doi:10.1017/S000305541000047X

How Large and Long-lasting Are the Persuasive Effects of Televised Campaign Ads? Results from a Randomized Field Experiment

ALAN S. GERBER *Yale University*
JAMES G. GIMPEL *University of Maryland*
DONALD P. GREEN *Yale University*
DARON R. SHAW *University of Texas at Austin*

We report the results of the first large-scale experiment involving paid political advertising. During the opening months of a 2006 gubernatorial campaign, approximately $2 million of television and radio advertising on behalf of the incumbent candidate was deployed experimentally. In each experimental media market, the launch date and volume of television advertising were randomly assigned. In order to gauge movement in public opinion, a tracking poll conducted brief telephone interviews with approximately 1,000 registered voters each day and a brief follow-up one month after the conclusion of the television campaign. Results indicate that televised ads have strong but short-lived effects on voting preferences. The ephemeral nature of these effects is more consistent with psychological models of priming than with models of on-line processing.

Paid television advertising commands the largest portion of the communications budget in campaigns for the most important elective offices and represents an important source of voter information about candidates. Despite increased use of Internet communications and renewed attention to voter mobilization fieldwork, big campaigns are still essentially paid media battles that aim to persuade voters. Our study addresses two unresolved questions regarding the persuasive influence of mass media campaigns: What is the effect of television and radio campaign advertising on voter preferences? How long do the effects last? After addressing these issues we use our empirical results to consider an important further question: What do our results suggest about how voters process political information?

We analyze the findings of a randomized field experiment measuring the size and duration of campaign effects caused by a $2 million television and radio buy. There are two main results. First, across a range of model specifications, television campaign advertisements have a large and statistically significant effect on voter preferences. Second, and perhaps most surprising, the effects of the advertisements dissipate rapidly. Nearly all previous research on advertising effects has ignored the issue of decay and implicitly assumes that decay, if it occurs, takes place over weeks or months. We find that just a week or two later, the advertisement's effects have all but disappeared.

The arresting finding of sizable effects and rapid decay has important implications for our understanding of campaign strategy and the effect of campaign spending on election outcomes. The results also have implications for alternative models of voter learning. As we explain at greater length in the discussion section of our paper, a large initial response followed by a quick return to pretreatment opinions does not fit well with models of on-line processing. In these models, existing opinions are adjusted when new information is received and then the resulting new opinions are maintained, even if the information that caused the change is forgotten. We find that campaign advertising causes a large initial boost in support for the sponsor, but this effect diminishes rapidly over time. This up and then down pattern is far more consistent with opinion change due to priming effects, where the advertising temporarily alters the mix of considerations that are salient for the respondent at the moment the respondent is asked to state an opinion.

Despite the substantial attention that the media's influence has received, no real-world experiments have measured the size and duration of effects of a candidate's media campaign. This is an important gap, as the use of random assignment in a field context addresses a number of potential methodological vulnerabilities in earlier work. Random assignment enables laboratory studies to detect causal effects, but the external validity of these measurements remains uncertain. Observational studies, despite recent advances in measurement and design, cannot identify the causal effects of advertising without invoking strong assumptions about unobserved factors that might be correlated with advertising and vote choice. To date, field experiments

Alan S. Gerber is Professor, Department of Political Science and Institution for Social and Policy Studies, Yale University, 77 Prospect Street, P.O. Box 208209, New Haven, CT 06520 (alan.gerber@yale.edu).

James G. Gimpel is Professor, Department of Government and Politics, University of Maryland, College Park, MD 20742 (jgimpel@gvpt.umd.edu).

Donald P. Green is Professor, Department of Political Science and Director, Institution for Social and Policy Studies, Yale University, 77 Prospect Street, P.O. Box 208209, New Haven, CT 06520 (donald.green@yale.edu).

Daron R. Shaw is Professor, Department of Government, University of Texas at Austin, University Station A1800, Austin, TX 78712 (dshaw@austin.utexas.edu).

The authors are grateful to David Carney and the Texans for Rick Perry Campaign for their willingness to conduct a randomized evaluation and to share their data with us. Special thanks go to Peter Aronow, who assisted with data analysis and manuscript preparation. The authors bear sole responsibility for any errors.

involving mass media have focused exclusively on nonpartisan messages, and no field experiment has come close to deploying television communications on the scale and intensity of an actual statewide campaign.[1]

Our research attempts to fill this gap by evaluating a $2 million television and radio campaign. This three-week campaign encompassed 18 television media markets (designated market areas, hereafter "DMAs") and more than 80 AM and FM radio stations. At its peak, the experimental campaign deployed up to 1,000 gross ratings points (GRPs[2]) of advertising per week into selected DMAs, which is comparable to what battleground states experience during the fall of a presidential election year (Johnston, Hagen, and Jamieson 2004). We track outcomes using a large daily tracking poll, which gathered approximately 1,000 interviews per day (for analyses of media effects using large tracking polls, see Hill et al. 2010; Huber and Arceneaux 2007; Johnston, Hagen, and Jamieson 2004).

The remainder of this essay is structured as follows. We describe the prior literature on the effects of media campaigns. No previous field experiments have measured the effects of a partisan media campaign, and very little work of any sort has estimated the decay rate of campaign advertising. We then describe the gubernatorial race that provides the setting for our experiment and the television and radio ads that were the focus of our experimental evaluation. Next we describe the design of the experiment and the survey that we used to assess outcomes. We then present an array of cross-sectional time-series models that parameterize media effects in different ways. We conclude by discussing the implications of our findings for competing hypotheses about how voters process campaign information.

PREVIOUS LITERATURE

Our research is the first field experiment assessing a large-scale partisan campaign and introduces a new approach to the decades-old scholarly effort to understand the effects of media campaigns on attitudes and behavior. The persuasiveness of mass media communications has been hotly debated since the dawn of modern social science during the 1930s. Although most political observers accorded enormous weight to radio, newsreels, and, later, television communication, early students of political propaganda tended to downplay these effects. The survey findings of Lazarsfeld, Berelson, and Gaudet (1944), coupled with the controlled experiments of Hovland, Lumsdaine, and Sheffield (1949), called into question the persuasiveness of political communication and ushered in the "minimal effects" thesis (Klapper 1960) that would hold sway among academics until the 1980s. In contrast, the overwhelming majority of more recent studies reach a different judgment, finding that media campaigns can have substantial effects. Using a variety of research designs, scholars have concluded that mass media campaigns can have important effects on voter attitudes and behavior.

One common research strategy employs survey data and correlates self-reported media exposure, measured by reported media usage or reported campaign advertising exposure, with voter opinions (e.g., Baum 2002). These studies tend to find very strong media effects, but the results are vulnerable to selection bias (respondent viewing patterns may be correlated with existing political views) and measurement bias (measurement error in self-reports of exposure may be correlated with political views).[3] Many recent studies improve on self-reported exposure measures by obtaining measurements of the advertising actually aired in the respondent's media market (Freedman, Franz, and Goldstein 2004; Freedman and Goldstein 1999; Huber and Arceneaux 2007; Johnston, Hagen, and Jamieson 2004). However, because advertising levels are chosen by the campaigns, the level of advertising exposure that a survey respondent's area receives may be correlated with other differences across places, including the level of targeted campaign activity (such as mailings, visits, phone calls), that are hard to measure accurately.

Several recent studies of media effects give careful attention to causal identification and aim to uncover and analyze "natural experiments" (Ashworth and Clinton 2006; DellaVigna and Kaplan 2007; Huber and Arcenaux 2007; Krasno and Green 2008). The study most similar in spirit to our work is DellaVigna and Kaplan (2007), which uses the staggered national rollout of the Fox News channel to estimate the effect of Fox News on voting in the 2000 presidential election. Comparing the change in Republican vote share between 1996 and 2000 in places that began receiving Fox News between 1996 and 2000 with those places that did not yet receive Fox News, the authors find that receiving Fox News boosted the Republican vote share by approximately 0.5 percentage points. The Della Vigna and Kaplan estimation strategy rests on the key assumption that, after conditioning on the control variables, regions that receive Fox News during the 1996 to 2000 period are politically similar to regions where Fox News was introduced after the 2000 election. Like DellaVigna and Kaplan, we employ a staggered introduction of the media "treatment," but we use random assignment to determine the timing of the ads. By design, the timing of our treatment is statistically independent of unmeasured differences in political opinions.

[1] Previous field experiments relating television to political attitudes have involved very small numbers of media markets. For example, the study reported in Ball-Rokeach, Rokeach, and Grube (1984) involved just two media markets. Green and Vavreck (2008) randomly assign nonpartisan television commercials to cable TV markets and study their effects on voter turnout. Panagopoulos and Green (2008) study the effects of nonpartisan radio messages on voter support for relatively unknown challengers in mayoral elections. Clinton and Lapinski (2004) use a Web-TV panel survey to study the effects of exposure to a presidential ad.

[2] One GRP is equal to 1% of the viewing audience; 1,000 points is ostensibly the equivalent of everyone seeing an ad 10 times, though we note below that this estimate may cover a variety of scenarios. We follow the convention of using GRPs as the measure of advertising volume (see Ansolabehere, Iyengar, and Simon 1999, 903; Johnston, Hagen, and Jamieson 2004, 70; Shaw 1999, 349).

[3] See Vavreck (2007) for evidence that measurement error in ad recall can produce spurious media effects.

Laboratory studies of campaign effects overcome many of the diverse threats to internal validity present in observational studies. Experiments have suggested that television news shapes what issues the public considers important (Iyengar, Peters, and Kinder 1982), that the tone of televised advertising influences voter cynicism (Ansolabehere and Iyengar 1995), and that the content of news stories shapes public opinion toward race-related policies (Gilliam and Iyengar 2000). Campaign advertisements are found to significantly increase voters' support for the sponsoring candidates (Ansolabehere and Iyengar 1995, 186; Brader 2005, 395; Kaid 1997, 1088; Valentino, Hutchings, and Williams 2004, 349). These results are highly suggestive. However, although laboratory studies often go to great lengths to simulate a naturalistic viewing environment, it is unclear whether the laboratory findings parallel the direction and magnitude of voter responses and campaign effects in real world settings. Rarely do laboratory studies gauge the duration of media effects (Gaines, Kuklinski, and Quirk 2007). It is unclear whether the effects produced through laboratory interventions are as long-lasting as those produced by campaigns in typical real world settings.

Surprisingly little research of any kind addresses the decay of political advertising effects. Although some work has shown decay in recall (Lodge, Steenbergen, and Brau 1995), decay is not estimated in leading studies of the impact of advertising on voter attitudes and behavior (e.g., Hillygus and Shields 2008; Huber and Arceneaux 2007; Johnston et al. 1992; Johnston, Hagen, and Jamieson 2004; Shaw 1999), which implicitly assume that advertising effects endure for several weeks or months. Empirical studies of campaign spending effects, starting with Jacobson (1978), correlate election results with total incumbent and challenger spending over the course of the campaign. Studies linking recalled exposure to survey responses do not make fine distinctions regarding how recent the exposure is. The prevailing inattention to the duration of effects is reflected in the standard ANES questions regarding campaign exposure. Respondents are asked questions such as "Do you recall seeing any ads for political candidates on television this fall?" This item does not account for the possibility that exposure last week is materially different from exposure the week before last, let alone exposure a month or two earlier.

Decay is generally neglected in political science research, but there is a small literature on decay of product advertising effects, which concludes that there is rapid decay in recall of the advertisement (Burke and Srull 1988; Vakratsas and Ambler 1999; Zielske and Henry 1980), attitudes toward the product (Haugtvedt et al. 1994), and purchase intention (Havlena and Graham 2004). The speed of decay can sometimes be determined from the data presented in these reports. For example, Zielske and Henry (1980) use survey data on recall of the advertisement and data on television advertising buys to estimate that recall falls by about 10% each week. Qualitatively similar findings are reported in the laboratory experiments of Haugtvedt et al. (1994), although decay rates vary across laboratory experimental conditions and are only measured a single time (1 week) after initial advertising exposure.

The only recent work that focuses on the decay rate of political advertising is a pair of observational studies by Hill et al. (2010). Hill et al. use survey data and data on advertising buys to estimate a dynamic model of decay rates across a variety of campaigns.[4] This study, which hinges on the assumption that the timing and extent of advertising volume are exogenous, indicates that the impact of advertising in House, Governor, and Senate elections in 2006 has a half-life of between two and three days, whereas presidential advertising in the 2 months prior to the 2000 election has a half-life of about 1 week. They also find some evidence that decay rates vary with political awareness; advertising effects decay more rapidly for the least politically aware. Our experimental results, which were obtained independently and based on an experiment performed prior to the November 2006 election, complement the Hill et al. research by providing a firm methodological basis for the finding that some campaign advertising may have a much shorter half-life than previously expected.

DESCRIPTION OF THE EXPERIMENTAL SETTING

Our analysis focuses on the 2006 reelection campaign of Texas governor Rick Perry. Aside from its receptiveness to experimental evaluation, the Perry campaign started off much like other big-state reelection campaigns. Perry became the Republican governor of Texas in 2000, succeeding George W. Bush when the latter resigned after winning the presidency. Perry had from 1985 to 1991 served two terms as the Commissioner of Agriculture and in the Texas House of Representatives. With deep Texas roots in a rural part of West Texas, Perry had solid conservative credentials that put him squarely in the middle of the Texas mainstream. In 2002, Perry was reelected governor, winning decisively over a well-financed Democratic opponent by a 58 to 40% margin.

The 2006 race presented a more difficult reelection test. Early signs indicated that the public had a far dimmer view of President Bush's performance heading into the 2006 race than they had heading into the 2002 contest. Perry also faced some internal strife within his own legislative majority, after action on an important education finance package stalled during the 2005 legislative session. During the first half of 2005, Perry's popularity ratings declined, and it appeared that he might face a stiff challenge in the GOP primary from U.S. Senator Kay Bailey Hutchinson. But in July 2005, Hutchinson decided not to challenge Perry, and in September Hurricane Katrina put the Governor's leadership on display through a surge in media coverage. His approval rating rebounded to its highest level in three years, just above 50%. Even so, a threatened primary challenge from state comptroller

[4] An earlier study by Shaw (1999) estimates the decay of some campaign activities but does not analyze decay of media advertising.

Carole Keeton Strayhorn kept the Perry camp on edge though the end of 2005. Although a less formidable opponent than Hutchinson, Strayhorn was a credible challenger, with her own distinguished record of public service. Strayhorn was the first female mayor of Austin and the only mayor elected for three successive terms. She was also the first woman elected Texas Railroad Commissioner—a position that has often been considered a stepping stone to the governorship.

The late filing deadline meant that final announcement of a Strayhorn primary challenge could wait until the beginning of January, which, with early voting in Texas, meant a potentially short and intense battle for the nomination in early March. In the fall of 2005, Strayhorn aired a series of radio ads that were critical of the governor. In December, she purchased several hundred thousand dollars worth of television time to air in January, presumably to kick off a nomination challenge; the Perry campaign countered by buying three weeks of advertising that form the basis of this experiment. However, on the day of the deadline, January 2, she announced her intention to run as an Independent, which meant that January 2 marked an end to the primary contest and the beginning of the general election campaign.

As noted below, the media campaign launched in January by Texans for Rick Perry was conceived in December, before it was known that Strayhorn would not be running in the primary. The Strayhorn campaign deployed a television and radio ad during the first three weeks of January in selected media markets, but did not advertise during the rest of the primary season. The other candidates, for their part, deployed no media advertising in January. During the early months of 2006, none of the candidates fielded more than a small ground campaign. Campaign communications were almost exclusively deployed through television and radio.

To be sure, the circumstances of this experiment are unusual: the start of a yearlong campaign in which a GOP incumbent governor squared off against two independent candidates and an as yet unnominated Democrat. At the same time, the manner in which ads were deployed closely approximates what Zaller (1996) describes as the ideal conditions for detecting media effects: well-measured, abrupt shifts in the quantity of advertising; a vacuum or profusion of opposing ads; a single ad that is deployed through the three-week experimental period, with no ads preceding or following it; and continuous tracking of opinion before, during, and after the flight of advertising. Although this experiment cannot tell us how media effects might play out under different conditions, it does speak with special clarity to the question of whether paid advertising is capable of producing noticeable shifts in voter support.

THE EXPERIMENTAL CAMPAIGN ADS

Texans for Rick Perry kicked off its campaign with an advertising message that highlighted the governor's accomplishments and charisma. In an attempt to appeal to a broad spectrum of Texas voters, the television ad sought to link positive images of the photogenic governor with voters' pride in the state of Texas. The scenes sweep from Texas landscapes to a schoolroom to a doctor's office, with Governor Perry's voiceover:

> I've never been more proud to call myself a Texan. In Texas we've set the national standard for economic development. We gained 300,000 new jobs. Lawsuit reform is bringing better healthcare to millions. We've invested ten billion new dollars in our public schools while improving standards of accountability for student performance. Our people are compassionate. Our vision, bold. Our values, strong. The best is yet to come. I'm proud of Texas. How 'bout you?

Two aspects of this ad are noteworthy. First, it appeals only indirectly to voter ideology, focusing instead on their retrospective performance evaluations and what Stokes (1963) dubbed "valence" issues. Oblique ideological reference is made to "lawsuit reform," "improved standards of accountability for student performance," and strong values, but these conservative themes are balanced by willingness to spend more on education. The emphasis throughout is on accomplishments and the pride that Texans should take in their state and the governor's stewardship. The visuals shift from one iconic Texas image to another, interspersed with images of a handsome suit-clad governor milling with workers and children or poised in front of the State Capitol.

Second, the ad makes no mention of opponents or their platforms or attributes. The script does not even conjure up implicit critiques of the challengers by referring to the governor's "proven leadership" or other personal traits. This tactic reflects the strategic setting at the time during which the ads aired. Carole Keeton Strayhorn had limited name recognition (at least with that particular name—she was known to some Texans by a last name that she dropped after her remarriage), as did the two Democratic contenders for their party's nomination. By "going positive," the Perry campaign sought to deny its opponents the salience that comes with back-and-forth in both the paid media and accompanying news coverage. Although no single ad can ever be said to be representative of the population of ads aired by well-funded media campaigns, it should be noted that the Wisconsin Advertising Project's content analysis of TV ads by U.S. Senate candidates in 2004 found that positive promotion ads, such as Perry's, comprised 46% of all ads aired in competitive races and 69% of all ads aired in uncompetitive races (Franz et al. 2008).

The radio ad followed a similar format but with more specific references to accomplishments. The text of the ad can be found in supplementary online Appendix A (available at http://www.journals.cambridge.org/psr2011005). The overall theme is that voters have ample reason to be proud of Texas given the many ways in which the state has excelled under Governor Perry's leadership. Although no mention is made of a political campaign—no requests for voters' support, for

example—the television and radio ads implicitly anticipate criticisms by other candidates.[5]

EXPERIMENTAL DESIGN

This section explains how the random assignment was performed and the implications for the statistical analysis that follows. For the broadcast television and radio experiments, the cross-sectional unit of observation is the DMA, or designated media market. Texas comprises 20 such markets of varying geographic and population size. These broadcast media markets also define the boundaries of cable TV systems, which are obligated to carry advertising purchased on broadcast stations. Approximately two-thirds of all households at this time subscribed to cable television.[6]

Of these 20 media markets, the campaign was willing to allow experiments in 18, regarding the other two (Houston and Dallas–Fort Worth) as too politically important to leave to chance. In light of the heterogeneity of the DMAs, we matched them as closely as possible based on demographic and socioeconomic attributes and then randomly assigned members of each stratum into an ordering that indicated the start date of the broadcast television campaign. See online Appendix C for a list of these matches. Within each weekly rollout bracket, we randomly assigned the quantity of weekly GRPs to be purchased: 250, 500, or 1,000. The rollout dates were then given to the campaign's television media buyer, who arranged to purchase the quantity of broadcast TV ads that we specified for each DMA each week.[7] Given the small number of DMAs, the power of the experiment derives from the over-time changes in advertising within DMAs, and the analysis presented below focuses on the within-subjects design.

The experiment did not randomize the stations and programs on which the ads were placed. Broadcast TV ads were purchased in a variety of stations based on the campaign consultants' strategic judgment. Table 1 illustrates the sort of advertising purchases that were made in three DMAs that varied in terms of the quantity of advertising to which they were randomly assigned. A DMA assigned 250 GRPs aired most of the Perry ads on morning shows and news programs. A DMA assigned 500 GRPs bolstered the morning and news programming with additional entertainment programming. At 1,000 GRPs, the quantity of entertainment programs increased further, and the range of shows extended into late night entertainment and news programs. Future work may attempt to estimate the distinct effects of a fixed supply of GRPs spread over a varying number of shows or the effects of greater GRP weight within a given set of shows.

Once the broadcast TV rollout brackets were determined, members of each bracket were randomly assigned to a period of radio advertisement. The GRP weight of the radio ads was not varied randomly. The purchaser of the radio ads was given discretion about how to place the ads, subject to the constraint that each radio station's signal remained largely confined to a single DMA. The research team obtained data on each station's signal propagation zone and worked with the radio purchaser to generate a list of advertising purchases. In the end, the radio purchases were somewhat more circumscribed geographically than would be usual for a political campaign, but still achieved a substantial audience.

The timing and intensity of the television and radio ads in each DMA are depicted in Table 2. For example, the Victoria DMA received a full dose of 1,000 television GRPs from the first week on, whereas the Lubbock and Austin DMAs did so from week two on, and the Corpus Christi DMA received a burst of 1,000 GRPs only during the third week. Looking at the TV and radio trajectories in conjunction, we see a range of different configurations. In 13 DMAs there were periods during which TV ran in the absence of radio. In 4 DMAs, radio ran in the absence of broadcast TV. In 10 DMAs, the two ran concurrently during at least one week. Also noteworthy are weeks in which nothing aired: 11 DMAs aired no media during week 1; 4 were quiescent during week 2; and 2 DMAs ran no ads during week 3. After week 3, no ads were run in any media markets for the duration of the study.

As we noted earlier, the Strayhorn campaign was on television and radio in some but not all of the media markets into which the Perry ads were introduced. The right panel of Table 2 shows the location and temporal relationship between the Perry and Strayhorn ads. As might be expected, given random assignment, the graph makes clear that the Strayhorn ads bear little relationship to the placement and timing of the Perry ads in the 18 experimental markets. Sometimes her ads precede the Perry ads; in other cases, they run concurrently. Her ads were localized both geographically and temporally, leaving ample statistical leverage for differentiating the effects of the two campaigns competing communications.

In forming the data for our pooled cross-sectional time-series analysis, the state of Texas has been divided into media markets. We divide the time units into weeks, as the campaign purchased its GRPs on

[5] The Perry campaign did not have a copy of Carole Keeton Strayhorn's ad but expected performance-related criticisms along the lines of Strayhorn's radio ads during the fall. Like the Perry ad, the Strayhorn televised ad features the candidate as narrator and focuses primarily on performance evaluations. Unlike the Perry ad, hers mentions party directly but parallels his strategy by attempting to stand above politics. It does not "go negative" in the conventional sense of attacking the personal failings of the opponent. Instead, it critiques the partisan deadlock of which the governor is a part. The Perry ads in some sense answer this critique by adducing evidence that conditions in Texas are strong and improving.

[6] The experiment also involved the random assignment of cable TV markets in select DMAs. This experiment, however, was underpowered and produced results with large standard errors. Including or excluding this aspect of the design in the overall analysis of TV advertising has no effect on the conclusions, which are driven by the broadcast TV results.

[7] In one DMA, the buyer was unable to place television ads totaling 1,000 GRPs and purchased 475 instead. Unless otherwise noted, we estimate intent-to-treat effects; that is, the number of GRPs used in the estimation process reflects treatment assigned, not treatment received. Using instrumental variables estimation to correct for the discrepancy between treatment assigned and treatment received does not substantively alter the causal estimates. See Table 4.

TABLE 1. Illustration of Programming on Which Broadcast TV Ads Aired in Three DMAs, by Gross Ratings Points

250 GRPs	500 GRPs	1,000 GRPs
5a Morning News	5a Morning News	5a News
6a News	6a News	6a News
7a Today Show	7a Today Show	7a Today Show
7a Good Morn. Amer.	7a Good Morn. Amer.	7a Good Morn. Amer.
7a Early Show	7a Early Show	7a Early Show
9a Regis	9a Regis	9a Regis
Noon News	10a Price is Right	9a Ellen
5p News	10a The View	10a Price is Right
6p News	11a Young & Restless	11a Young & Restless
630p Wheel of Fort.	Noon News	Noon News
630p Millionaire	3p Dr. Phil	1230p Soaps
10p News	4p Oprah	3p Dr. Phil
Sunday Today	430p Jeopardy	4p Oprah
Sunday AM News	5p News	430p Jeopardy
Sunday 5p News	6p News	5p News
Dateline 6-7p	630p Wheel of Fort.	6p News
60 Minutes	630p Enter. Tonight	630p Wheel of Fort.
Sunday 10p News	630p Millionaire	630p Millionaire
Friday 20/20 9-10p	9p News	9p News
	10p News	10p News
	1030p Tonight Show	1030p Tonight Show
	Sunday Today	1030p Letterman
	Sun Meet the Press	1030p Nightline
	Sunday AM News	Sat Jeopardy
	Sunday 5p News	Sat 6p News
	60 Minutes	Sat 10p News
	Sunday 10p News	Sunday Today
	Fri 20/20 9-10p	Sunday AM News
	Fri Law & Order 9-10p	Sunday Today
	Sat Cops	Meet the Press
	Sat Am Most Wanted	Face the Nation
	Tues Navy NCIS	This Week
		NFL Playoff Game
		Sunday 5p News
		Dateline 6-7p
		60 Minutes6-7p
		Ext. Home Makeover
		Sun Cold Case 8-9p
		Law & Order 9-10p
		Sun Movie 8-10p
		Sunday 10p News
		Sunday Sports Ext.
		Tues Navy NCIS
		Weds Law & Order
		Thurs. CSI
		Thurs Primetime Live

a weekly basis. The overall pattern of results is little affected by whether we operationalize time in terms of weeks, which was the unit of experimental assignment, or days.[8] The advantage of coarser time units is the ability to track opinion change more reliably, due to greater numbers of survey interviews.

SURVEY DESIGN

In light of the fact that the unit of assignment is the DMA, we sought to allocate our survey sample in a way that would make the most of the experiment's power. The basic strategy was to spread the surveys more or less evenly across each of the geographic units.

[8] Earlier versions of this paper report results based on daily data using daily polling results and dividing weekly GRPs by seven (we do not have daily GRPs figures, and only weekly totals were randomly assigned). These results, which are similar in terms of effect size and statistical significance to what we present here, are available from the authors.

TABLE 2. Television and Radio Advertising Purchases, by Media Market, Week, and Campaign (Entries are TV GRPs/Radio GRPs)

	Perry Campaign			Strayhorn Campaign		
DMA	Week 1	Week 2	Week 3	Week 1	Week 2	Week 3
Abilene	0/0	0/0	0/0	0/0	0/0	0/0
Amarillo	0/0	0/184	500/184	0/0	0/0	0/0
Austin	0/0	1,000/0[a]	1,000/0[a]	0/80	0/80	0/23.4
Beaumont–Port Arthur	0/0	0/0	250/0	0/0	0/0	0/0
Corpus Christi	0/0	0/185	1,000/185	0/0	0/0	0/0
Dallas–Ft. Worth	300/135	300/135	167/75	0/31	0/31	0/9.4
El Paso	500/0	500/0	500/187	0/0	0/0	0/0
Houston	300/113	300/113	180/113	515/49	546/49	0/15.5
Laredo	0/0	1,000[b]/0	1,000[b]/0	0/0	0/0	0/0
Lubbock	0/0	1,000/0	1,000/0	339/0	483/0	0/0
McAllen–Brownsville–Harlingen	500/0	500/162	500/0	0/0	0/0	0/0
Odessa–Midland	250/160	250/160	250/160	304/0	293/0	0/0
San Angelo	0/180	0/0	0/0	0/0	0/0	0/0
San Antonio	0/0	500/0	500/184	456/60	526/60	0/17.5
Sherman	0/0	250/0	250/0	0/0	0/0	0/0
Shreveport	0/0	250/0	250/0	300/0	260/0	0/0
Tyler–Lufkin–Nacogdoches	0/0	0/0	250/0	423/0	515/0	0/0
Victoria	1,000/0	1,000/159	1,000/159	0/0	0/0	0/0
Waco–Temple–Bryan	500/0	500/178	500/0	475/0	558/0	0/0
Wichita Falls	0/180	500/0	500/0	187/0	434/0	0/0

[a] In weeks 2 and 3, the Perry campaign aired 10 radio GRPs on Christian-format stations. Including or excluding these GRPs has no material effect on the results presented below.
[b] Although Laredo was assigned to receive 1,000 GRPs, media buyers were only able to purchase 475. This issue is addressed in Table 4, where we use instrumental variables regression to correct for the disjunction between intended and actual GRPs.

By reweighting the data according to the probability of selection from the voter file, we can approximate the results that would have obtained for a simple random sample, and our results conform closely to the results from a concurrent poll based on a simple random sample from the voter file.[9]

The survey itself was conducted by Advantage Inc., a firm that specializes in "voter identification" calls—which is to say, brief and inexpensive surveys. One of the practical innovations of this experiment was the use of this type of survey, which cost one-tenth as much as a conventional survey per completed interview, allowing approximately 1,000 completed interviews per day.[10] (See online Appendix D for more information about response rates by week and media market.) The polling firm and its callers were blind to the purpose of the study and to the deployment of Perry advertising; the scripts did not link the survey in any way to a political party or campaign. The calls were conducted only in English, which limits our ability to generalize beyond the population of English-proficient registered voters. According to the 2006 American Community Survey, 14% of the Texas population reported speaking English less than "very well." According to the 2006 Current Population Survey, 2.3% of U.S. *citizens* residing in Texas reported that "Spanish is the only language spoken by all members of the household who are 15 years of age or older." (The survey did not ask about other languages.) From these figures, we infer that English proficiency excluded roughly 5–10% of the Texas electorate from our survey. Because the ads themselves were aired in English, this sample restriction parallels the likely audience for the ads themselves.

After a brief introduction,[11] the first two questions assessed the favorability ratings of the two leading

[9] A poll by Baselice & Associates during January 16–19 put Perry ahead with 41%, the Democratic candidate with 14%, Strayhorn with 24%, and Friedman with 8%, with the remainder undecided or mentioning other names. Perry has a 17–percentage point lead over Strayhorn in our poll and a 19-point lead in the Baselice poll. The main difference between the two polls is that ours has a much larger proportion of undecided voters, perhaps due to the fact that those with lower levels of interest in politics received less weight in the likely voter weighting scheme used by Baselice & Associates.

[10] A new sample was released every other day, following the first day's survey. Thus, the survey is a tracking poll, with independent samples every two days. The short duration of callbacks is obviously a potential source of bias, but this bias is constant over time. Of the phone numbers attempted, 13% yielded interviews; excluding nonworking phones, unanswered numbers, and answering machines, the response rate was 43%. Note that the experimental treatment did not begin until January 5; the first two days of the survey provided a pretreatment baseline.

[11] This introduction was shortened after the first day of interviews in order to increase the response rate. This change in format prevents direct comparisons between the first day of interviewing (during the period before the ads were aired) and the days that follow. This change in format has no adverse effect on the results below, because our models use polling results from the preintervention period as covariates rather than outcomes.

TABLE 3. Survey Results over Time (Unweighted *N*)

	January 5–11	January 12–18	January 19–25	January 26–29	March 5–6
(a) Ballot Test Results by Week of Study					
Perry	33.8%	32.8%	32.7%	31.7%	33.7%
Strayhorn	13.3%	14.1%	14.9%	15.3%	11.4%
Friedman	4.1%	4.0%	4.5%	4.7%	5.3%
Democrat	11.5%	12.8%	13.0%	12.2%	12.0%
Other	2.0%	3.1%	3.7%	2.3%	3.2%
Don't Know	35.3%	33.2%	31.3%	33.7%	34.4%
Total %	100.0%	100.0%	100.0%	100.0%	100.0%
N of Cases	7,040	7,059	7,108	4,032	2,044
(b) Favorability Ratings of Rick Perry					
Strong Positive	30.3%	33.3%	34.9%	35.1%	38.5%
Weak Positive	14.1%	12.9%	13.0%	11.8%	17.1%
No Opinion	33.2%	31.8%	29.5%	29.5%	24.1%
Weak Negative	7.1%	6.7%	6.4%	5.7%	6.8%
Strong Negative	15.2%	15.3%	16.2%	18.0%	13.5%
Total %	100.0%	100.0%	100.0%	100.0%	100.0%
N of Cases	7,157	7,178	7,221	4,087	2,044
(c) Favorability Ratings of Carole Keeton Strayhorn					
Strong Positive	16.2%	20.1%	23.8%	22.9%	21.1%
Weak Positive	11.7%	12.2%	11.9%	10.7%	16.2%
No Opinion	57.7%	52.1%	47.8%	48.8%	45.7%
Weak Negative	5.9%	5.7%	5.6%	4.7%	7.1%
Strong Negative	8.6%	10.0%	11.0%	13.0%	9.9%
Total %	100.0%	100.0%	100.0%	100.0%	100.0%
N of Cases	7,094	7,114	7,146	4,056	2,044

Note: For question wording, see the supplementary online Appendix (http://www.journals.cambridge.org/psr2011005). Results here are unweighted for sampling probabilities, but weighted results are similar.

candidates, Perry and Strayhorn. (See online Appendix B for text.) The third question asked voters whom they would vote for if the election were held today. The fourth and final question on the survey instrument was rotated among three randomly selected alternatives. The first item in the rotation was designed as a manipulation check to assess whether, as expected, respondents in treatment areas were more likely to recall seeing Perry advertisements than respondents in control areas. One-third of survey respondents received this question. A parallel question asking about exposure to Strayhorn ads was directed to one-third of the sample. The final one-third of the sample were asked about their radio listening habits. Note that these concluding questions were asked after candidate preference, so that they would not contaminate the outcome measures. The survey contains no measures of slow-moving variables such as age, education, or group membership, as the influence of time-invariant variables is eliminated in the within-subjects analysis presented below.

Table 3 describes statewide trends in public opinion. Although these trends do not speak to the question of advertising effectiveness, they provide a useful description of how the campaign unfolded in early 2006. The incumbent governor enjoyed a lead, but approximately one-third of the respondents declined to express a voting preference. As we move from the ballot test to candidate evaluations, somewhat stronger over-time patterns are apparent. During January, evaluations of Rick Perry became more polarized, with larger proportions of the sample offering strongly favorable or strongly unfavorable evaluations. Less pronounced but still noticeable is the week-by-week increase in the proportion of people who provide an evaluation. Evaluations of Carole Keeton Strayhorn were confined to a much smaller proportion of registered voters, as a large proportion of respondents conceded that they were unfamiliar with her. This proportion declined gradually, from 58% in the first week after she declared her candidacy to 49% three weeks later.

GAUGING THE EFFECTS OF BROADCAST TELEVISION AND RADIO

Our analysis is based on aggregate survey data, where the level of geographic aggregation is the media market and the level of temporal aggregation is the week (see online Appendix D). The resulting dataset contains 90 observations ($N = 18$ experimental media markets; $T = 5$ weeks, comprising three treatment weeks, one immediate follow-up week, and a long-term follow-up week five weeks later). For each DMA, we also have a pretreatment reading of voter preference during the days prior to the launch of the media campaign. Because the aggregate units encompass varying numbers of survey observations, the analyses below show how the results change when the data are weighted

analytically to reflect the number of individual observations in each aggregate unit.

Within each randomization stratum,[12] we model the aggregate survey response as a linear function of the campaign's television and radio advertising, along with fixed effects for both time and geography:

$$Y_{it} = \alpha + \beta_1 \text{TV GRPs}_{it} + \beta_2 \text{Radio GRPs}_{it} + \gamma_1 \text{Week}_{it}^1 + \cdots + \gamma_{T-1} \text{Week}_{it}^{T-1} + \delta_1 \text{Market}_{it}^1 + \cdots + \delta_{K-1} \text{Market}_{it}^{K-1} + u_{it}. \quad \textbf{(1)}$$

In Equation (1), "TV GRPs" are weekly gross ratings points (in 1,000s) associated with Perry's TV ads, and "Radio GRPs" are weekly gross ratings points (in 1,000s) associated with Perry's radio ads. The dependent variable (Y_{it}) is the percentage expressing an intention to vote for Perry (without excluding "don't know" responses).[13] The use of fixed effects for geography tracks a given DMA over time as ads are randomly rolled out; in effect, we have 18 distinct time series in which the media buy at any given time is randomly determined. Fixed effects for geography also have the virtue of controlling for regional differences in opinion caused by different survey response rates. The use of weekly fixed effects controls for statewide shocks that result from events or the vagaries of survey administration in any given week. When pooling all of the strata together into a single regression, we expand Equation (1) to include interactions between the weekly dummy variables and dummy variables marking each stratum.

An alternative modeling approach is to treat the δ_k parameters associated with each media market as random draws from a normal distribution. This random effects model uses fewer degrees of freedom and potentially generates more efficient estimates than the fixed effects model. Because DMAs are randomly assigned to treatments, the random effects are statistically independent of media exposure, which is a precondition for unbiased estimation. However, with just 18 experimental units, there is risk of correlation between the media treatments and the random effects, which can produce distorted estimates. In our experiment, the pretreatment measures of Perry support turn out to be negatively correlated with the television GRPs that were assigned to subsequent weeks. We present the random effects estimates, but suspect that they underestimate the effects of advertising.

One aim of this study is to examine the rate at which advertising effects decay. One way to study time-series dynamics is to introduce a lagged dependent variable as a regressor in Equation (1):

$$Y_{it} = \alpha + \rho Y_{i,t-1} + \beta_1 \text{TV GRPs}_{it} + \beta_2 \text{Radio GRPs}_{it} + \gamma_1 \text{Week}_{it}^1 + \cdots + \gamma_{T-1} \text{Week}_{it}^{T-1} + \delta_1 \text{Market}_{it}^1 + \cdots + \delta_{K-1} \text{Market}_{it}^{K-1} + u_{it}. \quad \textbf{(2)}$$

This specification is a nested alternative to Equation (1), in that the two are identical when $\rho = 0$. For $0 > \rho > 1$, media effects decay geometrically in time, with larger values of ρ imply greater persistence in the effects of media advertisements. For example, if ρ were 0.5, TV GRPs would affect vote share by β_1 in the current week, by $(0.5)\beta_1$ in the subsequent week, $(0.5^2)\beta_1$ the week after, and so on.

As an alternative to the geometric decay model, we can also explore the time decay of ads using a variant of Equation (1) in which lagged media buys are included as regressors. A finite distributed lag model with a single lag, for example, augments the model in (1) by adding the preceding week's TV and radio GRPs:

$$Y_{it} = \alpha + \beta_1 \text{TV GRPs}_{it} + \beta_2 \text{Radio GRPs}_{it} + \beta_3 \text{TV GRPs}_{it-1} + \beta_4 \text{Radio GRPs}_{it-1} + \gamma_1 \text{Week}_{it}^1 + \cdots + \gamma_{T-1} \text{Week}_{it}^{T-1} + \delta_1 \text{Market}_{it}^1 + \cdots + \delta_{K-1} \text{Market}_{it}^{K-1} + u_{it}. \quad \textbf{(3)}$$

Again, when pooling across randomization strata, we include interactions between strata dummies and week dummies to account for the fact that observations in different strata have different probabilities of being treated in each week (or in a previous week). Finally, a polynomial distributed lag (PDL) model operates on a principle similar to that in Equation (3) but sidesteps the problems of collinearity that result when multiple lags are introduced as regressors (Almon 1965). The PDL specification constrains the lagged effects of a regressor to follow a polynomial equation.[14] We show below that the results of the PDL specification are robust for varying lag lengths and polynomial orders.

Each of these regression models may be augmented by including time-varying control variables. (Time-invariant control variables, such as pretreatment support for Perry in each market, are effectively included in the fixed effects model already, as they are perfectly predicted by the dummy variables for each media market.) Of particular interest are the radio and television GRPs purchased by the opposing Strayhorn campaign.

[12] A stratum refers to a group of DMAs that all have the same probability of being assigned to the treatment. We have three strata: those with two DMAs, those with three DMAs, and those with four DMAs.

[13] The results are unchanged when the dependent variable is calculated only for those respondents who express a candidate preference or when, instead of using vote preference, we analyze the net ratings of Perry and Strayhorn.

[14] In the PDL model, the K lag coefficients $\beta_0, \beta_1, \beta_2, \ldots, \beta_K$ are assumed to fall on a polynomial of order Q such that $\beta_k = \psi_0 + \psi_1 k + \cdots + \psi_Q k^Q$ for $k = 0,1, \ldots K$. Determining the polynomial order is typically done via a sequential testing procedure starting with $Q = K$ and reducing Q in order to conserve degrees of freedom. See Judge et al. (1988, Ch. 17). Amemiya and Morimune (1974) find that second- and third-order polynomials work well for many time series. Another way to conserve degrees of freedom is to stipulate that this polynomial equals zero at $j = K + 1$; see Amemiya (1985, Ch. 5).

Because the opponent's points were not randomly assigned, their causal effects are not identified. Including measures of Strayhorn's advertising nevertheless allows us to estimate the Perry ads' effects while controlling for whatever incidental correlation might exist between the randomly assigned Perry ads and their nonrandomly assigned counterparts. In operational terms, this simply involves adding control variables to the equations above—something that turns out to have no real consequence for the estimated effects of the Perry ads. A further control variable is the average partisanship of a DMA. Partisanship is imputed based on the voter file, which indicates the number of Republican or Democratic primaries in which a person has voted. This pretreatment measure is a significant predictor of candidate preference at the individual level, and post-stratifying daily or weekly samples by partisanship dampens the sampling variability associated with our DMA-level averages of Y_{it}.

RESULTS

This section begins by estimating simple models that ignore time-series dynamics and gradually builds up to more complex dynamic models. To preview the findings, we find relatively little evidence of time-series dynamics. Simple models, in other words, lead to roughly the same substantive conclusions as more elaborate models.

The columns of Table 4 show how the estimated effects of television and radio advertising change across alternative modeling approaches that do not involve dynamics. Table 4 focuses on the first four weeks of the advertising campaign. The first two columns report least-squares estimates with fixed effects for media market and week. Without covariates, the estimated effects of 1,000 GRPs of television and radio advertising are 5.27 (SE = 1.50) and 4.25 (SE = 5.91), respectively.[15] To put these estimates in perspective, it is helpful to bear in mind the fact that the campaign purchased up to 1,000 weekly GRPs of television advertising and 187 weekly GRPs of radio ads. The maximum dosage of television advertising apparently boosted Perry's relative standing by approximately six percentage points. Advertising thus appears to have the capacity to induce a substantial shift in vote preferences. The variance in radio advertising volume is much smaller, and the standard errors are therefore larger. We detect no significant effects of radio advertising, but this may reflect the much larger range of statistical uncertainty. As can be seen in column (2), the magnitude and statistical significance of the TV and radio coefficients remain largely unchanged when one introduces controls for the airing of opposing ads or the partisan composition of the sample.

The next three columns of Table 4 explore the consequences of altering the model to allow for random rather than fixed effects for each DMA. Without covariates, the estimated effect of television is smaller (2.25) and short of statistical significance. When we control for each DMA's level of Perry support prior to the start of the media campaign, the coefficient rises to 3.83 (SE = 1.26). Controlling further for partisanship and Strayhorn media raises the random effects estimate to 4.20 (SE = 1.32).

Specification (4) reports the results of a regression in which the data are weighted by sample size. The estimated effect of TV ads remains substantively unchanged when weights are applied. The coefficient (5.02, SE = 1.58) is very similar to the unweighted fixed effects estimate reported in column (2).

The final column in Table 4 represents the model that most closely reflects the nuances of the experimental design. The observations are weighted by sample size. Instrumental variables estimation is used to correct for the fact that in one of the markets, 1,000 TV GRPs were assigned but only 475 were actually aired, whereas in another DMA, 10 radio GRPs were aired when none were assigned. In this regression, randomly assigned GRPs serve as an instrument for actual GRPs (see Angrist, Imbens, and Rubin 1996). The results of this more exacting specification are similar to the other fixed-effects estimates in terms of magnitude and statistical significance. Overall, Table 4 shows that across a variety of model specifications and estimation methods, one generally obtains strong and significant effects of TV advertising, whereas the effects of radio advertising are properly signed but statistically indistinguishable from zero.

Table 5 shows how the estimated effects of TV and radio change when dynamic parameters are added to the fixed-effects model. To illuminate the way in which treatment effects decay, Table 5 includes data from week 9 of the study, roughly one month after the conclusion of the ad campaign. The regression models presented in columns (3) and (4) include a lagged dependent variable, as in Equation (2).[16] Comparing these columns to columns (1) and (2), which omit the lagged dependent variable, shows that the geometric lag specification has essentially no effect on the estimates. The lagged dependent variable turns out to be a weak predictor, with estimated ρ values that are statistically indistinguishable from zero. This pattern is not altogether surprising, because survey error in a relatively short weekly time-series attenuates the estimate of ρ. More interesting are the regression models that include lagged values of the media variables. Although including these lags introduces collinearity, thereby reducing our ability to estimate the effect of any particular lag with precision, the overall pattern is nonetheless suggestive. Television's effects appear to peak during the week in which the advertisements air. In column (1), for example, we see that the current week's advertising raises Perry's vote share by 4.73 percentage points per 1,000 GRPs (SE = 1.42); a week

[15] When the model in column (1) is estimated stratum by stratum, we obtain estimates for TV GRPs of 6.21 (SE = 2.86), 4.82 (SE = 2.28), and 3.97 (SE = 2.57) for strata containing 2, 3, and 4 DMAs, respectively.

[16] We exclude radio ads from these models for ease of presentation, but the pattern of results would be substantively unchanged if we were to include them.

TABLE 4. Estimates of TV and Radio Advertising's Effects on Voter Preference, Measured Weekly

Independent Variables	(1) Fixed Effects with no Covariates OLS	(2) Fixed Effects with Covariates OLS	(3a) Random Effects with no Covariates GLS	(3b) Random Effects Controlling for Pretreatment Vote Preference GLS	(3c) Random Effects Controlling for Pretreatment Vote Preference and Covariates GLS	(4) Weighted Fixed Effects with Covariates WLS	(5) Fixed Effects Two-stage Least Squares Weighted 2SLS
TV GRPs (in 1,000s)	5.27**	5.12**	2.25	3.83**	4.20**	5.02**	5.44**
(Standard Error)	(1.50)	(1.52)	(1.34)	(1.26)	(1.32)	(1.58)	(1.77)
Radio GRPs (in 1,000s)	4.25	4.89	4.78	3.37	3.94	4.85	4.83
(Standard Error)	(5.91)	(6.01)	(5.68)	(5.01)	(5.35)	(5.81)	(5.99)
Fixed Effects for Media Markets	Yes	Yes	No	No	No	Yes	Yes
Random Effects for Media Markets	No	No	Yes	Yes	Yes	No	No
Controls for Strayhorn TV and Radio	No	Yes	No	No	Yes	Yes	Yes
Controls for Voter Partisanship	No	Yes	No	No	Yes	Yes	Yes
Controls for Pretreatment Vote Preference	No	No	No	Yes	Yes	No	No
Randomization Strata Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	72	72	72	72	72	72	72

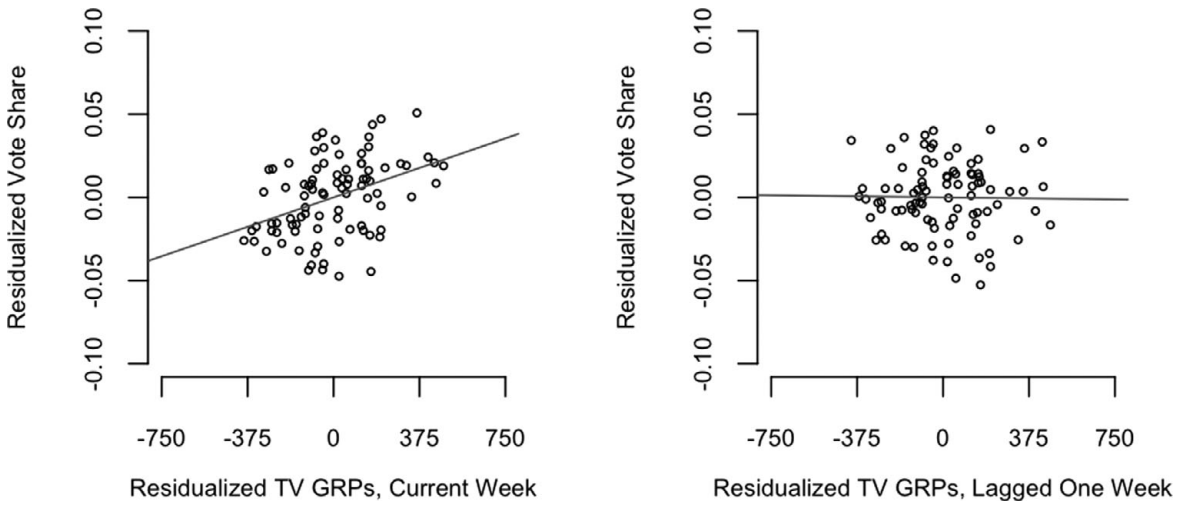
*$p < .05$.
**$p < .01$.
Note: Dependent variable is the percent of all respondents (including "don't know" responses) who, in a given DMA and in given week, favor Perry in a trial ballot. Partisanship is measured as the number of GOP primaries in which the voter has voted, minus the number of Democratic primaries. Specification (5) reports results from a weighted two-stage least squares regression (with a small-sample correction for the SE) in which the assigned TV GRPs and radio GRPs are used as instrumental variables for actual TV GRPs and radio GRPs.

TABLE 5. Dynamic Effects of Television Advertising on Voter Preference, Measured Weekly

	Finite Distributed Lag Models		Geometric Lag Models		Polynomial Distributed Lag Models			
Independent Variables	(1)	(2)	(3)	(4)	2nd Order, 3 Lags (5)	2nd Order, 3 Lags, Assumes No Effect after 3 Lags (6)	3rd Order, 3 Lags (7)	3rd Order, 3 Lags, Assumes No Effect after 3 Lags (8)
TV (No Lag)	4.72	4.73	5.18	5.23	5.78	5.44	6.48	4.07
	(1.41)	(1.42)	(1.50)	(1.54)	(1.72)	(1.56)	(1.85)	(1.28)
TV, 1-week Lag		−0.17		0.42	1.86	2.19	0.44	3.05
		(1.42)		(1.82)	(1.27)	(1.05)	(1.88)	(1.12)
TV, 2-week Lag					0.23	0.20	2.15	−0.01
					(1.25)	(1.24)	(2.25)	(1.37)
TV, 3-week Lag					0.91	−0.53	−0.34	−2.05
					(3.21)	(0.97)	(3.43)	(1.57)
Lag Vote Preference			0.00	−0.01				
			(0.11)	(0.13)				

Note: $N = 90$, except for specifications with lagged vote preference, for which $N = 72$. All models include fixed effects for week, DMA, randomization strata and lagged randomization strata. Finite distributed lag model (1) coincides with the "no covariates" specifications in Table 4, except that radio advertising has been omitted from the model. PDL models allow lagged effects of up to 3 weeks in duration, but models (6) and (8) impose the added constraint that cumulative effects go to zero after 4 weeks.

FIGURE 1. Illustration of the Contrasting Effects of Current and Lagged TV Exposure, Weekly Data*



*Specification is the same as column (2) of Table 5 and includes fixed effects for DMA, week, and randomization strata. $N = 90$.

later, the effects of these ads have receded to −0.17 percentage points (SE = 1.42).[17]

Figure 1 depicts the regression estimates in columns (1) and (2) to illustrate how television advertising boosts the sponsor when it is aired. The left panel plots residualized vote preference by residualized television GRPs aired in the current period. The graph suggests a strong and approximately linear relationship between TV exposure and voter preference. The right panel plots residualized vote preference by residualized television GRPs aired during the preceding week. The flat regression line illustrates the fact that advertising has a weak and statistically insignificant effect on voter opinion a week later.

Turning to PDL models to assess the decay in effects over a longer time frame, we find a consistent pattern across the various specifications: TV ads exert a strong and significant effect in the current week, smaller and statistically equivocal effects a week later, and no effects thereafter. The only models to show significant effects of 1-week lags produce no evidence of lagged

[17] Introducing a spurious 1-week lead into the specification in equation (1), as expected, shows no effect.

TABLE 6. Assessing Whether the Effects of TV Diminish with Prior Exposure

	Interaction with 1-week Lag		Interactions with 1-week Lag and 2-week Lag	
Independent Variables	(1)	(2)	(3)	(4)
TV (No Lag)	4.96**	4.99**	4.96**	5.03**
	(1.62)	(1.68)	(1.64)	(1.67)
TV, 1-week Lag	0.31	0.61	2.13	2.24
	(2.09)	(2.17)	(2.74)	(2.79)
TV, 2-week Lag			−2.37	−2.12
			(2.71)	(2.88)
TV × TV, 1-week Lag	−8.02e-04	−1.06e-03	−3.36e-03	−3.47e-03
	(2.57e-03)	(2.64e-03)	(3.32e-03)	(3.41e-03)
TV × TV, 2-week Lag			7.05e-03	7.19e-03
			(4.41e-03)	(4.58e-03)
Fixed Effects for Week, Media Markets, Randomization Strata, and Lagged Randomization Strata	Yes	Yes	Yes	Yes
Covariates	No	Yes	No	Yes
N	90	90	90	90
F-test *p*-value for all lag and interaction terms	.95	.92	.60	.58

*$p < .05$.
**$p < .01$.

effects beyond 1 week.[18] Overall, the analysis of dynamics suggests that the effects of TV, although powerful, were short-lived. Not only did they dissipate before Election Day in November, they vanished before the March primary.

We can press the data a bit further in order to distinguish between decay and diminishing returns. Decay refers to the reduced influence of exposure over time, whereas diminishing returns refers to the idea that the influence of ads decreases with cumulative exposure to date. A straightforward way to test the diminishing returns hypothesis is to introduce an interaction between this week's ads and last week's ads:

$$\begin{aligned} Y_{it} = {} & \alpha + \beta_1 \text{TV GRPs}_{it} + \beta_2 \text{TV GRPs}_{it-1} \\ & + \beta_3(\text{TV GRPs}_{it} \times \text{TV GRPs}_{it-1}) + \gamma_1 \text{Week}_{it}^1 \\ & + \cdots + \gamma_{T-1} \text{Week}_{it}^{T-1} + \delta_1 \text{Market}_{it}^1 \\ & + \cdots + \delta_{K-1} \text{Market}_{it}^{K-1} + u_{it}. \end{aligned} \quad \textbf{(4)}$$

A negative interaction ($\beta_3 < 0$) is consistent with the diminishing returns hypothesis. The main effect of lagged GRPs (β_2) gives a sense of how lagged exposure influences opinion when the current week's exposure is zero. The main effect of current GRPs (β_1) indicates the effect of "fresh" TV ads in markets where no ads have been shown previously. Table 6 presents the results of this analysis. Columns (1) and (2) interactions with a 1-week lag, and columns (3) and (4) interactions with both 1-week and 2-week lags. No evidence whatsoever is found for diminishing returns, and we again find evidence of strong short-term effects that decay rapidly. The same conclusions emerge when Table 6 is estimated using a random effects model, as in column (3c) of Table 4.

The fact that opinion equilibrates quickly in the wake of a televised ad campaign such as the one studied here underscores the potential costliness of waging a battle for votes on TV. The immediate effect of a televised ad campaign makes it a cost-effective option in the short run. To shift opinion 5 percentage points among an electorate of 14 million people means persuading 700,000 voters. To expose the electorate to 1,000 GRPs of television at a rough average cost of $150 per point amounts to approximately $3 million per week. At less than $5 per vote, this expenditure is a bargain, but the difficulty is that there is little to show for one's money a week or two later.

DISCUSSION

Our study offers the first experimental estimate of campaign advertising's causal effects in a field setting. This research is not the first study to show advertising effects, but demonstrating these effects with a field experiment is an important advance because the

[18] These particular specifications assume that effects die out after 4 weeks, which in effect assumes no effect on opinion by the time of the March primary. In other analyses, we find that Perry support, as measured by the March wave of interviews, is not positively affected by the cumulative TV GRPs shown a month earlier.

research design sidesteps criticisms that are often levied against other research methods. The large effects observed in the laboratory are routinely challenged on the grounds that they fail to tell us how media exposure translates into votes in the context of an actual campaign, and rarely do scholars trace the over-time decay of lab-induced effects. The large effects found in observational studies are similarly open to the charge that campaigns target their ads strategically. Even when observational researchers have perfect measures of media exposure, they rarely have access to the inside information and strategic considerations that determine how a campaign allocates resources over time or space. In the gubernatorial campaign studied here, the volume, timing, and location of TV and radio ads were determined randomly.

The most provocative finding to emerge from this experiment concerns the rapid decay of advertising effects. With the important exception of the recent observational work by Hill et al. (2010), very little attention has been given to the decay of advertising effects, and most previous research has ignored decay or implicitly assumed it is a much slower process than our findings suggest.

The rapid decay of advertising effects has important theoretical implications. In the course of a typical campaign, advertising targets are usually bombarded with messages without interruption; our experimental campaign enabled us to study how opinion equilibrates after the TV messages subside. This unusual design feature sheds light on the psychological mechanisms by which an ad designed to prime positive associations with the governor influences voter preferences.

Two basic psychological perspectives have dominated how scholars approach campaign effects. The first perspective emphasizes the *information content* of advertising. Recognizing that people forget the details of the information they receive, it has been argued that voters process information "on-line," forming new opinions and then discarding the information on which those opinions were based. As Lodge, Steenbergen, and Brau (1995) argue, "memory for campaign messages not only fades but fades quickly . . . the half-life of the message typically being less than a week," yet "recall is not a necessary condition for information to be influential" (pp. 315, 317–18). This interpretation has gradually won acceptance among public opinion scholars, who now concur that messages may have enduring effects even after their content is forgotten.

A second perspective is that advertisements prime voters to *invoke different evaluative criteria* when assessing candidates. An advertisement that stresses the threat posed by international terrorism might increase the extent to which voters draw on their foreign policy attitudes when evaluating presidential candidates. Iyengar, Peters, and Kinder (1982) and Miller and Krosnick (1996), for example, show that news stories have the capacity to strengthen the relationship between targeted attitude domains and candidate evaluations. Another argument with similar implications holds that the emotional content of advertisements accentuates their effects (Brader 2005). The priming hypothesis implies that the effects of advertising decay as the cognitive accessibility of the primed considerations fades.

These two competing perspectives – learning and priming – have quite different empirical and theoretical implications that have not, to our knowledge, been integrated within a common theoretical framework. Although this is beyond the scope of the current paper, we have formalized a model that incorporates both types of effects.[19] Intuitively, from the standpoint of the rational learning model, a transitory advertising effect represents an anomaly. Rapid change is not in itself an anomaly, but to explain both rapid change in the wake of a commercial and rapid reversal of this change, one would have to posit that the information in the commercial contained a good deal of pertinent information to cause the initial change and that, subsequently, the outside environment happened to supply a comparable amount of pertinent information in rebuttal over the course of the next few days. This coincidence seems highly implausible given the uneventful way in which the campaign unfolded during the period of our study. An abrupt shift toward a candidate followed by a return to the prior state is, in contrast, compatible with an altogether different psychological process, akin to what Zaller (1992) refers to as "sampling considerations."

The pattern of results we observe suggests priming rather than on-line processing.[20] The precise way in which priming occurs is subject to alternative but complementary explanations. One possibility is that the ads made people more likely to place certain considerations in working memory. Another is that the ads increased the weight that people accorded these considerations when evaluating the candidates. Whether these ads jogged memory or focused attention or

[19] Formalizations based on Bayes's rule gained prominence in political science with Achen's (1992) model of rational learning, which posits that rational voters have prior beliefs about an unobserved quantity—the net utility gain associated with the election of a given candidate—and update these beliefs based on efficient use of new information. Gerber and Green (1998) generalized this model to account for the possibility that the underlying value of a given candidate changes over time. Their learning model implies that the speed of opinion change is a function of the signal-to-noise ratio, that is, the rate of true change in the underlying attributes of the candidates as opposed to nondiagnostic variation in rhetoric and events. The two corollaries of the rational learning model are that (1) voters change their opinions slowly unless confronted with an especially telling piece of new information suggesting change in the candidates' underlying attributes and (2) as voters learn, they replace old information with new, a proposition that comports with the "on-line processing" hypothesis of Lodge, Steenbergen, and Brau (1995). Combined, these propositions mean that if an ad contains information that causes voters to abruptly prefer candidate A over candidate B, this new state should persist until new evidence is presented to change this view.

[20] One important proviso is that our conclusions regarding priming versus on-line processing are based on analysis of aggregate movements of opinion rather than panel data in which individuals are tracked over time. It would therefore be useful for future research to follow the same set of individuals over time to confirm that the patterns we observe at the aggregate level are repeated at the individual level. It is, however, unclear what conjunction of offsetting individual movements in opinion would produce the aggregate results we observe, especially given the positive tone of the advertising appeals.

both, they do not appear to have propagated enduring beliefs.

The finding that campaign effects diminish sharply over time opens up a new research agenda. The theoretical arm of this broader research project involves formally characterizing the properties of different types of ads and the mechanisms through which they influence voter preference. Rational learning models are relatively well developed formally; more work needs to focus on formalizing propositions about memory, emotion, and attention that might account for the impact and dynamics of advertising.

This theoretical enterprise must be complemented by empirical exploration. Progress requires assessment of a range of different ads' effects in different political contexts. Do the dynamics of opinion change look different when, instead of offering evocative imagery, an advertisement campaign reveals new information? Would, for example, a negative campaign ad that leveled a specific charge about an opponent's malfeasance in office produce an enduring shift in opinion? Do the effects of emotion-evoking ads dissipate more rapidly than ads with comparable content that lack emotional content? Are the effects of advertising muted toward the end of a campaign, when voters have acquired more information about the candidates, or do the priming or emotional mechanisms remain undiminished in influence?

As the empirical consequences of different types of ads come to be understood, scholars will be better positioned to develop arguments about the strategic logic of different types of advertising campaigns. The current study suggests that valence appeals have the capacity to influence large segments of the electorate, but although such ads may succeed in associating a candidate with a popular value in the short run, they face the challenge of presenting information that will have a lasting impact on the way that voters think about the candidates. Hard-hitting attack ads (e.g., charges of corruption) also appeal to valence dimensions, but may adduce evidence that has a more enduring impact on the way that a candidate is evaluated. From a strategic standpoint, the latter strategy seems riskier, which may explain why the incumbent front-runner studied here, who faced an array of challengers from across the ideological spectrum, chose to open his campaign with a positive broad-based appeal (Gerber 2004).

It should come as no surprise that a study such as this one generates as many research questions as it answers. This experiment is among the first in any discipline to estimate the effects of a large-scale media campaign using random assignment. At a minimum, this study must be replicated in other political contexts in order to answer basic questions, such as how much the effectiveness of advertising depends on the tone and content of the ads, proximity to Election Day, and the competitiveness of the race. In a similar vein, the relative effectiveness of radio and television advertising awaits further experimentation, given the limited power of the current study to detect radio's influence.

That said, the current study represents an important advance, both because it shows that large-scale experiments of this sort are possible and because it provides a reliable assessment of mass media's causal effects. Although, like any scientific effort, this study requires replication and refinement, the causal estimates it generates force a rethinking of several key topics in the field of political psychology. If campaign ads have powerful but short-lived effects on voter preference, what are we to conclude about the "minimal effects" thesis? It can hardly be said that people are so inattentive or resistant to new information that they fail to change, but on the other hand – consistent with models of priming or emotion but inconsistent with on-line processing – the effects of these messages appear to fade quickly. The advent of field experimental investigation of campaign advertising's effects, in sum, has the potential to profoundly shape the empirical foundation on which theories of political communication rest.

REFERENCES

Achen, Christopher H. 1992. "Breaking the Iron Triangle: Social Psychology, Demographic Variables and Linear Regression in Voting Research." *Political Behavior* 14: 195–211.

Almon, Shirley. 1965. "The Distributed Lag Between Capital Appropriations and Expenditures." *Econometrica* 33: 178–96.

Amemiya, Takeshi. 1985. *Advanced Econometrics*. Cambridge, MA: Harvard University Press.

Amemiya, Takeshi, and Kimio Morimune. 1974. "Selecting the Optimal Order of Polynomial in the Almon Distributed Lag." *Review of Economics and Statistics* 56 (3): 378–86.

Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 91 (434): 444–55.

Ansolabehere, S., and S. Iyengar. 1995. *Going Negative: How Attack Ads Shrink and Polarize the Electorate*. New York: Free Press.

Ansolabehere, Stephen D., Shanto Iyengar, and Adam Simon. 1999. "Replicating Experiments Using Aggregate and Survey Data: The Case of Negative Advertising and Turnout." *American Political Science Review* 93: 901–9.

Ashworth, Scott, and Joshua D. Clinton. 2006. "Does Advertising Exposure Affect Turnout?" *Quarterly Journal of Political* 2: 27–41.

Ball-Rokeach, Sandra J., Milton Rokeach, and Joel W. Grube. 1984. *The Great American Values Test: Influencing Behavior and Belief Through Television*. New York: Free Press.

Baum, Matthew A. 2002. "Sex, Lies, and War: How Soft News Brings Foreign Policy to the Inattentive Public." *American Political Science Review* 96: 91–109.

Brader, Ted. 2005. "Striking a Responsive Chord: How Political Ads Motivate and Persuade Voters by Appealing to Emotions." *American Journal of Political Science* 49: 388–405.

Burke, Raymond R., and Thomas K. Srull. 1988. "Competitive Interference and Consumer Memory for Advertising." *Journal of Consumer Research* 15: 55–68.

Clinton, Joshua D., and John S. Lapinski. 2004. "'Targeted' Advertising and Voter Turnout: An Experimental Study of the 2000 Presidential Election." *Journal of Politics* 66: 69–96.

DellaVigna, Stefano, and Ethan Kaplan. 2007. "The Fox News Effect: Media Bias and Voting." *Quarterly Journal of Economics* 122: 1187–1234.

Franz, Michael M., Paul B. Freedman, Kenneth M. Goldstein, and Travis N. Ridout. 2008. *Campaign Advertising and American Democracy*. Philadelphia: Temple University Press.

Freedman, Paul, Michael Franz, and Kenneth Goldstein. 2004. "Campaign Advertising and Democratic Citizenship." *American Journal of Political Science* 48: 723–41.

Freedman, Paul, and Kenneth Goldstein. 1999. "Measuring Media Exposure and the Effects of Negative Campaign Ads." *American Journal of Political Science* 43: 1189–1208.

Gaines, Brian J., James H. Kuklinski, and Paul J. Quirk. 2007. "The Logic of the Survey Experiment Reexamined." *Political Analysis* 15: 1–20.

Gerber, Alan. 2004. "Does Campaign Spending Work?" *American Behavioral Scientist* 47: 541–74.

Gerber, Alan S., and Donald P. Green. 1998. "Rational Learning and Partisan Attitudes." *American Journal of Political Science* 42: 794–818.

Gilliam, Franklin D., Jr. and Shanto Iyengar. 2000. "Prime Suspects: The Influence of Local Television News on the Viewing Public." *American Journal of Political Science* 44: 560–74.

Green, Donald P., and Lynn Vavreck. 2008. "Analysis of Cluster-Randomized Experiments: A Comparison of Alternative Estimation Approaches." *Political Analysis* 16 (2): 138–52.

Haugtvedt, Curtis P., David W. Schumann, Wendy L. Schneier, and Wendy L. Warren. 1994. "Advertising Repetition and Variation Strategies: Implications for Understanding Attitude Strength." *Journal of Consumer Research.* 21: 176–89.

Havlena, William J. and Jeffrey Graham. 2004. "Decay Effects in Online Advertising: Quantifying the Impact of Time Since Last Exposure on Branding Effectiveness." *Journal of Advertising Research* 44: 327–32.

Hill, Seth J., James Lo, Lynn Vavreck, and John R. Zaller. 2010. "How Quickly We Forget: Campaign Advertising and Electoral Advantage" Working Paper, UCLA Department of Political Science.

Hillygus, D. Sunshine, and Todd G. Shields. 2008. *The Persuadable Voter: Wedge Issues in Presidential Campaigns*. Princeton, NJ: Princeton University Press.

Hovland, C., A. Lumsdaine, and F. Sheffield. 1949. *Experiments in Mass Communication: Studies of Social Psychology in World War II*. Ann Arbor, MI: Association for Consumer Research.

Huber, Gregory, and Kevin Arceneaux. 2007. "Uncovering the Persuasive Effects of Presidential Advertising." *American Journal of Political Science* 24: 957–77.

Iyengar, Shanto, Mark Peters, and Donald R. Kinder. 1982. "Experimental Demonstration of the Not-So-Minimal Consequences of Television News Programs." *American Political Science Review* 76: 848–58.

Jacobson, Gary C. 1978. "The Effects of Campaign Spending in Congressional Elections." *American Political Science Review* 72: 469–91.

Johnston, Richard, Andre Blais, Henry Brady, and Jean Crete. 1992. *Letting the People Decide: Dynamics of a Canadian Election*. Montreal: McGill-Queen's University Press.

Johnston, Richard, Michael G. Hagen, and Kathleen Hall Jamieson. 2004. *The 2000 Presidential Elections and the Foundations of Party Politics*. Cambridge: Cambridge University Press.

Judge, George G., R. Carter Hill, William E. Griffiths, Helmut Lutkepohl, and Tsoung-Chao Lee. 1988. *Introduction to the Theory and Practice of Econometrics*, 2nd ed. New York: Wiley.

Kaid, Lynda Lee. 1997. "Effects of the Television Spots on Images of Dole and Clinton." *American Behavioral Scientist* 40 (8): 1085–94.

Klapper, J. T. 1960. *The Effects of Mass Communications*. New York: Free Press.

Krasno, Jonathan S., and Donald P. Green. 2008. "Do Televised Presidential Ads Increase Voter Turnout? Evidence from a Natural Experiment." *Journal of Politics* 70: 245–61.

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

Lodge, Milton, Marco Steenbergen, and Shawn Brau. 1995. "The Responsive Voter: Campaign Information and the Dynamics of Candidate Evaluation." *American Political Science Review* 89: 309–26.

Miller, Joanne M., and Jon A. Krosnick. 1996. "News Media Impact on the Ingredients of Presidential Evaluations: A Program of Research on the Priming Hypothesis." In *Political Persuasion and Attitude Change*, eds. Diana Mutz, Richard A. Brody, and Paul M. Sniderman. Ann Arbor: University of Michigan Press.

Panagopoulos, Costas, and Donald P. Green. 2008. "Field Experiments Testing the Impact of Radio Advertisements on Electoral Competition." *American Journal of Political Science* 52 (1): 156–68.

Shaw, Daron R. 1999. "The Effect of TV Ads and Candidate Appearances on Statewide Presidential Votes, 1988–96." *American Political Science Review* 93: 345–61.

Stokes, Donald E. 1963. "Spatial Models of Party Competition." *American Political Science Review* 57: 368–77.

Vakratsas, Demetrios, and Tim Ambler. 1999. "How Advertising Works: What Do We Really Know?" *Journal of Marketing* 63: 26–43.

Valentino, Nicholas A., Vincent L. Hutchings, and Dmitri Williams. 2004. "The Impact of Political Advertising on Knowledge, Internet Information Seeking, and Candidate Preference." *Journal of Communication* 54: 337–54.

Vavreck, Lynn. 2007. "The Exaggerated Effects of Advertising on Turnout: The Dangers of Self-reports." *Quarterly Journal of Political Science* 2: 325–43.

Zaller, John. 1992. *The Nature and Origins of Mass Opinion*. New York: Cambridge University Press.

Zaller, John. 1996. "The Myth of Massive Media Impact Revived: New Support for a Discredited Idea." In *Political Persuasion and Attitude Change*, eds. Diana Mutz, Richard A. Brody, and Paul M. Sniderman. Ann Arbor: University of Michigan Press 17–78.

Zielske, Hubert, and Walter Henry. 1980. "Remembering and Forgetting Television Advertisements." *Journal of Advertising Research* 20: 7–13.