Mobilizing African-American Voters Using Direct Mail and Commercial Phone Banks: A Field Experiment

Donald P. Green
Department of Political Science
Yale University

January 2, 2003

Abstract: This essay summarizes the results of a large-scale randomized experiment conducted during the 2000 election campaign by the NAACP National Voter Fund, which sought to mobilize African-American voters. The study focuses primarily on 980,280 participants residing in single-voter households, 1.7% of whom were randomly assigned to a control group. The experiment permits us to estimate (1) the extent to which the National Voter Fund's phone calls and direct mail increased voter turnout and (2) the approximate cost per vote. The NVF's two pieces of GOTV mail increased turnout minimally, if at all. Three live and two recorded phone calls had modest and statistically insignificant effects, raising turnout by approximately 2.3 percentage-points.

Introduction

But for a few hundred votes cast in Florida, the 2000 presidential election might have been remembered as one in which voter mobilization efforts, spearheaded by labor unions and African-Americans, carried the day. Trailing in the polls days before the November 7th election, Democratic candidate Al Gore went on to win the popular vote and carry several closely contested populous states, such as Pennsylvania and Michigan. Attributing Gore's success to voter mobilization, strategists of both parties have taken a renewed interest in get-out-the-vote (GOTV) campaign activity, and it has become an active topic of both political discussion and scholarly research.

Building on previous experimental studies of voter mobilization campaigns, this paper examines the role of two widely used techniques for mobilizing voters, direct mail and calls from commercial phone banks. During the 2000 general election campaign, the NAACP-National Voter Fund (NVF) alone sent out a total of 6,697,755 mailings and completed 5,246,225 phone calls. Several recent studies have examined the effectiveness of campaign mail and phone calls. Using a randomized experimental design and very large samples, Gerber and Green (2000) found nonpartisan direct mail to have weak positive effects on voter turnout. Gerber, Green, and Green (2003), summarizing several large-scale experiments, found partisan mail to have little if any mobilizing influence. Gerber and Green (2000, 2001) found nonpartisan appeals made by commercial phone banks to be ineffective, although phone calls from noncommercial phone banks do seem to stimulate voter turnout (Green and Gerber 2001).

The present study fills an important gap in this emergent literature. Here we examine the effectiveness of a mobilization appeal expressed by and directed to African-Americans. The NVF crafted its appeal so that it would speak to the concerns of African-Americans (e.g., hate crime, racial profiling, educational opportunity). The phone bank's callers were African Americans. While these messages did not explicitly advocate the election of a particular candidate, they were on the whole much more partisan than the League of Women Voters' appeals studied by Gerber and Green (2000, 2001). And unlike the partisan mail that has been studied in previous research (Gerber, Green, and Green 2002), the mail sent by the NVF was part of a campaign that focused primarily on mobilizing an audience whose preferences lopsidedly favored one candidate.

What makes this study of direct mail and phone calls particularly interesting is that these are not GOTV techniques that are ordinarily associated with the mobilization of African-American voters (Leighley 2001). African-Americans, by comparison to the population as a whole, have lower proportions of listed telephone numbers. Their lower rates of homeownership and higher levels of residential mobility make them more elusive targets for direct mail. Although the NVF also devoted its energies to mass media appeals and to door-to-door canvassing, millions of dollars were allocated to phone calls and direct mail. The question is whether these staples of contemporary campaigning are effective in mobilizing African-American voters.

This study uses randomized field experimentation to address this issue. Registered voters were randomly assigned to treatment and control conditions. Those in the treatment group were encouraged to vote by means of direct mail and phone calls. After the election, public records were used to assess the voter turnout rates of the treatment and control groups. With nearly one million participants, this study represents the largest field experiment ever conducted in political science.

Unfortunately, it is also a study that confronted a number of operational problems. In some states, certain members of the control group received calls from the NVF phone bank. Some members of the treatment group were never called for lack of a listed phone number, and some were never mailed because their names were added to the master list after the mailings had already been sent. Fortunately from an analytic standpoint, these types of complications have become familiar to statisticians in other experimental disciplines, such as public health, where medical treatments frequently find their way into the hands of the control group rather than the treatment group. Our statistical analysis makes explicit allowance for these infirmities and takes care to produce robust inferences.

Had the experiment been executed flawlessly, the data would speak to the effectiveness the NVF's campaign with a very high degree of precision. As it stands, the experiment leaves some room for uncertainty but nonetheless supports a clear conclusion: direct mail was ineffective; phone calls were marginally effective; and neither was an economically efficient means for producing votes. The principal advantage of direct mail and phone banking seems to be that it enables campaigns to get their messages to very large numbers of people with short lead time.

Study Description

Scope and nature of the mobilization campaign. In the months leading up to the November election, the National Voter Fund purchased data files containing the names, addresses, and in some cases phone numbers of more than 3.7 million registered voters presumed to be African-American.[1] We focus here on approximately 1 million registered voters living in ten states for which we have records of voter turnout in 2000. For reasons noted below, each of these voters is also the only registered voter living at a given address.

During the last four weeks of the campaign, the GOTV campaign attempted to call every person with a known phone number. The calling campaign consisted of three conversations with African-American callers employed by a commercial phone bank as well as two automated calls made in the waning days of the campaign. Each call was intended to motivate voters, and calls made toward the end of the campaign placed special emphasis on encouraging voter turnout. The phone script read between October 17-22 contained the following appeal:

[1] In states covered by the Voting Rights Act, race was ascertained using a racial identifier in the registration list. In other states, race was inferred based on the density of African-Americans residing in a given Census tract.

> We have a lot at stake in this year's election – good schools, safe neighborhoods, jobs, as well as hate crimes and racial profiling – all of these issues will be decided by who we choose to elect. It's simple – if you vote, you decide your future and your community's future. If you don't vote, you lose that chance, and someone else will decide. Can the NAACP National Voter Fund count on you to go to the polls on November 7?

The script read during the period October 28-November 1 took a sharper partisan tone:

> I am calling today on behalf of the NAACP national voter fund to discuss a serious issue: hate crimes. African Americans and other minorities have been singled out by groups and individuals as the targets of many violent and despicable acts. In Texas, James Byrd was kidnapped, chained to the bumper of a pickup truck and dragged three miles to his death. When his daughter begged Governor George Bush to support a hate crimes bill he refused. Call governor George Bush to support a hate crimes bill at 512-463-2000. Tell him that we won't let our community be dragged down. Tell him to support hate crimes legislation.

On November 3-4, the phone calls switched to recorded messages. One, from radio personality Tom Joyner, advertised local rallies. A second, from President Bill Clinton, contained the following text:

> Hello, this is President Clinton calling for the NAACP National Voter Fund, urging you to vote on November 7. We have made a lot of progress in the last few years, but there is more to be done. To make sure our communities have better schools, safer neighborhoods, and more and better jobs, your voice must be heard – and it will be if you vote. Make no mistake about it – issues like racial profiling and hate crimes will be decided at the national, state, and local level by the people we elect. I hope the National Voter Fund and I can count on you to vote on November 7.

Finally, a 15 second reminder was delivered by callers from the phone bank on November 7:

> Hello, this is the NAACP National Voter Fund, reminding you to vote today. If you vote, you decide your future and your community's future. If you don't, someone else will decide. It's important to vote all the way down the ballot. Please remember to vote on November 7.

Direct mail was sent to selected individuals based on their age and past voting history. The selection criteria varied by state, as shown in Table 1. Direct mail targeted young and middle-aged voters who had abstained from voting during one or more recent elections.[2] Prior to random assignment, approximately one-third of the list was sent a postcard emphasizing the importance of voter registration. This postcard was designed to

[2] The "propensity" score assigned by the campaign is simply the fraction of recent elections in which a person voted. The number of recent elections listed in the registration records varies by state, which accounts for the state-to-state variation in the criteria used.

weed out bad addresses. Another mailer dealing with the subject of hate crime was sent to roughly one-third of the list. Since our treatment and control groups are equally likely to have received these preliminary mailings, they play no direct role in our analysis. One may argue, however, that the initial mailings diminish the influence of subsequent mailings. We will take up this concern later on when examining the effects of the experimental intervention on those who received no preliminary mail.

[insert Table 1 here]

The GOTV mail that is the subject of this experimental analysis focused on the issue of discrimination. One mailer, for example, depicted a black motorist stopped by a white police officer brandishing a nightstick. The text surrounding the photo reads, “Stopped again? Start voting.” The text on the reverse side reads:

> Getting stopped for “Driving While Black” is swerving out of control. We all know someone. Or we’ve had to deal with the frustration of it. And racial profiling hasn’t just stopped there. Now’s the time to do something about it. Vote. Let’s start pulling the ballot lever and stop getting pulled over. Stop the madness. Vote. And get in the real driver’s seat. Vote on Election Day, Tuesday November 7.

In certain states, listed in Table 1, voters who had not earlier received a mailing on hate crime were sent a different version of this mailing, which again implored them to “Stop the madness. Write Governor George W. Bush and tell him that we need hate crime legislation." The phone and mail campaigns thus struck similar themes, and indeed were designed to reinforce each other insofar as the target population received both types of messages.

Method of assessment. In order to gauge the effectiveness of phone calls and direct mail, a random subset of the master list was assigned to a control group. This control group comprised 1.7% of the total list. This rate is virtually identical across the ten states analyzed here. Because people in the control group were selected at random, they were nearly identical to the treatment group in terms of age, past voting history, and other background characteristics. For example, among the 980,208 voters in our sample, the experimental treatment correlates -.001 with age, .001 with abstaining in at least one past election, .000 with whether the voter's phone number could be identified by a commercial vendor a year after the election, and -.001 with whether the voter was sent a preliminary mailing.[3] None of these correlations is significant at $p < .10$ using a two-tailed test. Since the treatment and control groups differ only by chance in advance of the experimental treatment, the effectiveness of the phone and mail campaign may be assessed by comparing turnout rates in the treatment and control groups.

Information about whether each person in the treatment and control group voted was compiled by the Aristotle International organization based on their own proprietary data. Voting records were merged to the existing data file by means of voter registration

[3] The correlation between experimental assignment and age was based on 963,010 cases, since information about age was unavailable for some participants.

numbers, names, and addresses. In order to check the accuracy of the results, we compared the voter history supplied by Aristotle International in selected jurisdictions to another commercial list vendor and to data supplied by a registrar of voters. The results indicate that Aristotle International reliably recorded voter turnout in 2000.

Operational problems. The experiment was hampered by a number of deviations from an ideal experimental design. First, phone calls were inadvertently made to some of the people in the control group. This mistake occurred when lists of voters were shipped directly from the list vendors to the phone banks; the proper procedure was to send lists to the central data processing firm, which would have extracted a control group before sending the data on to the phone bank. Second, tens of thousands of names in the treatment group were neither called nor mailed. The lack of contact from the campaign in many cases occurred because vendors were unable to supply a person's working phone number. Although it is tempting to focus our data analysis solely on those people for whom a working number could be found, a third problem prevents this: the central processing firm, contrary to our instructions, did not attempt to obtain matching phone numbers for subjects in the control group. Thus, we do not have a matching control group for those with working numbers. Fourth, randomization was performed at the level of the individual name rather than the household. Since the chances of being assigned to the control group were just 1.7%, only a tiny fraction of those living in households with two or more voters were assigned entirely to the control group. This report therefore focuses on those households containing a single registered voter. This approach reduces the number of observations in our study and attenuates somewhat its generalizability to the African-American population as a whole.

[insert Table 2 here]

Despite these difficulties, it remains possible to tease out the effects of the experimental treatments so long as we make allowances for the mis-assignment of treatments. Table 2 shows the rate at which the treatment and control groups in the phone universe and the phone/mail universe were contacted by the campaign in each state. The fact that some of the treatment group never actually received the mail or phone treatments complicates the interpretation of the data, as does the fact that some of the control group received phone calls. Any observed differences in turnout between the treatment and control groups must be adjusted by a factor known as the *contact rate*. The contact rate is defined as the difference between the percentage of the treatment group who received the phone/mail minus the percentage of the control group who received the phone/mail. When an experiment treats everyone in the treatment group and no one in the control group, the contact rate is 100%, and no adjustment is required. In our study, contact rates are well below 100%. In Pennsylvania, for example, just 31% of the treatment group received a phone call as opposed to 11% of the control group, yielding a phone contact rate of 20%. Similarly, just 46% of the treatment group received the experimental mailings; since none of the control group received these mailings, the mail contact rate is 46% – 0% = 46%.

In order to calculate the effects of direct mail or phone calls among those who actually receive them, one divides the observed difference in turnout between treatment and control groups by the contact rate (Angrist, Imbens, and Rubin 1996; Gerber and Green 2000). This approach is identical to a two-stage least squares regression in which turnout is regressed on actually receiving the mail or calls, with random assignment as the instrumental variable. The one complication is that participants were not randomly assigned to combinations of mail and phone calls; rather, as noted in Table 1, they were assigned to a single regimen based on their demographic profile and past voting history. In order to distinguish between the effects of mail and phone calls, we make use of the fact that some participants were ineligible for mailings and that phone contact rates varied widely across states.

Results

Table 2 summarizes the pattern of treatments and turnout in each state. In Florida, Michigan, Missouri, and Pennsylvania, roughly half of the treatment group received the experimental mailings. In the remaining states, the mail contact rate was lower, and in North Carolina and Virginia it was zero. Phone contact rates also varied from a high of 30% in North Carolina to a low of 4% in Ohio. The large samples in Georgia, Michigan, and Pennsylvania make them useful, although their phone contact rates range from 20% to 26%. The last column of Table 2 shows the estimated intent-to-treat effects in each state; this figure is simply the observed turnout difference between treatment and control groups. These estimates hover around zero. Four of the ten are negative, and the largest estimate occurs in the state with the smallest control group. The overall intent-to-treat effect may be estimated by an OLS regression of voter turnout on a dummy variable marking whether participants were assigned to the treatment or control group, using dummy variables indicating each state as covariates. This regression yields a 0.3 percentage-point effect with a standard error of 0.4. Based on this estimate, one would infer that direct mail and phone campaigns produced an additional 3300 votes among the 963,496 names assigned to the treatment group.

The size of the intent-to-treat effect reflects both the weakness of the stimulus and the infrequency with which voters were contacted by mail or phone. In order to tease apart these two explanations, we disaggregate the data according to whether participants were eligible to receive direct mail. Tables 3 and 4 present the contact and turnout rates for each subgroup. The group that was ineligible for mail, as expected, voted at higher rates than its mail-eligible counterpart, but neither group shows a consistent pattern of positive experimental effects. Four of the ten intent-to-treat estimates are negative in Table 3, and four of eight are negative in Table 4. We will return momentarily to a more detailed statistical analysis of these two tables.

[insert Tables 3 and 4 here]

In an effort raise the statistical precision of these comparisons, we restricted our attention to those names for which a matching phone number could be found in the fall of 2001, a year after the experiment occurred. Although it may seem odd to append data after the experiment was conducted, the match rates for those assigned to the treatment and control groups was identical. Tables 5 and 6 show the contact and turnout rates for the subgroup with matching phone numbers. As expected, this phone contact rate is higher for this subgroup. Although we lose observations by restricting the sample in this way, we gain statistical precision by increasing the contact rate. The intent-to-treat estimates in Table 5 change a bit from Table 3, with Georgia and New Jersey's effect rising and Michigan and Pennsylvania's falling, but the overall pattern remains similar across tables. Six of the ten intent-to-treat estimates are greater than zero in Table 5. Four of the eight intent-to-treat estimates in Table 6 are greater than zero.

[insert Tables 5 and 6 here]

Multivariate analysis. Table 7 summarizes the statistical analysis of actual treatment effects. These estimates are obtained using two-stage least squares. Using the subscript *i* to refer to individuals, the second-stage equation may be expressed:

$$V_i = b_0 + b_1 P_i + b_2 M_i + b_3 D_{FL} + \ldots + b_{11} D_{PA} + \ldots + b_{12} D_{FL}E_i + \ldots + b_{19} D_{PA}E_i + u_i.$$

The notation is as follows:

V_i is scored 1 if the participant voted, 0 otherwise.

P_i is scored 1 if the participant was called by the phone bank, 0 otherwise.

M_i is scored 1 if the participant received experimental mail, 0 otherwise.

D_k is scored 1 if the participant resides in a specified state.

E_i is scored 1 if the participant was eligible to receive experimental mail, 0 otherwise.

u_i is an unobserved disturbance term.

In this equation, b_3 through b_{19} represent intercepts marking each statewide group of eligible and ineligible participants. These coefficients, which are all statistically significant due to the enormous sample sizes, have little theoretical meaning and will be ignored for purposes of presentation. The key coefficients of interest are b_1 and b_2, which represent the effect of receiving phone calls and mail, respectively.

[insert Table 7 here]

The first-stage equations model the receipt of phone calls and mail.

$$P_i = a_0 + a_1 T_i + a_2 T_iE_i + a_3 D_{FL} + \ldots + a_{11} D_{PA} + \ldots + a_{12} D_{FL}E_i + \ldots + a_{19} D_{PA}E_i + e_{i.}$$

$$M_i = c_0 + c_1 T_i + c_2 T_iE_i + c_3 D_{FL} + \ldots + c_{11} D_{PA} + \ldots + c_{12} D_{FL}E_i + \ldots + c_{19} D_{PA}E_i + q_{i.}$$

Here, T_i is scored 1 if the participant was randomly assigned to the treatment group, 0 otherwise. The system of equations has two endogenous regressors (P_i, M_i) and two instrumental variables (T_i, T_iE_i), which permits the estimation of b_1 and b_2, the actual treatment effects.

The first column in Table 7 shows the results for the entire sample. Those receiving calls from the phone bank were 2.3 percentage-points more likely to vote. The standard error of this estimate, however, is also 2.3 percentage-points, which means that the estimate cannot be distinguished statistically from zero. The estimated effect of mail is small and statistically insignificant (-0.2 percentage-points, SE=0.8).

The apparent effect of phone calls increases slightly when we add controls for age, past voting history, and whether a matching phone number could be found a year later. The estimate (2.7) remains statistically insignificant, however, as does the effect of direct mail (-0.3). A similar estimate of the effectiveness of phone calls is obtained when we restrict the sample to those who are ineligible to receive direct mail. Those who received no prior direct mail do seem to have been a bit more responsive to the regimen of mail that they received as part of the experimental treatment. This estimate (1.0 percentage-points, SE=1.2) is the only hint we find that mail increased turnout. This estimate implies that each of the three mailers increased turnout by roughly one-third of a percentage-point. Again, however, we note that this estimate cannot be distinguished statistically from zero.

Focusing exclusively on those whose phone numbers could be matched a year after the experiment reduces both the phone and mail effects to essentially zero.

Cost per vote. In sum, across a range of alternative models, we find the maximal estimate of b_1 to be 2.7 percentage-points and the maximal estimate of b_2 to be 1.0

percentage-points. In our sample of single-voter households, the number of votes produced may be calculated using these maximal estimates:

$$\text{Total Votes Produced} = 333{,}126 \times .027 + 362{,}416 \times .01 = 12{,}619.$$

In order to calculate the approximate cost per vote, we may conservatively estimate the series of phone calls to have cost $2.00 per participant and the 2-3 pieces of direct mail to have cost $1.25 per participant. These rates imply a total expenditure of $1,119,272, or $89 per vote. Obviously, lowering the effect estimates raises the estimated cost per vote. If we use, for example, the estimates derived from the model that includes covariates (Table 7, column 2), we obtain:

$$\text{Total Votes Produced} = 333{,}126 \times .027 + 362{,}416 \times -.003 = 7{,}907.$$

This figure implies a cost-per-vote of $142. To put this figure in perspective, Gerber and Green (2000) and Green, Gerber, and Nickerson (2003) report that face-to-face canvassing generates one additional voter for each $12-$20 expenditure.

Spillover. One concern about these estimates is that they fail to take into account the possibility that those in the control group are influenced indirectly by the massive phone and direct mail campaigns going on around them. The control group represents 1.7% of the entire NVP list, and if the NVP campaign creates a wellspring of enthusiasm that indirectly affects the control group, the effectiveness of the treatment will be underestimated. One problem with this argument is that even in a state like Pennsylvania, where the campaign took place on a large scale, fewer than half of the people assigned to the treatment group received mail, and only about one in ten received phone calls. These figures are somewhat higher for multi-voter households in the NVF target population. Still, it should be remembered that just 21% (N=168) of the African American respondents to the 2000 National Election Study post election survey report having received direct mail from a group other than a political party and just 8% claim to have been called by such a group.

Nevertheless, the issue of spillover remains an important one, and in anticipation of this problem we performed the following experiment using randomly selected precincts in Missouri, Ohio, New York, and Virginia. Within each selected precinct, 20%, 40%, 60%, or 80% of the names were randomly assigned to the control group. If spillover from treatment to control groups indeed attenuates the estimated effect of the treatment, we should observe stronger treatment effects in these precincts. This hypothesis is not borne out by an intent-to-treat regression of turnout on treatment group, controlling for each state. Among the 3,703 subjects in this experiment (none of whom were included in previous analyses), the intent-to-treat effect is -0.5 (SE=1.7). Including covariates for age, voting history, and matched telephone number raises this estimate to 0.7 (SE=1.6). It does not appear that spillover significantly attenuates the estimated effectiveness of direct mail or phone calls.

Finally, one might wonder how the analysis would have turned out had the study randomized at the level of the household rather than the level of the individual. Consider the case of two-voter households. Of the 437,390 households two-voter households in the primary experiment, just 126 were randomly omitted from any treatment whatsoever, but 14,469 households had just one person assigned to the treatment group. In addition, the spillover experiment described above allows us to compare an additional 334 untreated households to 490 half-treated households and to 755 fully-treated two-voter households. Consistent with our earlier findings, we find meager intent-to-treat effects among the full set of two-voter households. Controlling for state and the rate at which individuals were assigned to treatment and control conditions, we find that households in which neither voter was assigned to the treatment group voted at 1.5 (SE=2.2) percentage-point lower rate than fully-treated households. Half-treated households are nearly indistinguishable from fully-treated households, voting at a 0.3 percentage-point lower rate (SE=0.3). The voting behavior of two-voter households, like that of their one-voter counterparts, suggests a weakly positive treatment effect.

Conclusion

The National Voter Fund's 2000 campaign consisted of many components, such as radio and television advertisements and door-to-door canvassing, that lie outside the scope of this evaluation.[4] The experiment described here focuses solely on the extent to which phone and direct mail increased turnout over and above other factors that might have impelled citizens to vote in the 2000 election.

[4] The NVF estimated that it fielded 9,970 volunteers during the campaign and 5,015 Election Day workers.

These other factors warrant emphasis, because the treatment effects we estimate here are conditional on the volume of communications that voters received from other sources. One potential explanation for why the mobilizing effects of commercial phone calls and direct mail proved to be so weak in this case is that the recipients were showered with communications from candidates, parties, and interest groups. As studies of this sort accumulate, a key empirical question will be whether campaign communications diminish in effectiveness when competing communications abound. If true, this hypothesis would imply that campaign mail and commercial phone calls are much more effective in midterm, off-year, and municipal elections. To date, experimental studies, which have focused disproportionately on non-presidential elections, have not found markedly greater effects for partisan mail and commercial phone banks (Gerber and Green 2000, 2001), but more studies are surely needed to address this question.

A related question concerns the comparative advantages of alternative mobilization techniques. The attractiveness of direct mail and phone banks stems from the fact that they provide political communication on a grand scale with very little lead time. Mailings may be prepared and distributed in a matter of days; phone banks may deploy messages in a matter of hours, which explains why money contributed very late in a campaign tends to be allocated to phone canvassing. Campaigns using door-to-door canvassing may be more cost effective (Gerber and Green 2000; Green, Gerber, and Nickerson 2003; Michelson 2002), particularly in densely populated urban areas, where many of the NVF's efforts centered, but reaching millions of voters in this manner requires a considerable investment in field organization and labor. Because this type of campaign (currently) cannot be subcontracted to commercial firms in the same way that direct mail and phone canvassing can, there is the very real danger that a large scale GOTV effort using door-to-door canvassing might never come to fruition. The NVF's campaign in 2000 might be viewed therefore as a diversified portfolio of mobilization strategies, hedging against the risks and limitations associated with any particular mode of communication with voters.

Finally, the current study paves the way for experimental evaluation of campaign activity by interest groups. While the flaws of the present study underscore the importance of maintaining tight control of the databases from which direct mail and phone calls emanate, the fact that this study was conducted in the first place attests to the willingness of interest groups and their financial backers to undertake a rigorous experimental evaluation. In contrast to most evaluations, this one focuses not on the quantity of campaign contacts but rather on how much these contacts altered voter turnout rates. No single experiment can resolve the myriad questions that could be asked about the effectiveness of alternative mobilization strategies; nevertheless, randomized evaluations such as this one provide a starting point for a scientific analysis of the conditions under which various campaign tactics stimulate voter turnout. Although experimental evaluation is now a rarity, one can readily imagine a point in time where such evaluations are not only standard operating procedure but serve as the basis for performance-based contracts.

References

Angrist, Joshua D. , Imbens, Guido W. and Donald B. Rubin. 1996. Identification of Casual Effects Using Instrumental Variables. *Journal of the American Statistical Association* 91(June): 444-455.

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* 94:653-63.

Gerber, Alan S., and Donald P. Green. 2001. Do Phone Calls Increase Turnout? *Public Opinion Quarterly* 65 (Spring): 75-85.

Gerber, Alan S., Donald P. Green, and Matthew N. Green. 2003. The Effects of Partisan Direct Mail on Voter Turnout. *Electoral Studies* (in press).

Donald P. Green, Alan S. Gerber, and David W. Nickerson. 2003. Getting Out the Vote in Local Elections: Results from Six Door-to-Door Canvassing Experiments. *Journal of Politics* (in press).

Leighley, Jan E. 2001. *Strength in Numbers? The Political Mobilization of Racial and Ethnic Minorities.* Princeton: Princeton University Press

Michelson, Melissa R. 2002. "Getting Out the Latino Vote: How door-to-door canvassing influences voter turnout in rural Central California." Paper presented at the annual meeting of the Western Political Science Association, Long Beach, CA, March 22-24.

Table 1
Experimental Treatment Groups, by State

State	Control N	Treatment N	Criteria Used to Determine Eligibility for Experimental Mailings	Percentage of Sample Sent Pre-Experimental Mailing on Hate Crime*	Type of Experimental Mailings Sent**
FL	2,038	119,070	Ages 25-65, Voting Propensity Less than 100%	31	2 GOTV, Hate Crime
GA	4,124	236,316	Ages 25-55, Voting Propensity Less than 61%	42	2 GOTV Mailings
MI	2,516	144,948	Ages 25-65, Voting Propensity Less than 100%	38	2 GOTV, Hate Crime
MO	389	23,594	Ages 25-65, Voting Propensity Less than 100%	23	2 GOTV, Hate Crime
NC	244	13,666	No Experimental Mailings	30	
NJ	483	29,186	Ages 25-55, Voting Propensity Less than 51%	35	2 GOTV Mailings
NY	3,082	175,887	Ages 25-55, Voting Propensity Less than 51%	26	2 GOTV Mailings
OH	1,241	71,377	Ages 25-55, Voting Propensity Less than 51%	26	2 GOTV Mailings
PA	1,937	111,569	Ages 25-65, Voting Propensity Less than 100%	26	2 GOTV, Hate Crime
VA	658	37,883	No Experimental Mailings	18	

* Note that the pre-experimental mailing was sent with equal probability to treatment and control groups.
**The experimental Hate Crime mailer was sent only to those who did not receive it in the pre-experimental period.

Table 2: Treatment Rates and Intent-to-Treat Effects, by State

State	Control N	Treatment N	% Receiving Mail	% Phoned in Control Group	% Phoned in Treatment Group	Phone Contact Rate	Turnout in Control Group (V_c)	Turnout in Treatment Group (V_t)	V_t-V_c Intent-to-Treat Effect (SE)
FL	2,038	119,070	56	0	12	12	43.2	43.0	-0.2 (1.1)
GA	4,124	236,316	39	1	22	21	43.1	43.4	0.3 (0.8)
MI	2,516	144,948	50	0	26	26	36.2	36.9	0.7 (1.0)
MO	389	23,594	45	39	51	12	51.4	50.3	-1.1 (2.6)
NC	244	13,666	0	0	30	30	36.5	43.1	6.6 (3.2)
NJ	483	29,186	32	45	53	8	66.7	65.0	-1.7 (2.2)
NY	3,082	175,887	25	47	56	9	73.8	74.3	0.5 (0.8)
OH	1,241	71,377	24	47	51	4	54.1	53.3	-0.8 (1.4)
PA	1,937	111,569	46	11	31	20	53.0	53.5	0.5 (1.1)
VA	658	37,883	0	52	67	15	58.8	60.4	1.6 (1.9)

Notes: Less than 1% of the control group received the experimental pieces of direct mail. The mailings consisted of two GOTV mailers and, for residents of FL, MI, MO, and PA, a third mailer concerning hate crime.

Table 3: Experimental Results for those Ineligible to Receive Experimental Mailings

State	Control N	Treatment N	% Phoned in Control Group	% Phoned in Treatment Group	Phone Contact Rate	Turnout in Control Group (V_c)	Turnout in Treatment Group (V_t)	V_t-V_c Intent-to-Treat Effect (SE)
FL	784	44,046	0	22	22	52.4	52.5	0.1 (1.8)
GA	1,569	90,015	2	37	35	55.6	55.5	-0.1 (1.3)
MI	940	56,313	0	34	34	51.4	51.3	-0.1 (1.6)
MO	183	11,441	47	59	12	56.8	55.1	-1.7 (3.7)
NC	244	13,666	0	30	30	36.5	43.1	6.6 (3.2)
NJ	286	15,890	42	55	13	72.4	67.7	-4.7 (2.8)
NY	2,075	118,367	45	59	14	76.7	77.8	1.1 (0.9)
OH	830	46,376	53	60	7	63.7	64.7	1.0 (1.7)
PA	924	55,485	10	38	28	64.0	64.7	0.7 (1.6)
VA	658	37,883	52	67	15	58.8	60.4	1.6 (1.9)

Table 4: Experimental Results for those Eligible to Receive Experimental Mailings

State	Control N	Treatment N	% Receiving Mail	% Phoned in Control Group	% Phoned in Treatment Group	Phone Contact Rate	Turnout in Control Group (V_c)	Turnout in Treatment Group (V_t)	V_t-V_c Intent-to-Treat Effect (SE)
FL	1,254	75,024	89	0	6	6	37.4	37.5	0.1 (1.4)
GA	2,555	146,301	63	1	13	12	35.4	35.9	0.4 (1.0)
MI	1,576	88,635	81	0	20	20	27.2	27.8	0.6 (1.1)
MO	206	12,153	88	32	43	11	46.6	45.8	-0.8 (3.5)
NJ	197	13,296	70	50	50	0	58.4	61.8	3.4 (3.5)
NY	1,007	57,520	75	51	51	0	67.7	67.1	-0.6 (1.5)
OH	411	25,001	70	33	35	2	34.5	32.3	-2.2 (2.3)
PA	1,013	56,084	91	12	23	11	43.0	42.4	-0.6 (1.6)

Table 5: Experimental Results for those Ineligible to Receive Experimental Mailings whose Phone Numbers were Known as of September 2001

State	Control N	Treatment N	% Phoned in Control Group	% Phoned in Treatment Group	Phone Contact Rate	Turnout in Control Group (V_c)	Turnout in Treatment Group (V_t)	V_t-V_c Intent-to-Treat Effect (SE)
FL	272	15,586	0	43	43	76.5	76.5	0.0 (2.6)
GA	766	42,904	2	54	52	67.9	69.6	1.7 (1.7)
MI	467	27,567	0	52	52	67.7	66.5	-1.2 (2.2)
MO	88	5,575	58	76	18	75.0	79.0	4.0 (4.4)
NC	101	6,866	0	48	48	48.5	56.1	7.6 (5.0)
NJ	165	8,422	49	71	22	89.7	83.2	-6.5 (2.9)
NY	1,238	70,312	57	74	17	82.6	82.9	0.3 (1.1)
OH	412	22,868	72	81	9	81.8	82.8	1.0 (1.9)
PA	453	27,877	14	54	40	84.5	82.7	-1.8 (1.8)
VA	335	19,793	68	84	16	85.4	84.6	-0.8 (2.0)

Table 6: Experimental Results for those Eligible to Receive Experimental Mailings whose Phone Numbers were Known as of September 2001

State	Control N	Treatment N	% Receiving Mail	% Phoned in Control Group	% Phoned in Treatment Group	Phone Contact Rate	Turnout in Control Group (V_c)	Turnout in Treatment Group (V_t)	V_t-V_c Intent-to-Treat Effect (SE)
FL	295	18,081	86	0	14	14	61.0	58.6	-2.4 (2.9)
GA	890	52,716	67	2	23	21	52.2	51.7	-0.5 (1.7)
MI	593	32,461	84	0	35	35	40.8	41.6	0.8 (2.0)
MO	92	4,798	92	53	64	11	64.1	74.6	10.5 (4.6)
NJ	87	6,058	73	63	65	2	78.2	78.9	0.7 (4.4)
NY	542	30,988	76	66	66	0	74.0	73.0	-1.0 (1.9)
OH	114	6,538	73	59	57	0	57.9	57.9	0.0 (4.7)
PA	380	21,066	91	16	34	18	61.1	61.7	0.6 (2.5)

Table 7: Two-Stage Least Squares Regression Estimates
(standard errors in parentheses)

	Entire Sample	Entire Sample (with covariates)**	Ineligible for Mail	Received No Pre-Experimental Mail	Matched Phone Numbers
Effects of Receiving Phone Calls	.023 (.023)	.027 (.022)	.023 (.023)	.021 (.023)	.004 (.022)
Effects of Receiving Mailings	-.002 (.008)	-.003 (.007)		.010 (.012)	-.001 (.011)
N	980,208	963,010	497,975	662,544	427,766

* Each model includes control variables that are not reported here: dummy variables for each state and interactions between each state and whether a person was eligible to receive mail. The first-stage equations include these covariates as well as a dummy variable indicating random assignment to the treatment group and an interaction between this dummy and eligibility to receive mail.

** First- and second-stage equations include age, a dummy variable indicating whether the participant had voted in all previous elections, and a dummy variable for whether the participant's phone number could be matched in 2001. Some observations are lost due to missing data for age.