is unknown and may be due to random chance. However, past voter history is not correlated with the original assignment to the treatment and control groups, so the original monolithic treatment and control groups are used in the present analysis. Attempting to control for past voter history and rolling respondents into the control group yields an estimate of 3.0% with a standard error of 1.4%. However, the model may not accurately capture the process generating the data (i.e., the correlation between voter history and list ordering), and the results are highly suspect. Thus, the analysis makes the most conservative possible assumptions and uses only the original assignment to treatment and control.

7. Precision-weighted averaging takes into account the uncertainty associated with a particular estimate. The average β and associated σ^2 for two experiments would be calculated as follows:

$$\beta_{average} = \frac{\beta_1\sigma_2^2}{\sigma_1^2 + \sigma_2^2} + \frac{\beta_2\sigma_1^2}{\sigma_1^2 + \sigma_2^2} \text{ and } \sigma^2_{average} = \frac{\sigma_1^2\sigma_2^2}{\sigma_1^2 + \sigma_2^2}.$$

8. If the true treatment effect was zero, one would expect to see all five intent-to-treat effects be positive due to random chance only 3 times in 100.

9. Omitting Stonybrook from the analysis alters the results modestly. The estimated treatment effect for the 2000 experiments would be 3.1 percentage points with a standard error of 2.0 ($p < .06$). The overall treatment effect for all eight experiments would be 2.8 percentage points with a standard error of 1.6 ($p < .04$). The conclusion therefore reached would remain that volunteer phone calls remain an effective means of increasing turnout, but the precise estimate would be 1 percentage point less than the reported finding.

10. Omitting the students registered by the Oregon Student Association in Eugene from the analysis alters the results modestly. The estimated treatment effect for the 2000 experiments would be 5.5 percentage points with a standard error of 1.9 ($p < .01$). The overall treatment effect for all eight experiments would be 4.3 percentage points with a standard error of 1.5 ($p < .01$). The conclusion therefore reached would remain that volunteer phone calls remain an effective means

of increasing turnout, but the precise estimate would be 1 percentage point greater than the reported finding.

11. The pooled result is calculated via precision-weighted averages. A conditional shrinkage model yields results identical to the numbers reported here.

12. Adding covariates for the experiments conducted in 2000 would not provide much information. The population consisted overwhelmingly of first-time voters who lived close together within student neighborhoods. Covariates that are traditionally used to explain levels of voter turnout simply do not possess enough variance to be informative.

13. The results from John McNulty's (in press) small, nonpartisan, volunteer phone voter mobilization experiment, 5.4 percentage point boost in turnout with a standard error of 3.0, are consistent with this finding. Pooling the two results, one would conclude that volunteer get-out-the-vote (GOTV) calls boost turnout by 4.1 percentage points with an associated standard error of 1.3.

14. In common parlance, a paid volunteer is oxymoronic. However, political campaigns often need to use monetary selective benefits to recruit labor. Paid volunteers refer to campaign workers who work piecemeal and are not professional telemarketers or campaign consultants. One could also use the $15 as a stand-in for the value of each hour worked by a volunteer.

15. This hypothesis would be easy to test. The phone numbers of registered voters could be divided into three groups: control, called using automatic dialer, or called using manual dialing. The experiment could then consider the differences in both the contact rates of the two strategies and the ultimate effect upon voter turnout.

References

Adams, W. C., & Smith, D. J. (1980). Effects of telephone canvassing on turnout and preferences: A field experiment. *Public Opinion Quarterly*, *44*, 389-395.

Arceneaux, K. T., Gerber, A. S., & Green, D. P. (2004, August). *Comparing experimental and matching methods using a large-scale field experiment on voter mobilization*. Paper presented at the annual meeting of the American Political Science Association, Chicago.

Balz, D., & Edsall, T. B. (2004, November 1). Unprecedented effort to mobilize voters begins. *The Washington Post*, p. A1.

Bennett, S. E. (1991). Left behind: Exploring declining turnout among non-college young Whites, 1964-1988. *Social Science Quarterly*, *72*, 314-333.

Bumiller, E. (2004, November 4). Turnout effort and Kerry, too, were G.O.P.'s keys to victory. *The New York Times*, p. A1.

Eldersveld, S. J. (1956). Experimental propaganda techniques and voting behavior. *American Political Science Review*, *50*, 154-165.

Gerber, A. S., & Green, D. P. (2000a). The effect of a nonpartisan get-out-the-vote drive: An experimental study of leafleting. *Journal of Politics*, *62*, 846-857.

Gerber, A. S., & Green, D. P. (2000b). The effects of canvassing, telephone calls, and direct mail on turnout: A field experiment. *American Political Science Review*, *94*, 653-663.

Gerber, A. S., & Green, D. P. (2001). Do phone calls increase turnout? *Public Opinion Quarterly*, *65*, 75-85.

Gerber, A. S., & Green, D. P. (2002, July 10). A note of correction to results published in: Gerber, Alan S., and Donald P. Green. 2000. The effects of canvassing, direct mail, and telephone contact on voter turnout: A field experiment. *American Political Science Review*, *94*, 653-663. Retrieved from http://research.yale.edu/vote/CORRECTEDAPSRREPLICATION TABLES.HTM

Gerber, A. S., Green, D. P., & Green, M. N. (2003). The effects of partisan direct mail on voter turnout. *Electoral Studies*, *22*, 563-579.

Green, D. P., & Gerber, A. S. (2001). *Getting out the youth vote: Results from randomized field experiments*. Unpublished report to the Pew Charitable Trusts, Institution for Social and Policy Studies.

Green, D. P., Gerber, A. S., & Nickerson, D. W. (2003). Getting out the vote in local elections: Results from six door-to-door canvassing experiments. *Journal of Politics*, *65*, 1083-1096.

Hayes, T. J., Dwyer, F. R., Greenwalt, T. J., & Coe, N. A. (1984). A comparison of two behavioral influence techniques for improving blood donor recruitment. *Transfusion*, *24*, 399-403.

Imai, K. (2002, July). *The importance of statistical methodology for analyzing data from field experiments*. Paper presented at the annual meeting of Political Science Methodology, Seattle, WA.

McNulty, J. E. (in press). Phone-based GOTV: What's on the line? Field experiments with varied partisan components. *Annals of Political and Social Science*.

Miller, R. E., Bositis, D. A., & Baer, D. L. (1981). Stimulating voter turnout in a primary: Field experiment with a precinct committeeman. *International Political Science Review*, *2*, 445-460.

A new Internet democracy? (2003, July 26). *The Economist*, *368*(8334), 32-33.

Nickerson, D. W. (in press-a). Hunting the elusive young voter. *Journal of Political Marketing*.

Nickerson, D. W. (in press-b). Scalable protocols offer efficient design for field experiments. *Political Analysis*.

Polman, D. (2004, November 1). In Ohio: A political fight like none before. *Philadelphia Inquirer*, p. A1.

Reams, M. A., & Ray, B. H. (1993). The effects of three prompting methods on recycling participation rates: A field study. *Journal of Environmental Systems*, *22*, 371-379.

Rosenstone, S. J., & Hansen, J. M. (1993). *Mobilization, participation, and democracy in America*. New York: Macmillan.

David W. Nickerson is an assistant professor of political science at the University of Notre Dame.