

# The Effect of Survey Mode and Sampling on Inferences about Political Attitudes and Behavior: Comparing the 2000 and 2004 ANES to Internet Surveys with Nonprobability Samples

**Neil Malhotra**

*Department of Political Science, Stanford University,  
Encina Hall West, Room 100, Stanford, CA 94305-6044  
e-mail: neilm@stanford.edu (corresponding author)*

**Jon A. Krosnick**

*Departments of Communication, Political Science, and Psychology,  
Stanford University, 434 McClatchy Hall, 450 Serra Mall, Stanford, CA 94305  
e-mail: krosnick@stanford.edu*

Since the inception of the American National Election Study (ANES) in the 1940s, data have been collected via face-to-face interviewing in the homes of members of area probability samples of American adults, the same gold-standard approach used by the U.S. Census Bureau, other federal agencies, and some nongovernment researchers for many of the most high-profile surveys conducted today. This paper explores whether comparable findings about voters and elections would be obtained by a different, considerably less expensive method: Internet data collection from nonprobability samples of volunteer respondents. Comparisons of the 2000 and 2004 ANES data (collected via face-to-face interviewing with national probability samples) with simultaneous Internet surveys of volunteer samples yielded many differences in the distributions of variables and in the associations between variables (even controlling for differences between the samples in reported interest in politics). Accuracy was higher for the face-to-face/probability sample data than for the Internet/volunteer sample data in 88% of the possible comparisons. This suggests that researchers interested in assuring the accuracy of their findings in describing populations should rely on face-to-face surveys of probability samples rather than Internet samples of volunteer respondents.

## 1 Introduction

In the early part of the last century, when contemporary survey research was born, all serious surveys were done via face-to-face interviewing. And since 1948, when the first survey in what became the American National Election Study (ANES) series was conducted, the primary survey data sets analyzed by political scientists have also involved

---

*Authors' note:* We thank Randy Thomas of Harris Interactive and Morris Fiorina for very helpful suggestions. Jon Krosnick is University Fellow at Resources for the Future.

© The Author 2007. Published by Oxford University Press on behalf of the Society for Political Methodology. All rights reserved. For Permissions, please email: journals.permissions@oxfordjournals.org

face-to-face interviewing of probability samples. Face-to-face interviewing has always yielded the highest response rates in surveys of the general public, and the presence of interviewers in respondents' homes has allowed for the development of trust and rapport, presumably enhancing respondent engagement in the reporting process. These are some reasons why the U.S. federal government's most visible and important surveys continue to rely primarily on face-to-face interviewing of probability samples today.

Since the 1970s, the cost of in-person interviewing has increased dramatically (Rossi, Wright, and Anderson 1983), so survey researchers have been motivated to explore more cost-effective alternatives. When telephone service to private households became nearly universal in the United States, a great deal of survey research shifted to random digit dialing (RDD) and telephone interviewing. Not only did the telephone cut costs considerably but also it allowed for centralized monitoring and supervising of interviewers and computer presentation of questionnaires, which were perceived to be tools for enhancing the quality of the data collection exercise. And because no interviewer travel time is required, telephone surveys can be conducted more quickly than face-to-face surveys.

Unfortunately, however, we have learned that these advantages of telephone data collection are countervailed by significant disadvantages (see, e.g., Holbrook, Green, and Krosnick 2003): telephone response rates are lower than face-to-face response rates and respondents and interviewers rush through telephone conversations, yielding lower accuracy of reporting. Furthermore, the physical and psychological distance between telephone interviewers and respondents reduces respondent trust and rapport, leading to less willingness to admit embarrassing attributes and more social desirability bias in reporting. Respondents become impatient and unwilling to continue telephone interviews after relatively few questions have been asked, and visual aids cannot be used during telephone interviews as they can in face-to-face interviews. The rising difficulty of reaching respondents when they are willing to complete a telephone interview has driven costs up significantly in recent years as response rates have been dropping.<sup>1</sup> So the cost savings offered by the telephone (relative to face-to-face interviewing) also bring with them significant compromises in measurement quality.

With the emergence of the Internet, survey researchers inside academia and out have been intrigued with the possibility that this medium might permit cost-effective and quick data collection. The possibility of using computers as such was pioneered by Willem Saris, Professor at the University of Amsterdam, who created the Dutch "Telepanel" in the 1980s by recruiting a representative sample of Dutch citizens and giving them computers and modems (Saris 1998). Each week, these panel members completed an electronic questionnaire and telephoned their responses in to a central computer via their modems. This approach was adopted by the U.S. firm Knowledge Networks and by firms in other countries as well (e.g., TNS Forsa in Germany). Holding constant the incentives offered, it has become increasingly difficult to recruit respondents to join and stay in such panels, so the effective response rates for weekly surveys have dropped to relatively low levels. Nonetheless, data collection from a probability sample can be accomplished via this medium in a timely fashion, though at a cost about equivalent to that of RDD telephone interviewing.

---

<sup>1</sup>Telephone surveys could achieve higher response rates at higher costs. However, there are several structural characteristics of telephone interviewing (e.g., cell phone usage, answering machines, caller ID) that reduce response rates regardless of expenditure.

In this light, researchers have been increasingly open to the possibility of scrapping representative sampling altogether and collecting data via the Internet from volunteer samples instead. Dozens of online survey data collection firms have been created in recent years and are doing brisk business, mostly with commercial clients. These firms recruit respondents by placing banner advertisements on Web pages, purchasing lists of e-mail addresses from commercial suppliers, and by other means, and once a person has signed up to be a member of an opt-in online panel, he or she is invited occasionally to complete a questionnaire (either for no reward or for a financial or nonfinancial reward). For each survey, the demographics of the sample can be fine-tuned by imposing quotas at the time of data collection (e.g., only half the respondents can be female) and/or at the time of sampling (e.g., invitees can be chosen in numbers designed to produce a sample that resembles the population with regard to some demographic variables). In addition, post-stratification weighting can be implemented to yield the appearance of demographic representativeness.

Political scientists have been particularly interested in exploring this alternative method of data collection because of the remarkably low price entailed. One possibility under consideration in some circles is replacing face-to-face interviewing of probability samples with Internet surveying of nonprobability, volunteer respondents. The latter sort of data can be used for at least two purposes: to gauge the prevalence of various characteristics in the population (e.g., what proportion of Americans voted in the last presidential election, and for whom?) and to gauge relations between variables (e.g., to what extent is turnout related to particular posited predictors of it?). In fact, a large-scale cooperative political science survey research project (the "Cooperative Congressional Election Study") was mounted at the time of the 2006 U.S. elections by the firm Polimetrix, whose Internet survey respondents are volunteers not sampled by probability methods.

In doing the work described in this paper, we set out to explore the viability of this method as a replacement for probability sampling and face-to-face interviewing in the United States. We compared two national face-to-face surveys of probability samples, done in 2000 and 2004, to surveys conducted at the same times with volunteer samples of Americans via the Internet. Our findings cannot be used to determine whether any observed differences are due to mode of data collection or method of sampling or both. But we can ascertain whether results would be different if we replaced face-to-face interviewing of probability samples with Internet data collection from volunteer samples.

As with other methods that employ volunteers and quotas, Internet surveying may produce misleading results partly because respondents select themselves into the sample in ways unrepresentative of the nation (Erikson and Tedin 2007, 37). For instance, if wealthier individuals are more likely to participate, then a data set will be skewed toward the preferences and behaviors of high-income respondents. Weighting low-income respondents more heavily to account for such discrepancies is not a panacea because the low-income respondents who choose to participate may not be representative of low-income individuals in the population at large. If that is true, then weighting to increase the proportion of low-income respondents would only introduce more sample composition bias. To assess whether weighting has constructive impact on results, we determined whether differences between the data collection methods varied depending upon whether we analyzed weighted Internet data or unweighted Internet data.

This article is organized as follows. We begin by describing the data sets we analyzed and our methods of analysis. Next, we present comparisons of results obtained by the two data collection modes/sampling methods. Finally, we discuss the implications of the findings for survey research practice in the future.

## 2 Data

We analyzed four data sets: (1) the 2004 ANES survey, conducted face-to-face with a national area probability sample, (2) survey data collected from a national sample of volunteer American adults by the firm YouGov at the time of the 2004 presidential election via the Internet, (3) the 2000 ANES, also conducted face-to-face with a national area probability sample,<sup>2</sup> and (4) survey data collected at the time of the 2000 presidential election from a national sample of volunteer American adults by Harris Interactive.

### 2.1 2004 ANES

For the 2004 ANES, the Survey Research Center at the University of Michigan designed a multistage area probability sample. Geographically defined sampling units of decreasing size were selected with probabilities proportionate to their total number of occupied housing units. Interviewers then visited the selected households to conduct face-to-face interviews with a randomly selected adult U.S. citizen residing in each household. From a respondent pool of 1833 people, 1212 people were interviewed preelection between September 7 and November 1, 2004, yielding a response rate of 66%. In all, 1066 of these respondents were reinterviewed after the election between November 3 and December 20, 2004, yielding a reinterview rate of 88%. Poststratification weights were built using the U.S. Census Bureau's Current Population Survey (CPS) so that the sample proportions matched the population proportions on age and education. In addition, because one adult was randomly selected from each household, the data were weighted by the number of adults in the household (capped at three) so as to correct for unequal respondent selection probabilities. And metropolitan statistical areas and regions that had lower response rates were weighted more heavily to compensate for their underrepresentation.

### 2.2 2004 YouGov<sup>3</sup>

The survey firm YouGov collected data from national samples of American adults ( $N = 1747\text{--}3634$  per wave) weekly for 18 weeks between July and November, 2004. We focus on the nine waves of data collected between September 7 and November 2, so as to overlap with the preelection administration period of the ANES, and one postelection wave collected in early November. All data were collected via the Internet. YouGov recruited a diverse set of panel members and then constructed samples that resembled the nation as a whole in terms of particular demographics.

Respondents were recruited using three methods. First, YouGov contacted previous participants in their American polls via e-mail and solicited their participation. Second, people who were on electronic mailing lists purchased from Internet marketing firms were contacted and invited to join the panel. Third, advertisements inviting people to join the panel appeared on Google search pages when people conducted searches on various hobby and interest topics, such as politics, sports, health and fitness, and entertainment.

The links and e-mail solicitations directed prospective respondents to a Web site that asked them to answer a survey question related to the topic of their search. For example,

---

<sup>2</sup>A portion of the 2000 ANES was conducted via telephone using RDD sampling methods. These cases are excluded in the analyses that follow.

<sup>3</sup>Information about the YouGov methodology was obtained via personal communication with Stephan Shakespeare, Director of Public Opinion Research.

individuals who typed the search term “music” into Google were asked a question such as: “Which of the following people is your favorite singer?” After answering the question, visitors were asked if they would like to participate in longer surveys in exchange for monetary compensation. If they agreed, YouGov then asked them for an e-mail address, as well as their gender, date of birth, educational attainment, religious affiliation, ethnicity, employment status, income, marital status, home ownership status, media consumption, and party identification.

Based on the demographic information collected, respondents were invited to participate in polls in proportions so that the final samples would resemble the nation in terms of distributions of age, race, sex, and education. For our analyses, the collected data were weighted using weights calculated by YouGov based on CPS figures.<sup>4</sup>

People were invited to participate in later survey waves if they responded favorably to initial participation invitations.<sup>5</sup>

### 2.3 2000 ANES

For the 2000 ANES, the Survey Research Center at the University of Michigan designed a national area probability sample and interviewed 1006 respondents face-to-face before the election (September 5 to November 6, 2000, response rate 64.3%) and interviewed 694 of these people again after the election (November 8 to December 18, 2000, reinterview rate 69%). Poststratification weights were built using the CPS.

### 2.4 2000 Harris Interactive

As of 2000, Harris Interactive had recruited a panel of more than 7 million volunteer respondents, more than three-quarters of them from one of the most popular Internet search engines: <http://www.excite.com>. On the main page of [excite.com](http://www.excite.com), a link appeared inviting visitors to participate in the poll of the day. Respondents who voted on the day’s issue then saw a link inviting them to become panel members for the Harris Poll Online (HPOL). A second source of panel members was the Web site of Matchlogic, Inc., an online marketing company and a subsidiary of Excite@Home. Matchlogic posted banner advertisements on the Internet to attract consumers with promises of incentives, such as free merchandise and sweepstakes. When a person registered for those incentives, he or she was invited to become a panel member for HPOL. Excite and Matchlogic accounted for about 90% of all panel members.

People visiting the HPOL registration Web site were told that HPOL offered them the opportunity to express their views, hopes, and opinions and that influential decision makers in government and business would hear what they think. Then respondents were asked for their e-mail addresses and demographic information (e.g., gender, date of birth, occupation, education, marital status, ethnicity, income, and primary language used at home).

Respondents could click on a link to a Web page explaining the benefits of HPOL membership: that they could influence important decision makers in government,

---

<sup>4</sup>The maximum weights for YouGov and the 2004 ANES were 13.82 and 3.03, respectively. The minimum weights were .10 and .36, respectively. In all surveys, weights were constructed only from demographic characteristics of the respondents. It may also be possible to weight post hoc based on the election result and actual turnout figures.

<sup>5</sup>Some respondents may have provided data during multiple months between September and November and during previous waves. However, YouGov did not inform us about who these respondents were. Consequently, we could not predict actual vote choice with variables measured during the preelection survey.

nonprofit organizations, and corporations; could help to shape policies, products, and services; would have access to some survey results prior to publication in the media; and might win cash, free consumer products, or discount coupons or receive other tangible incentives.

After individuals signed up to be panel members, they received e-mail invitations to complete questionnaires. A panel member who was invited to do a survey could not be invited to do another survey for at least 10 days, and most panel members received an invitation at least once every few months.

For this study, Harris selected the sample by stratifying their panel on gender, age, and region. Respondents were selected such that they were similar to population parameters (from CPS data) in terms of distributions of age and region of residence, and males were oversampled because Harris expected that nonrespondents were more likely to be male than female. Each selected panel member was sent an e-mail invitation that described the topic of the survey and provided a hyperlink to the Web site where the questionnaire was posted and a unique password allowing access to the questionnaire.

Data were collected between October 5 and 25, 2000, concurrent with the fielding dates of the ANES preelection survey. A total of 41,393 participants from the HPOL database were invited to complete the preelection questionnaire, and 3980 of them did so, yielding a participation rate of 9.6%. The postelection interviews were conducted during November, 2000. Of the original participants, 2024 completed the follow-up questionnaire, yielding a participation rate of 50.9%.

Harris Interactive computed weights based on gender, age, education, race, ethnicity, and Internet access to match the sample to estimates from the CPS and one of Harris Interactive's monthly RDD telephone surveys.<sup>6</sup> We were given two sets of weights to use: uncapped and capped. The uncapped weights were produced for poststratification adjustments. Survey firms often impose maxima (or caps) on such weights so that no single respondent will be weighted disproportionately heavily. We chose to focus on the capped weights for this reason. However, we also analyzed the Harris data using uncapped weights, and the results were statistically and substantively similar to those reported in the text.

### 3 Measures

Ideally, we would compare responses to questions worded identically in the two modes/sampling methods and asked within identical questionnaires. Unfortunately, the questionnaires employed in the two modes/sampling methods were not identical and relatively few questions were asked with identical wordings, because the surveys were not designed to permit the comparisons reported here. Therefore, the conclusions we reach about items not asked identically should be treated as tentative, pending conduct of comparable studies in the United States using identical questionnaires in the two modes/sampling methods. For this investigation, we analyzed all questions that measured the same construct in comparable ways in the two simultaneous surveys. Political attitude questions involving rating scales must have offered the same number of points on the rating scale in both surveys in order to be analyzed. The question wordings, response options, and codings are presented in the Appendix. All variables were coded such that the minimum possible value was 0 and

---

<sup>6</sup>The maximum weights for the Harris and the 2000 ANES preelection surveys were 5.00 and 3.20, respectively. The minimum weights were .20 and .34, respectively. The maximum weights for the Harris and 2000 ANES postelection surveys were 5.00 and 3.15, respectively. The minimum weights were .20 and .29, respectively.

the maximum possible value was 1. Variables with some political valence were coded such that 1 represented the response most supportive of President Bush and his party and policy positions and 0 represented the response least supportive of President Bush and his party and policy positions.

In some cases, judgment calls were made to produce comparable coding of slightly different response categories. For example, “quite well” and “very well” were the second-highest response options on four-point rating scales for questions measuring perceptions of candidate personality traits in the 2000 ANES and Harris questionnaires, respectively. These two answers were coded identically (.67) in our analyses. In some cases, response categories were combined to allow direct comparisons between data sets, and this sometimes produced different sets of response categories for the same variable in 2000 and 2004. For example, “bachelor’s degree” and “associate’s degree” were combined into the “College” category for the 2004 ANES education variable, and “college graduate” and “completed some graduate school, no degree” were combined into “College” for the 2004 YouGov survey. Such inexact matches of measures are further reason to view the present findings involving such measures as suggestive but not conclusive.

#### 4 Analytic Method

To evaluate the differences between the data sets (2004 ANES versus 2004 YouGov, and 2000 ANES versus 2000 Harris), we used two approaches. First, we compared the distributions of responses, testing for statistically significant differences. Second, we estimated bivariate logistic regression equations predicting predicted vote choice (reported preelection), actual vote choice (reported postelection), and actual turnout (reported postelection) with theoretically sensible predictors, testing for interactions of survey mode/sampling method with political and demographic variables to gauge whether the predictors were more powerfully related to the dependent variable in one mode/sampling method than the other. We did not conduct multivariate regressions because the sets of predictors in such regressions can be specified in many ways, and results may differ depending upon which predictors are included or excluded. Because researchers often predict both predicted and actual vote choice using variables measured preelection, we took this same approach.

The logistic regression equations we estimated took this form:

$$\text{Logit}(\text{Pr}(Y_i = 1)) = \alpha + \beta_1 X_i + \beta_2 I_i + \beta_3 (X_i I_i) + \varepsilon_i, \quad (1)$$

where  $Y_i$  is either predicted or actual vote choice (vote for President Bush = 1, vote for Senator Kerry or Vice President Gore = 0) or actual turnout (voted in the election = 1, did not vote = 0),  $X_i$  represents the predictor of interest,  $I_i$  represents a dummy variable for survey mode/sampling method (Internet/nonprobability = 1, face-to-face/probability = 0), and  $\varepsilon_i$  represents random error.<sup>7</sup> The interaction between the predictor and the survey mode/sampling method dummy tests whether the association of the predictor with the dependent variable differed between modes/sampling methods.

We estimated equation (1) separately for each predictor and dependent variable in the 2004 and 2000 data. For the logistic regressions predicting predicted and actual vote choice, we tested every  $X_i$  for which the question wording and response options were identical or highly similar across the data sets. Because YouGov did not provide us with

<sup>7</sup>For all statistical analyses using ANES data, cases were modeled as being clustered by primary sampling unit (PSU). In analyses pooling ANES and Internet data, each Internet case was treated as a separate PSU to permit proper representation of the clustering of the face-to-face cases.

respondent identifiers to link data from the preelection survey to data in the postelection survey, we could only estimate logistic regressions predicting actual vote choice with the 2000 data. In the logistic regressions predicting postelection turnout, we included as predictors the demographics that have been shown to be related to turnout in past studies.

We also tested a specific substantive hypothesis. As is typical of Internet surveys with volunteer samples, YouGov and Harris invited their respondents to participate in these particular surveys with e-mails mentioning that the topic was politics. This makes it possible for respondents to choose whether or not to participate in the surveys based upon how interested in politics they were. That is, people more interested in politics may have been more likely to participate. In the case of the ANES surveys, potential respondents were told in advance that the survey was about elections, but interviewers made aggressive efforts to recruit all selected respondents regardless of their interest in politics. So this self-selection bias may have been less, since people with low interest in politics would have had to opt “out” of the ANES (perhaps relatively more difficult to do in the face of an interviewer), whereas the YouGov and Harris surveys would have involved people with low interest in politics failing to opt “in” (perhaps easier to do, since no force encouraged participation).

If in fact political interest was higher in the Internet/volunteer samples than among the ANES samples, then we might expect to see greater consistency among the political judgments reported by the former individuals than among those reported by the latter individuals (see, e.g., Converse 1964). To test whether differences between the modes/sampling methods in the strength of the bivariate associations were attributable to differences in political interest, we made use of a question that was asked in the two 2004 questionnaires to measure interest in campaigns, which allowed us to estimate the following equation:

$$\text{Logit}(\Pr(Y_i = 1)) = \alpha + \beta_1 X_i + \beta_2 I_i + \beta_3 P_i + \beta_4 (X_i I_i) + \beta_5 (X_i P_i) + \varepsilon_i, \quad (2)$$

where  $P_i$  represents political interest. The interaction between  $X_i$  and survey mode/sampling method gauges the effect of mode/sampling method on the strength of association, controlling for the moderating impact of political interest.

We conducted all of our analyses twice, first using weighted data and then using unweighted data. The weights provided by the survey firms are each only one of many possible sets of weights that could be used, so results might have been different if different weights had been employed. But because most investigators probably rely on weights provided to them by survey firms, we have done the same here.

## 5 Results

### 5.1 Marginal Distributions

The face-to-face surveys were more accurate in comparisons of unweighted frequencies of the sample demographics with the CPS (see Table 1). In general, deviations of the ANES samples from the CPS were smaller than were deviations of the Internet samples from the CPS.<sup>8</sup> Furthermore, statistically significant differences appeared between the survey firms with respect to the distributions of nearly all the demographic and political variables, both unweighted and weighted (see Tables 2–4). Some of these differences are

<sup>8</sup>Both the CPS and ANES data were collected from U.S. citizens only. The Internet survey samples may have included residents who were not U.S. citizens.



**Table 1** Marginal distributions for demographic variables compared to the CPS (unweighted)

<i>2004 data</i>	<i>ANES</i> (%)	<i>YouGov</i> (%)	<i>Nov. CPS</i> (%)	<i>ANES –</i> <i>CPS (%)</i>	<i>YouGov –</i> <i>CPS (%)</i>	N ( <i>ANES</i> )	N ( <i>YouGov</i> )
<b>Gender</b>							
Male	46.7	40.3	47.6	-0.9	-7.3***	566	10,642
Female	53.3	59.7	52.4	0.9	7.3***	646	15,730
						1212	26,372
<b>Race</b>							
Black	14.8	6.0	10.5	4.3†	-4.5***	180	1580
Others	85.2	94.0	89.5	-4.3***	4.5***	1032	24,792
						1212	26,372
<b>Education</b>							
Less than high school	9.2	3.2	14.3	-5.1†	-11.1***	111	834
High school	29.3	15.4	32.3	-3.0	-16.9***	355	4061
Some college	21.8	36.0	19.2	2.6	16.8***	264	9480
College	28.3	33.3	25.5	2.8	7.8***	343	8791
Graduate	11.5	12.2	8.7	2.8	3.5***	139	3206
						1212	26,372
<b>Age</b>							
18–29	19.7	15.9	19.9	-0.2	-4.0***	239	4199
30–39	16.2	21.7	18.3	-2.1	3.4***	197	5733
40–49	19.2	28.2	21.2	-2.0	7.0***	233	7444
50–64	27.6	27.7	23.8	3.8	3.9***	335	7307
65+	17.2	6.4	16.9	0.3	-10.5***	208	1689
						1212	26,372
<i>2000 data</i>	<i>ANES</i> (%)	<i>Harris</i> (%)	<i>Nov. CPS</i> (%)	<i>ANES –</i> <i>CPS (%)</i>	<i>Harris –</i> <i>CPS (%)</i>	N ( <i>ANES</i> )	N ( <i>Harris</i> )
<b>Gender</b>							
Male	44.6	45.5	48.0	-3.4	-2.5*	449	1811
Female	55.4	54.5	52.0	3.4	2.5*	557	2169
						1006	3980
<b>Race</b>							
Black	12.9	3.3	11.9	1.0	-8.6***	129	132
Others	87.1	96.7	88.1	-1.0	8.6***	869	3819
						998	3951
<b>Education</b>							
Less than high school	11.1	1.9	16.9	-5.8†	-15.0***	111	76
High school	30.4	12.3	32.8	-2.4	-20.5***	306	487
Some college	19.4	40.7	19.8	-0.4	20.9***	195	1612
College	29.4	32.3	23.0	6.4*	9.3***	296	1280
Graduate	9.7	12.8	7.5	2.2	5.3***	98	506
						1006	3961
<b>Age</b>							
18–24	9.1	11.8	13.2	-4.1	-1.4	92	471
25–34	17.8	18.9	18.7	-0.9	0.2	179	752
35–44	23.1	25.8	22.1	1.0	3.7**	232	1027
45–54	17.1	23.4	18.3	-1.2	5.1***	172	931
55–64	13.9	12.8	11.6	2.3	1.2	140	508
65+	19.0	7.3	15.9	3.1	-8.6***	191	291
						1006	3980

\*\*\* $p < .001$ ; \*\* $p < .01$ ; \* $p < .05$ ; † $p < .10$  (two-tailed).

**Table 2** Marginal distributions of demographic variables

2004 data	Unweighted data (%)			Weighted data (%)		
	ANES	YouGov	Difference	ANES	YouGov	Difference
Gender						
Male	46.7	40.3	6.4**	48.5	48.9	-0.4
Female	53.3	59.7	6.4**	51.5	51.1	0.4
Race						
Black	14.8	6.0	8.8**	15.6	12.9	2.7
Others	85.2	94.0	-8.8***	84.4	87.1	-2.7*
Education						
Less than high school	9.2	3.2	6.0*	14.5	3.2	11.3***
High school	29.3	15.4	13.9***	31.4	13.6	17.8***
Some college	21.8	36	-14.2***	19.6	37.2	-17.6***
College	28.3	33.3	-5.0*	24.9	32.4	-7.5**
Graduate	11.5	12.2	-0.7	9.5	13.6	-4.1
Age						
18-29	19.7	15.9	3.8	20.7	20.5	0.2
30-39	16.2	21.7	-5.5*	17.7	19.1	-1.4
40-49	19.2	28.2	-9.0***	20.3	23.9	-3.6
50-64	27.6	27.7	-0.1	24.1	20.2	3.9
65+	17.2	6.4	10.8***	17.2	16.2	1.0
2000 data						
	ANES	Harris	Difference	ANES	Harris	Difference
Gender						
Male	44.6	45.5	-0.9	44.6	47.7	-3.1
Female	55.4	54.5	0.9	55.4	52.3	3.1
Race						
Black	12.9	3.3	9.6**	13.3	9.3	4.0
Others	87.1	96.7	-9.6***	86.7	90.7	-4.0***
Education						
Less than high school	11.1	1.9	9.2**	16.3	6.3	10.0*
High school	30.4	12.3	18.1***	35.0	39.0	-4.0
Some college	19.4	40.7	-21.3***	18.1	29.1	-11.0***
College	29.4	32.3	-2.9	23.6	20.4	3.2
Graduate	9.7	12.8	-3.1	7.0	5.1	1.9
Age						
18-24	9.1	11.8	-2.7	14.0	13.5	0.5
25-34	17.8	18.9	-1.1	17.9	17.4	0.5
35-44	23.1	25.8	-2.7	21.1	24.7	-3.6
45-54	17.1	23.4	-6.3*	16.0	19.9	-3.9
55-64	13.9	12.8	1.1	12.5	11.4	1.1
65+	19.0	7.3	11.7***	18.5	13.1	5.4

\*\*\* $p < .001$ ; \*\* $p < .01$ ; \* $p < .05$  (two-tailed).

not only highly significant but also substantively large. Thus, analysts would reach different conclusions about trends over time in distributions if comparing ANES data collected in a prior year with these data, depending upon which mode's/sampling method's data were used.

In general, the mode/sampling method differences in terms of demographic variable distributions were much more pronounced with unweighted data than with weighted data (see Table 2). This is because the weights were explicitly constructed to match the two samples to the same benchmarks (i.e., CPS distributions). Thus, it is especially surprising that the weights failed to eliminate all differences between the modes/sampling methods in this regard, which is attributable to suboptimal construction of the weights. For example, the proportions of respondents in the two lowest education categories in 2004 were between 11.3 and 17.8 percentage points lower in the Internet sample than the ANES sample, even after weighting.

It is equally striking that applying the weights did almost nothing to reduce the differences between modes/sampling methods in the distributions of the political variables (see Tables 3 and 4). These differences were often strikingly large, in some cases becoming even larger after the weights were applied. For instance, the reported 2004 turnout rate with weighted data was 15 percentage points higher in the Internet sample than in the ANES, up from 12.5 percentage points unweighted. In 2000, the difference was even larger weighted: 18.7 percentage points. Weighted turnout exceeded 88% in both years' Internet data, regardless of whether the weights were or were not applied, considerably higher than the overestimates provided by the ANES data. This suggests that it may be difficult to identify predictors of turnout in the Internet samples with confidence because those samples included almost no nonvoters to study. Also as expected, interest in politics was considerably larger in the 2004 Internet sample than in the ANES sample, by 9.9 percentage points weighted. Differences between the modes/sampling methods in the proportions of respondents providing various responses to other substantive variables were often large. For example, the economic evaluation questions sometimes manifested mode/sampling method differences exceeding 15 percentage points, even after weighting.

Only two of the 27 substantive variables we examined failed to manifest a statistically significant difference between the modes/sampling methods in the distributions of responses. Responses to the predicted and actual vote choice questions were remarkably similar in the two 2004 data sets, using both unweighted and weighted data. Interestingly, both data sets showed John Kerry winning the popular vote for the weighted data. In 2000, the ANES data showed Al Gore winning the popular vote, whereas the Internet data showed George Bush winning, and the gap between modes/sampling methods was considerable: 12.1 percentage points unweighted for predicted vote choice and 10.0 percentage points unweighted for actual vote choice.

## 5.2 *Logistic Regressions Predicting Predicted and Actual Vote Choice*

The relations of substantive and demographic variables with predicted vote choice often differed considerably between the ANES and Internet data. Parameter estimates for equation (1) with unweighted and weighted data appear in Table 5. The first and seventh columns of coefficients show the effect of each predictor in the ANES data, and the third and ninth columns show the increase or decrease in that effect in the Internet data. All the estimated effects of substantive variables are in the expected, positive direction in all four data sets. Furthermore, the one demographic effect that we can predict with most confidence (African Americans were more supportive of the Democratic candidate) appeared in all four data sets.

However, striking differences between the modes/sampling methods appeared in both years. In 2004, party identification was a considerably weaker predictor of vote choice in

**Table 3** Marginal distributions for 2004 ANES and YouGov data, political variables

	<i>Unweighted data (%)</i>			<i>Weighted data (%)</i>			<i>Unweighted Ns</i>	
	<i>ANES</i>	<i>YouGov</i>	<i>Difference</i>	<i>ANES</i>	<i>YouGov</i>	<i>Difference</i>	<i>ANES</i>	<i>YouGov</i>
<b>Predicted vote choice</b>								
Bush	49.4	52.8	-3.4	48.9	49.7	-0.8	545	11,306
Kerry	50.6	47.2	3.4	51.1	50.3	0.8	558	12,624
							1103	23,930
<b>Actual vote choice</b>								
Bush	49.8	52.4	-2.6	49.6	49.5	0.1	354	1746
Kerry	50.2	47.6	2.6	50.4	50.5	-0.1	357	1587
							711	3333
<b>Actual turnout</b>								
Voted	82.1	94.6	-12.5***	79.9	94.9	-15.0***	441	3428
Did not vote	17.9	5.4	12.5**	20.1	5.1	15.0***	96	194
							537	3622
<b>Party ID</b>								
Strong Democrat	17.0	19.1	-2.1	16.5	21.2	-4.7†	203	2122
Weak Democrat	15.0	13.3	1.7	15.6	13.5	2.1	179	1474
Lean Democrat	17.6	9.1	8.5**	17.5	9.0	8.5**	210	1013
Independent	9.9	15.6	-5.7*	9.7	15.2	-5.5†	118	1728
Lean Republican	11.6	8.8	2.8	11.7	8.6	3.1	138	971
Weak Republican	12.9	12.9	0.0	12.5	12.2	0.3	154	1428
Strong Republican	16.2	21.2	-5.0†	16.5	20.4	-3.9	193	2351
							1195	11,087
<b>Bush job approval</b>								
Strongly approve	33.7	27.4	6.3**	34.0	26.1	7.9**	398	5956
Somewhat approve	17.2	20.4	-3.2	16.9	18.8	-1.9	203	4435
Somewhat disapprove	11.6	15.0	-3.4	11.6	14.8	-3.2	137	3261
Strongly disapprove	37.5	37.2	0.3	37.5	40.3	-2.8	442	8100
							1180	21,752
<b>Iraq worth it</b>								
Worth it	39.4	50.9	-11.5***	39.9	47.2	-7.3**	465	3991
Not worth it	60.6	49.1	11.5***	60.1	52.8	7.3***	714	3851
							1179	7842
<b>Sociotropic retrospective</b>								
Gotten much better	3.5	9.1	-5.6†	3.7	9.7	-6.0†	42	848
Gotten somewhat better	19.7	28.9	-9.2***	19.9	26.6	-6.7*	236	2698
Stayed about the same	31.6	22.2	9.4***	30.9	22.6	8.3**	378	2075
Gotten somewhat worse	25.7	22.4	3.3	25.7	22.9	2.8	308	2093
Gotten much worse	19.5	17.5	2.0	19.7	18.2	1.5	234	1631
							1198	9345
<b>Sociotropic prospective</b>								
Get much better	8.3	12.7	-4.4	8.9	13.2	-4.3	96	1123
Get somewhat better	27.3	36.6	-9.3***	27.7	35.6	-7.9**	316	3235
Stay about the same	46.4	29.0	17.4***	46.2	29.3	16.9***	537	2564

Continued

Table 3 (continued)

	Unweighted data (%)			Weighted data (%)			Unweighted Ns	
	ANES	YouGov	Difference	ANES	YouGov	Difference	ANES	YouGov
Get somewhat worse	13.8	14.1	-0.3	13.5	14.3	-0.8	160	1249
Get much worse	4.2	7.6	-3.4	3.7	7.7	-4.0	48	674
							1157	8845
Pocketbook retrospective								
Gotten much better	12.0	6.2	5.8*	12.5	6.3	6.2*	144	576
Gotten somewhat better	31.5	22.8	8.7***	30.9	21.7	9.2***	378	2136
Stayed about the same	25.0	29.6	-4.6†	25.0	30.1	-5.1†	301	2770
Gotten somewhat worse	22.1	24.2	-2.1	21.8	23.8	-2.0	265	2267
Gotten much worse	9.5	17.2	-7.7**	9.8	18.1	-8.3**	114	1611
							1202	9360
Pocketbook prospective								
Get much better	15.7	15.8	-0.1	16.6	15.9	0.7	183	1415
Get somewhat better	23.0	39.0	-16.0***	22.9	37.7	-14.8***	267	3488
Stay about the same	52.5	29.0	23.5***	52.5	29.6	22.9***	610	2589
Get somewhat worse	6.7	10.7	-4.0	5.8	10.7	-4.9†	78	955
Get much worse	2.2	5.5	-3.3	2.2	6.1	-3.9	25	495
							1163	
Interest in campaigns								
Very much interested	41.1	47.5	-6.4**	39.8	49.7	-9.9***	498	12,513
Somewhat interested	43.6	39.5	4.1†	43.7	37.7	6.0**	528	10,390
Not much interested	15.4	13.0	2.4	16.5	12.6	3.9	186	3432
							1212	26,335

Note. Actual vote choice and actual turnout were measured postelection.

\*\*\* $p < .001$ ; \*\* $p < .01$ ; \* $p < .05$ ; † $p < .10$  (two-tailed).

the Internet data than in the ANES data. But for almost all other substantive predictors in both years, the observed association was considerably stronger in the Internet data than in the ANES data. Significant mode/sampling method differences also appeared for some of the demographic predictors as well. Especially striking, the effects of some of the education dummy variables were reversed in sign in the two modes/sampling methods. Even when controlling for differences in political interest between the 2004 samples using equation (2), the mode/sampling method differences in associations mostly remained (see the top panel of Table 6).

In addition to being statistically significant, the differences between modes/sampling methods in the strengths of the bivariate relations are substantively large as well. The 6th and 12th columns of Table 5 show our measure of "effect size," which is the difference of the differences between predicted probabilities when moving across the interquartile range of the independent variable between the two modes/sampling methods.<sup>9</sup> For instance,

<sup>9</sup>Since "Male" and "Black" are dummy variables, we compared the differences in the predicted probabilities of having a value of 0 compared to having a value of 1. We did not calculate effect sizes for education or age because the constructs were measured by a series of dummy variables with a reference category, rendering our approach of examining the interquartile range nonimplementable.

**Table 4** Marginal distributions for 2000 ANES and Harris data, political variables

	<i>Unweighted data (%)</i>			<i>Weighted data (%)</i>			<i>Unweighted Ns</i>	
	<i>ANES</i>	<i>Harris</i>	<i>Difference</i>	<i>ANES</i>	<i>Harris</i>	<i>Difference</i>	<i>ANES</i>	<i>Harris</i>
<b>Predicted vote choice</b>								
Bush	44.1	56.2	-12.1***	44.1	53.4	-9.3***	382	1838
Gore	55.9	43.8	12.1***	55.9	46.6	9.3***	485	1433
							867	3271
<b>Actual vote choice</b>								
Bush	45.7	55.7	-10.0**	46.6	53.1	-6.5†	223	1527
Gore	54.3	44.3	10.0**	53.4	46.9	6.5†	265	1215
							488	2742
<b>Actual turnout</b>								
Voted	73.6	91.3	-17.7***	69.4	88.1	-18.7***	634	1986
Did not vote	26.5	8.7	17.7***	30.6	11.9	18.7***	228	190
							862	2176
<b>Party ID</b>								
Strong Democrat	19.6	18.8	0.8	19.6	23.8	-4.2	194	694
Weak Democrat	16.6	13.2	3.4	17.0	13.0	-4.0	164	488
Lean Democrat	16.7	9.4	7.3*	16.7	8.6	8.1*	165	349
Independent	10.8	9.4	1.4	12.0	11.2	0.8	107	349
Lean Republican	14.0	10.0	4.0	14.0	9.4	4.6	138	371
Weak Republican	11.3	14.4	-3.1	10.5	13.3	-2.8	112	534
Strong Republican	10.9	24.8	-13.9***	10.3	20.6	-10.3**	108	917
							988	3702
<b>Gun control</b>								
A lot more difficult	45.5	38.4	7.1**	44.3	37.6	6.7*	454	1527
Somewhat more difficult	13.3	18.9	-5.6†	13.9	19.1	-5.2	133	754
About the same	36.7	30.0	6.7*	37.6	31.3	6.3*	366	1195
Somewhat easier	2.5	7.0	-4.5	2.5	6.2	-3.7	25	279
A lot easier	2.0	5.7	-3.7	1.8	5.8	-4.0	20	225
							998	3980
<b>Clinton retrospective: economy</b>								
Much better	28.7	36.3	-7.6†	27.9	35.3	-7.4†	145	1444
Somewhat better	33.7	33.7	0.0	34.9	32.8	2.1	170	1343
No difference	32.7	20.3	12.4**	31.9	21.1	10.8**	165	808
Somewhat worse	3.4	6.7	-3.3	3.7	7.0	-3.3	17	267
Much worse	1.6	3.0	-1.4	1.6	3.8	-2.2	8	118
							505	3980
<b>Clinton retrospective: crime</b>								
Much better	9.9	8.9	1.0	8.7	8.5	0.2	50	354
Somewhat better	27.8	31.2	-3.4	24.9	29.5	-4.6	140	1240
No difference	31.2	33.9	-2.7	33.0	34.4	-1.4	157	1349
Somewhat worse	16.1	18.2	-2.1	17.2	19.2	-2.0	81	725
Much worse	15.1	7.8	7.3†	16.3	8.5	7.8†	76	312
							504	3980

Continued

Table 4 (continued)

	<i>Unweighted data (%)</i>			<i>Weighted data (%)</i>			<i>Unweighted Ns</i>	
	<i>ANES</i>	<i>Harris</i>	<i>Difference</i>	<i>ANES</i>	<i>Harris</i>	<i>Difference</i>	<i>ANES</i>	<i>Harris</i>
<b>Gore: moral</b>								
Extremely well	26.4	19.7	6.7*	25.7	19.8	5.9†	249	783
Quite well	49.5	27.0	22.5***	49.6	28.8	20.8***	466	1074
Not too well	18.5	31.0	-12.5***	18.7	29.4	-10.7***	174	1232
Not well at all	5.6	22.4	-16.8***	6.0	22.0	-16.0***	53	891
							942	3980
<b>Gore: cares about people</b>								
Extremely well	17.2	14.1	3.1	18.2	14.5	3.7	165	562
Quite well	44.2	20.2	24.0***	43.5	21.5	22.0***	423	803
Not too well	27.5	31.9	-4.4	27.0	31.1	-4.1	263	1271
Not well at all	11.1	33.8	-22.7***	11.4	32.8	-21.4***	106	1344
							957	3980
<b>Gore: intelligent</b>								
Extremely well	30.4	26.6	3.8	30.3	25.1	5.2†	296	1059
Quite well	56.2	33.1	23.1***	55.4	33.3	22.1***	547	1318
Not too well	10.9	29.2	-18.3***	11.4	29.5	-18.1***	106	1161
Not well at all	2.5	11.1	-8.6*	2.9	12.1	-9.2*	24	442
							973	3980
<b>Gore: strong leader</b>								
Extremely well	16.5	14.7	1.5	16.3	15.1	1.2	158	586
Quite well	43.5	23.0	20.5***	43.7	23.4	20.3***	418	915
Not too well	30.8	32.5	-1.7	30.3	31.7	-1.4	296	1295
Not well at all	9.2	29.8	-20.6***	9.3	29.9	-20.6***	88	1184
							960	3980
<b>Bush: moral</b>								
Extremely well	19.4	27.8	-8.4*	19.3	27.8	-8.5**	176	1107
Quite well	53.0	32.2	20.8***	52.3	30.3	22.0***	482	1280
Not too well	22.4	28.8	-6.4*	22.6	30.5	-7.9*	204	1148
Not well at all	5.2	11.2	-6.0†	5.9	11.4	-5.5	47	445
							909	3980
<b>Bush: cares about people</b>								
Extremely well	8.0	16.2	-8.2*	9.2	16.7	-7.5*	76	643
Quite well	37.5	23.5	14.0***	36.7	21.7	15.0***	358	934
Not too well	35.5	29.8	5.7†	35.2	30.2	5.0†	339	1185
Not well at all	19.1	30.6	-11.5***	18.9	31.4	-12.5***	183	1218
							956	3980
<b>Bush: intelligent</b>								
Extremely well	18.8	14.3	4.5	19.8	15.0	4.8	181	570
Quite well	56.8	35.2	21.6***	57.2	35.3	21.9***	546	1399
Not too well	17.7	34.9	-17.2***	16.8	34.8	-18.0***	170	1388
Not well at all	6.7	15.7	-9.0**	6.2	14.8	-8.6**	64	623
							961	3980

Continued

Table 4 (continued)

	Unweighted data (%)			Weighted data (%)			Unweighted Ns	
	ANES	Harris	Difference	ANES	Harris	Difference	ANES	Harris
Bush: strong leader								
Extremely well	14.7	23.3	-8.6**	16.0	22.3	-6.3†	138	926
Quite well	51.0	29.8	21.2***	50.0	29.2	20.8***	480	1185
Not too well	25.8	29.2	-3.4	25.2	29.4	-4.2	234	1163
Not well at all	8.6	17.7	-9.1**	8.8	19.2	-10.4**	81	706
							942	3980
Feeling thermometers								
Clinton	42.3	60.4	-18.1***	41.4	59.2	-17.8***	1002	3977
Gore	40.0	54.2	-14.2***	39.8	54.3	-14.5***	992	3976
Bush	56.0	54.5	1.5	56.7	53.3	3.4**	989	3975

Note. Actual vote choice and actual turnout were measured postelection. The figures presented for the feeling thermometers are means, not proportions.

\*\*\* $p < .001$ ; \*\* $p < .01$ ; \* $p < .05$ ; † $p < .10$  (two-tailed).

in the 2000 unweighted data, moving from the 25th to the 75th percentile of the “Gun Control” variable resulted in a 48.5 percentage point increase in the probability of voting for President Bush in the Harris data but only a 28.0 percentage point increase in the ANES data, a difference of 20.5 percentage points. For most of the variables, particularly in the 2000 data, the effect sizes approached or exceeded 10 percentage points, a substantively large difference in probabilities.

Another method of assessing substantive significance is to compare the cross-tabulations of predicted vote choice with the independent variables in the two modes/sampling methods. Using this approach, we again found large and meaningful differences. For presentational purposes, Table 7 presents cross-tabulations for the independent variables that manifested the largest effect sizes in the 2004 and 2000 data: “Pocketbook Retrospective” and “Clinton Crime Retrospective,” respectively. The relations of these variables with vote choice were much stronger in the Internet data than in the ANES data. For instance, in the 2004 YouGov sample (unweighted), the percentage of respondents who voted for President Bush was 71.4 percentage points higher among people whose personal finances were “a lot better off” than among people whose personal finances were “a lot worse off.” In the ANES data, this difference was only 44.4 percentage points.

We also observed statistically and substantively significant mode differences in regressions predicting actual vote choice with political and demographic variables measured during the preelection survey. As shown in Table 8, associations between many substantive variables and actual vote choice were stronger in the Internet data, particularly for candidate traits and policy issues.

### 5.3 Logistic Regressions Predicting Actual Turnout

When predicting actual turnout with demographics, many of the expected associations appeared in the ANES data (see Table 9). For example, turnout was higher among more



**Table 5** Comparison of face-to-face and Internet modes in logistic regressions predicting predicted vote choice

	<i>Unweighted data</i>						<i>Weighted data</i>					
	<i>Main effect</i>	<i>House</i>	<i>House interaction</i>	<i>Constant</i>	<i>N (pseudo R<sup>2</sup>)</i>	<i>Effect size (%)</i>	<i>Main effect</i>	<i>House</i>	<i>House interaction</i>	<i>Constant</i>	<i>Pseudo R<sup>2</sup></i>	<i>Effect size (%)</i>
2004 data												
Party ID	7.60***	0.68*	-1.39**	-3.64***	11,224 (0.49)	-7.7	7.65***	0.61†	-1.31*	-3.72***	0.51	-7.1
Presidential job approval	7.47***	-0.09	1.51***	-3.87***	20,994 (0.71)	2.2	7.31***	-0.30	1.84***	-3.83***	0.72	2.9
Iraq worth it	3.58***	-0.73***	0.61*	-1.34***	8203 (0.48)	8.5	3.52***	-0.80***	0.77**	-1.37***	0.50	9.7
Sociotropic retrospective	4.94***	-0.55**	0.74*	-2.06***	9594 (0.32)	7.2	4.91***	-0.82***	1.02**	-2.07***	0.34	9.7
Sociotropic prospective	2.98***	-1.30***	2.36***	-1.68***	9111 (0.22)	13.0	2.79***	-1.59***	2.62***	-1.62***	0.22	15.2
Pocketbook retrospective	2.26***	-0.28	1.48***	-1.25***	9607 (0.15)	15.2	2.29***	-0.46*	1.64***	-1.29***	0.17	17.3
Pocketbook prospective	1.26***	-0.98***	1.82***	-0.77***	9195 (0.09)	11.0	1.19***	-1.09***	1.83***	-0.76***	0.08	11.2
Male	0.23†	0.16	-0.03	-0.13	25,033 (0.00)	-0.7	0.24	0.01	0.04	-0.16	0.00	1.0
Black	-1.99***	-0.01	-0.09	0.21†	25,033 (0.03)	-1.0	-2.00***	-0.01	-0.07	0.20†	0.06	-0.8
High school	0.49*	0.69*	-0.46	-0.44*	25,033 (0.01)	—	0.68**	1.01**	-1.03**	-0.57*	0.00	—
Some college	0.57*	0.69*	-0.59†	-0.44*	25,033 (0.01)	—	0.74**	1.01**	-1.09**	-0.57*	0.00	—
College	0.50*	0.69*	-0.70*	-0.44*	25,033 (0.01)	—	0.59*	1.01**	-1.14**	-0.57*	0.00	—
Graduate	0.04	0.69*	-0.59*	-0.44*	25,033 (0.01)	—	0.17	1.01**	-0.81*	-0.57*	0.00	—
Age												
30–39	0.64**	0.21	-0.17	-0.47**	25,033 (0.01)	—	0.82***	0.20	-0.32	-0.58***	0.01	—
40–49	0.46*	0.21	-0.02	-0.47**	25,033 (0.01)	—	0.58**	0.20	-0.14	-0.58***	0.01	—
50–64	0.63**	0.21	-0.19	-0.47**	25,033 (0.01)	—	0.77**	0.20	-0.26	-0.58***	0.01	—
65+	0.45*	0.21	-0.06	-0.47**	25,033 (0.01)	—	0.51*	0.20	-0.06	-0.58***	0.01	—
2000 data												
Party ID	6.29***	-0.30	0.72	-3.05***	3984 (0.56)	4.10%	6.09***	-0.64*	1.26*	-2.86***	0.57	7.4%
Gun control	2.34***	-0.09	1.92***	-0.83***	4133 (0.17)	20.5	2.00***	-0.12	1.43**	-0.75***	0.12	16.3
Clinton economy retrospective	4.76***	0.49*	2.38**	-1.74***	3702 (0.28)	18.3	4.60***	0.42	1.57†	-1.68***	0.24	14.7
Clinton crime retrospective	1.05**	-1.28***	4.57***	-0.88***	3703 (0.22)	44.1	1.21**	-1.31***	4.24***	-0.97***	0.21	42.9
Gore: moral	4.87***	-0.94***	1.70***	-1.93***	4096 (0.40)	11.9	4.67***	-1.06***	1.83***	-1.88***	0.39	13.1

*Continued*

**Table 5** (continued)

	<i>Unweighted data</i>						<i>Weighted data</i>					
	<i>Main effect</i>	<i>House</i>	<i>House interaction</i>	<i>Constant</i>	<i>N (pseudo R<sup>2</sup>)</i>	<i>Effect size (%)</i>	<i>Main effect</i>	<i>House</i>	<i>House interaction</i>	<i>Constant</i>	<i>Pseudo R<sup>2</sup></i>	<i>Effect size (%)</i>
Gore: care about people	5.63***	-1.38***	1.70***	-2.73***	4102 (0.46)	15.4	5.53***	-1.80***	2.22***	-2.65***	0.47	19.3
Gore: intelligent	3.17***	-0.64***	2.51***	-1.14***	4114 (0.28)	25.3	2.96***	-0.79***	2.36***	-1.08***	0.26	25.3
Gore: strong leader	6.01***	-1.60***	2.41***	-2.91***	4109 (0.50)	15.8	5.76***	-2.22***	3.24***	-2.78***	0.52	20.8
Bush: moral	4.16***	-0.63†	1.99***	-2.87***	4070 (0.33)	14.4	4.19***	-0.91*	2.15**	-2.86***	0.35	16.7
Bush: care about people	4.86***	-0.26	2.99***	-2.57***	4106 (0.46)	27.1	4.44***	-0.76**	3.80***	-2.40***	0.48	31.7
Bush: intelligent	2.98***	-0.39†	2.66***	-2.15***	4108 (0.27)	21.1	3.07***	-0.42	2.35***	-2.22***	0.25	19.9
Bush: strong leader	5.39***	-0.67*	2.91***	-3.49***	4091 (0.47)	22.9	5.33***	-1.02*	3.35***	-3.46***	0.49	25.1
Clinton thermometer	5.92***	-1.37***	1.57***	-2.69***	4135 (0.51)	13.5	5.29***	-1.41***	1.49*	-2.39***	0.47	14.8
Gore thermometer	10.52***	-1.61***	2.76**	-4.39***	4130 (0.67)	6.6	9.52***	-2.15***	3.39*	-3.98***	0.66	9.4
Bush thermometer	9.48***	-1.44*	3.85***	-5.92***	4125 (0.64)	9.8	9.29***	-1.96**	4.49***	-5.86***	0.65	11.5
Male	0.45***	0.29**	0.44**	-0.44***	4138 (0.04)	10.3	0.36**	0.09	0.56**	-0.39***	0.03	13.6
Black	-1.55***	0.37***	-0.76*	-0.07	4112 (0.03)	-14.0	-1.68***	0.35**	-0.60	-0.05	0.06	-11.7
High school	0.61*	1.20**	-0.77†	-0.91***	4125 (0.01)	—	0.60†	1.11**	-0.70	-0.86**	0.01	—
Some college	0.98***	1.20**	-0.92*	-0.91***	4125 (0.01)	—	0.98**	1.11**	-1.11**	-0.86**	0.01	—
College	0.67*	1.20**	-0.68	-0.91***	4125 (0.01)	—	0.68†	1.11**	-0.78†	-0.86**	0.01	—
Graduate	0.81**	1.20**	-1.15**	-0.91***	4125 (0.01)	—	0.76*	1.11**	-1.14*	-0.86**	0.01	—
Age												
25–34	0.42	0.61*	-0.35	-0.56*	4138 (0.01)	—	0.49†	0.63*	-0.64†	-0.57*	0.01	—
35–44	0.46†	0.61*	-0.23	-0.56*	4138 (0.01)	—	0.48†	0.63*	-0.34	-0.57*	0.01	—
45–54	0.37	0.61*	-0.16	-0.56*	4138 (0.01)	—	0.45	0.63*	-0.55	-0.57*	0.01	—
55–64	0.15	0.61*	0.27	-0.56*	4138 (0.01)	—	0.20	0.63*	0.23	-0.57*	0.01	—
65+	0.31	0.61*	0.01	-0.56*	4138 (0.01)	—	0.27	0.63*	0.00	-0.57*	0.01	—

*Note.* The dependent variable was coded 1 if the respondent intended to vote for Bush and 0 if the respondent intended to vote for Kerry or Gore. All variables were coded to range from 0 to 1. House = 1 for YouGov/Harris and 0 for ANES.

\*\*\* $p < .001$ ; \*\* $p < .01$ ; \* $p < .05$ ; † $p < .10$  (two-tailed).

**Table 6** Analyzing the effect of political interest in comparing face-to-face and Internet modes in logistic regressions predicting 2004 predicted vote choice and actual turnout

	<i>Unweighted data</i>							<i>Weighted data</i>						
	<i>Main effect</i>	<i>House</i>	<i>Interest</i>	<i>House interaction</i>	<i>Interest interaction</i>	<i>Constant</i>	<i>N (pseudo R<sup>2</sup>)</i>	<i>Main effect</i>	<i>House</i>	<i>Interest</i>	<i>House interaction</i>	<i>Interest interaction</i>	<i>Constant</i>	<i>Pseudo R<sup>2</sup></i>
Dependent variable: vote														
Party ID	6.29***	0.82**	-1.74***	-1.59**	2.28***	-2.65***	11,216 (0.50)	6.54***	0.79*	-1.57***	-1.56**	2.00***	-2.86***	0.51
Presidential														
job approval	6.40***	-0.08	-0.77***	1.49***	1.76***	-3.39***	20,982 (0.71)	6.11***	-0.25	-0.89***	1.75***	2.02***	-3.29***	0.72
Iraq worth it	2.29***	-0.63***	-1.32***	0.50†	2.13***	-0.56***	8201 (0.50)	2.31***	-0.66***	-1.21***	0.60*	2.04***	-0.68***	0.51
Sociotropic														
retrospective	3.54***	-0.46**	-1.82***	0.58†	2.39***	-0.98***	9589 (0.34)	3.34***	-0.66**	-2.03***	0.75†	2.80***	-0.91***	0.35
Sociotropic														
prospective	0.80*	-1.24***	-2.73***	2.27***	3.53***	0.00	9107 (0.23)	0.49	-1.49***	-2.89***	2.48***	3.76***	0.14	0.24
Pocketbook														
retrospective	0.76*	-0.22	-1.73***	1.37***	2.47***	-0.19	9602 (0.17)	0.85*	-0.36†	-1.66***	1.46***	2.43***	-0.29	0.18
Pocketbook														
prospective	-0.17	-0.94***	-1.92***	1.75***	2.28***	0.44†	9190 (0.10)	-0.03	-1.00***	-1.65***	1.71***	1.95***	0.27	0.09
Male	-0.01	0.19	-0.61***	-0.05	0.43***	0.24†	25,006 (0.01)	0.00	0.04	-0.61***	0.02	0.42***	0.21†	0.01
Black	-2.03***	0.00	-0.39***	-0.06	0.02	0.47***	25,006 (0.03)	-1.74***	0.01	-0.32***	-0.01	-0.47	0.41***	0.06
High school	-0.35	0.98**	-1.50**	-0.75*	1.57***	0.36	25,006 (0.01)	0.06	1.17***	-0.81†	-1.19**	1.11*	-0.13	0.01
Some college	-0.07	0.98**	-1.50**	-0.86*	1.23**	0.36	25,006 (0.01)	0.43	1.17***	-0.81†	-1.24**	0.6	-0.13	0.01
College	0.09	0.98**	-1.50**	-0.97**	0.93*	0.36	25,006 (0.01)	0.60†	1.17***	-0.81†	-1.27***	0.14	-0.13	0.01
Graduate	-0.33	0.98**	-1.50**	-0.89**	0.96*	0.36	25,006 (0.01)	0.20	1.17***	-0.81†	-0.97**	0.23	-0.13	0.01
Age: 30–39	0.16	0.32*	-1.17***	-0.26	0.89***	0.17	25,006 (0.01)	0.39†	0.31†	-1.07***	-0.42†	0.80***	-0.01	0.01
Age: 40–49	0.04	0.32*	-1.17***	-0.13	0.85***	0.17	25,006 (0.01)	0.21	0.31†	-1.07***	-0.25	0.77***	-0.01	0.01
Age: 50–64	0.24	0.32*	-1.17***	-0.29	0.82***	0.17	25,006 (0.01)	0.48†	0.31†	-1.07***	-0.36	0.65***	-0.01	0.01
Age: 65+	-0.44†	0.32*	-1.17***	-0.21	1.53***	0.17	25,006 (0.01)	-0.29	0.31†	-1.07***	-0.21	1.38***	-0.01	0.01

*Continued*

**Table 6** (continued)

		Unweighted data						Weighted data							
		Main effect	House	Interest	House interaction	Interest interaction	Constant	N (pseudo R <sup>2</sup> )	Main effect	House	Interest	House interaction	Interest interaction	Constant	Pseudo R <sup>2</sup>
Dependent variable: turnout															
	Male	0.01	1.13***	1.87***	0.26	-0.39	0.58**	4157 (0.09)	0.43	1.32***	2.08***	0.17	-0.91*	0.29	0.10
	Black	-0.26	1.10***	1.77***	1.08*	-0.88	0.66***	4157 (0.09)	-0.88†	1.21***	1.65***	1.27*	0.09	0.61**	0.10
	High school	0.28	1.71***	1.07†	-0.36	0.96	-0.35	4157 (0.12)	0.10	1.97***	0.38	-0.67	1.64*	-0.19	0.14
	Some college	1.15†	1.71***	1.07†	-0.77	0.76	-0.35	4157 (0.12)	1.16†	1.97***	0.38	-1.31*	1.51†	-0.19	0.14
	College	1.92**	1.71***	1.07†	-0.77	-0.09	-0.35	4157 (0.12)	1.77**	1.97***	0.38	-0.92†	0.55	-0.19	0.14
	Graduate	3.79**	1.71***	1.07†	-2.54*	-0.06	-0.35	4157 (0.12)	3.36*	1.97***	0.38	-2.39*	0.77	-0.19	0.14
	Age														
	30-39	1.02*	1.20***	1.87***	-0.28	-0.46	0.03	4157 (0.10)	1.12*	1.53***	1.70**	-0.46	-0.43	-0.11	0.11
	40-49	0.49	1.20***	1.87***	0.19	-0.32	0.03	4157 (0.10)	0.44	1.53***	1.70**	-0.20	0.05	-0.11	0.11
	50-64	0.93†	1.20***	1.87***	-0.08	-0.16	0.03	4157 (0.10)	1.23*	1.53***	1.70**	-0.41	-0.22	-0.11	0.11
	65+	0.23	1.20***	1.87***	0.56	0.17	0.03	4157 (0.10)	0.47	1.53***	1.70**	0.67	-0.02	-0.11	0.11

*Note.* For the regressions predicting predicted vote choice, the dependent variable was coded 1 if the respondent intended to vote for President Bush and 0 if the respondent intended to vote for Senator Kerry. For the regressions predicting turnout, the dependent variable was coded 1 if respondent voted and 0 if respondent did not vote. All variables were coded to range from 0 to 1. House = 1 for YouGov/Harris and 0 for ANES.

\*\*\* $p < .001$ ; \*\* $p < .01$ ; \* $p < .05$ ; † $p < .10$  (two-tailed).

**Table 7** Cross-tabulations between predicted vote choice and independent variables

	<i>Unweighted data</i>				<i>Weighted data</i>			
	<i>ANES</i>		<i>YouGov/Harris</i>		<i>ANES</i>		<i>YouGov/Harris</i>	
	<i>Percent voting for</i>		<i>Percent voting for</i>		<i>Percent voting for</i>		<i>Percent voting for</i>	
	<i>Bush</i>	<i>Kerry/ Gore</i>	<i>Bush</i>	<i>Kerry/ Gore</i>	<i>Bush</i>	<i>Kerry/ Gore</i>	<i>Bush</i>	<i>Kerry/ Gore</i>
2004 data: pocketbook retrospective								
A lot better off	73.9	26.1	93.2	6.8	73.2	26.8	92.8	7.2
A little better off	62.2	37.8	79.9	20.1	60.8	39.2	78.0	22.0
About the same	45.6	54.4	56.0	44.0	47.5	52.5	53.7	46.3
A little worse off	30.3	69.7	33.6	66.4	30.0	70.0	29.0	71.0
A lot worse off	29.5	70.5	21.8	78.2	25.4	74.7	18.6	81.4
2000 data: Clinton crime retrospective								
Much worse	48.7	51.3	95.8	4.2	49.6	50.4	94.1	5.9
Somewhat worse	60.9	39.1	86.3	13.7	64.3	35.7	83.6	16.4
About the same	51.8	48.2	68.2	31.8	48.8	51.2	63.9	36.1
Somewhat better	34.6	65.4	31.7	68.3	35.7	64.4	27.8	72.2
Much better	33.3	66.7	7.7	92.3	34.3	65.7	6.6	93.4

educated and older respondents and among whites (e.g., Wolfinger and Rosenstone 1980; Verba, Schlozman, and Brady 1995). No significant mode/sampling method differences appeared in the 2000 data, but five of the 20 tests conducted with the 2004 data revealed significant mode/sampling method differences. And these mode/sampling method differences remained significant when controlling for differences between the modes/sampling methods in political interest (see the bottom panel of Table 6).

## 6 Discussion

In this article, we have explored how data collection mode and sample composition affect the statistical and substantive inferences researchers might make about voting behavior. Comparing the 2000 and 2004 ANES surveys to two parallel Internet surveys, we found evidence of considerable differences. Although no mode/sampling effects appeared in some analyses, such effects were substantial in others and modest but reliable in still others. This was the case in distributions of responses and in logistic regressions predicting both vote choice and turnout, using unweighted and weighted data. These mode/sampling effects were not simply the result of the fact that the Internet samples were more interested in politics on average. Because true population regression parameters are unknown, we cannot speculate as to which mode/sampling method yielded the more valid parameter estimates. We can therefore simply conclude that researchers may reach different conclusions about the distributions of variables and associations between them and about changes in these phenomena over time, depending upon which mode/sampling method is used to collect the data.

**Table 8** Comparison of face-to-face and Internet modes in logistic regressions predicting 2000 actual vote choice

2000 data	Unweighted data						Weighted data					
	Main effect	House	House interaction	Constant	N (pseudo R <sup>2</sup> )	Effect size (%)	Main effect	House	House interaction	Constant	Pseudo R <sup>2</sup>	Effect size (%)
Party ID	6.72***	0.18	-0.02	-3.34***	2132 (0.55)	-0.4	6.57***	0.10	-0.18	-3.11***	0.53	-1.2
Gun control	2.59***	-0.10	1.78**	-0.84***	2213 (0.18)	18.5	2.19***	-0.27	1.62*	-0.71**	0.14	17.7
Clinton economy retrospective	5.85***	0.77*	1.33	-2.01***	1972 (0.29)	9.3	5.88***	0.72*	0.22	-1.96***	0.25	3.1
Clinton crime retrospective	1.15*	-1.35***	4.36***	-0.79**	1969 (0.21)	42.0	1.47**	-1.26**	3.66***	-0.97**	0.20	37.5
Gore: moral	4.38***	-0.90***	1.48*	-1.65***	2201 (0.36)	11.2	4.43***	-0.82	1.02	-1.65***	0.33	9.0
Gore: care about people	5.04***	-1.45***	1.76*	-2.36***	2199 (0.42)	17.7	4.97***	-1.75***	2.06*	-2.29***	0.43	20.8
Gore: intelligent	2.74***	-0.65***	2.35***	-0.94***	2210 (0.25)	26.2	2.86***	-0.65**	1.83**	-0.99***	0.23	21.0
Gore: strong leader	6.18***	-1.09***	1.59*	-3.07***	2202 (0.47)	8.9	6.06***	-1.27**	1.56†	-3.01***	0.46	7.5
Bush: moral	4.32***	-0.16	1.37*	-2.99***	2183 (0.31)	8.9	4.18***	-0.42	1.40†	-2.81***	0.31	10.9
Bush: care about people	5.53***	0.36	1.47*	-2.79***	2201 (0.43)	16.9	5.29***	0.03	1.70*	-2.68***	0.43	18.0
Bush: intelligent	3.74***	0.13	1.81**	-2.57***	2199 (0.27)	15.2	3.85***	0.20	1.28	-2.65***	0.24	12.6
Bush: strong leader	5.52***	-0.15	1.81**	-3.39***	2192 (0.43)	14.7	5.79***	0.00	1.24	-3.52***	0.42	11.3
Clinton thermometer	6.78***	-0.77*	0.52	-3.20***	2212 (0.52)	6.2	7.12***	-0.53	-0.29	-3.28***	0.51	3.3
Gore thermometer	8.48***	-1.14**	1.64	-3.54***	2210 (0.57)	7.7	7.45***	-1.60***	2.39	-3.13***	0.57	11.8
Bush thermometer	9.13***	0.22	1.12	-5.65***	2210 (0.58)	1.4	9.53***	1.08	-0.70	-5.90***	0.54	-4.3
Male	0.58***	0.35*	0.26	-0.44**	2214 (0.03)	5.8	0.45*	-0.03	0.55*	-0.33*	0.04	13.3
Black	-2.74***	0.31*	0.24	0.01	2213 (0.04)	-3.7	-3.24***	0.18	0.47	0.09	0.07	-1.1
High school	0.39	0.79	-0.53	-0.55	2214 (0.01)	—	0.39	0.73	-0.55	-0.49	0.00	—
Some college	0.47	0.79	-0.37	-0.55	2214 (0.01)	—	0.41	0.73	-0.60	-0.49	0.00	—
College	0.23	0.79	-0.12	-0.55	2214 (0.01)	—	0.29	0.73	-0.28	-0.49	0.00	—
Graduate	0.78†	0.79	-1.00†	-0.55	2214 (0.01)	—	0.68	0.73	-0.99	-0.49	0.00	—

*Continued*

Table 8 (continued)

2000 data	Unweighted data						Weighted data					
	Main effect	House	House interaction	Constant	N (pseudo R <sup>2</sup> )	Effect size (%)	Main effect	House	House interaction	Constant	Pseudo R <sup>2</sup>	Effect size (%)
Age												
25–34	1.16*	1.23*	–0.94†	–1.28**	2214 (0.01)	—	1.08†	0.94†	–0.82	–1.06*	0.02	—
35–44	1.12*	1.23*	–0.82	–1.28**	2214 (0.01)	—	0.95†	0.94†	–0.71	–1.06*	0.02	—
45–54	1.09*	1.23*	–0.75	–1.28**	2214 (0.01)	—	1.04*	0.94†	–1.20*	–1.06*	0.02	—
55–64	1.20*	1.23*	–0.69	–1.28**	2214 (0.01)	—	1.10*	0.94†	–0.34	–1.06*	0.02	—
65+	1.21*	1.23*	–0.72	–1.28**	2214 (0.01)	—	0.92†	0.94†	–0.49	–1.06*	0.02	—

Note. The dependent variable was coded 1 if the respondent voted for Bush and 0 if the respondent voted for Gore. All variables were coded to range from 0 to 1. House = 1 for Harris and 0 for ANES.

\*\*\* $p < .001$ ; \*\* $p < .01$ ; \* $p < .05$ ; † $p < .10$  (two-tailed).

**Table 9** Comparison of face-to-face and Internet modes in logistic regressions predicting actual turnout

		<i>Unweighted data</i>				<i>Weighted data</i>					
		<i>Main effect</i>	<i>House</i>	<i>House interaction</i>	<i>Constant</i>	<i>N</i> <i>(pseudo R<sup>2</sup>)</i>	<i>Main effect</i>	<i>House</i>	<i>House interaction</i>	<i>Constant</i>	<i>Pseudo R<sup>2</sup></i>
309	2004 data										
	Male	-0.02	1.28***	0.19	1.54***	4159 (0.04)	0.12	1.49***	0.11	1.33***	0.06
	Black	-0.76**	1.20***	1.06*	1.65***	4159 (0.04)	-0.87**	1.35***	1.29*	1.53***	0.06
	High school	0.83*	1.82***	-0.46	0.17	4159 (0.08)	0.99**	1.99***	-0.71	0.00	0.10
	Some college	1.60**	1.82***	-0.75	0.17	4159 (0.08)	1.97***	1.99***	-1.17†	0.00	0.10
	College	1.97***	1.82***	-0.78	0.17	4159 (0.08)	2.11***	1.99***	-0.83	0.00	0.10
	Graduate	4.03***	1.82***	-2.63*	0.17	4159 (0.08)	4.03***	1.99***	-2.38*	0.00	0.10
	Age										
	30-39	0.83*	1.41***	-0.38	0.95***	4159 (0.05)	0.97*	1.75***	-0.59	0.72**	0.08
	40-49	0.49	1.41***	0.07	0.95***	4159 (0.05)	0.59	1.75***	-0.26	0.72**	0.08
	50-64	1.01**	1.41***	-0.20	0.95***	4159 (0.05)	1.29***	1.75***	-0.57	0.72**	0.08
	65+	0.58	1.41***	0.50	0.95***	4159 (0.05)	0.68	1.75***	0.61	0.72**	0.08
	2000 data										
	Male	0.26	1.24***	0.27	0.90	2869 (0.06)	0.18	1.17**	0.12	0.71***	0.05
	Black	-0.19	1.35***	-0.46	1.04	2857 (0.06)	-0.13	1.22***	-0.14	0.81***	0.05
	High school	0.58*	1.21**	-0.02	-0.18	2861 (0.12)	0.78**	1.53**	-0.30	-0.37	0.11
	Some college	1.40***	1.21**	-0.07	-0.18	2861 (0.12)	1.51***	1.53**	-0.30	-0.37	0.11
	College	1.98***	1.21**	-0.40	-0.18	2861 (0.12)	2.18***	1.53**	-0.60	-0.37	0.11
	Graduate	2.76***	1.21**	-0.22	-0.18	2861 (0.12)	2.92***	1.53**	-0.43	-0.37	0.11
	Age										
25-34	0.68†	1.70***	-0.49	-0.21	2869 (0.11)	0.59†	1.23**	-0.12	-0.33	0.13	
35-44	1.21***	1.70***	-0.15	-0.21	2869 (0.11)	1.21***	1.23**	0.16	-0.33	0.13	
45-54	1.67**	1.70***	-0.41	-0.21	2869 (0.11)	1.61***	1.23**	-0.23	-0.33	0.13	
55-64	1.72***	1.70***	-0.21	-0.21	2869 (0.11)	1.84***	1.23**	-0.02	-0.33	0.13	
65+	1.78***	1.70***	1.95†	-0.21	2869 (0.11)	1.77***	1.23**	1.51	-0.33	0.13	

*Note.* The dependent variable was coded 1 if the respondent voted and 0 if the respondent did not vote. All variables were coded to range from 0 to 1. House = 1 for YouGov/Harris and 0 for ANES.

\*\*\* $p < .001$ ; \*\* $p < .01$ ; \* $p < .05$ ; † $p < .10$  (two-tailed).



**Table 10** Marginal distributions for actual vote choice and turnout compared to election outcomes

		<i>Unweighted (%)</i>		<i>Weighted (%)</i>		<i>Unweighted (%)</i>		<i>Weighted (%)</i>		<i>Unweighted Ns</i>		
		<i>ANES</i>	<i>YouGov/Harris</i>	<i>ANES</i>	<i>YouGov/Harris</i>	<i>Election outcome (%)</i>	<i>ANES – outcome</i>	<i>YouGov/Harris – outcome</i>	<i>ANES – outcome</i>	<i>YouGov/Harris – outcome</i>	<i>ANES</i>	<i>YouGov</i>
310	2004 data											
	Actual vote choice											
	Bush	49.8	52.4	49.6	49.5	51.2	-1.4	1.2	-1.6	-1.7	354	1746
	Kerry	50.2	47.6	50.4	50.5	48.8	1.4	-1.2	1.6	1.7	357	1587
											711	3333
	Actual turnout											
	Voted	82.1	94.6	79.9	94.9	60.3	21.8***	34.3***	19.6***	34.6***	441	3428
	Did not vote	17.9	5.4	20.1	5.1	39.7	-21.8***	-34.3***	-19.6***	-34.6***	96	194
											537	3622
	2000 data											
	Actual vote choice											
	Bush	45.7	55.7	46.6	53.1	49.7	-4.0	6.0***	-3.1	3.4**	223	1527
Gore	54.3	44.3	53.4	46.9	50.3	4.0	-6.0***	3.1	-3.4*	265	1215	
										488	2742	
Actual turnout												
Voted	73.6	91.3	69.4	88.1	54.2	19.4***	37.1***	15.2***	33.9***	634	1986	
Did not vote	26.4	8.7	30.6	11.9	45.8	-19.4***	-37.1***	-15.2***	-33.9***	228	190	

*Note.* In the “Election Outcome” column, “Actual Vote Choice” is the two-party vote share, and “Actual Turnout” is the percentage of the voting-eligible population that voted. \*\*\**p* < .001; \*\**p* < .01; \**p* < .05; †*p* < .10 (two-tailed).

Two political variables offered opportunities to compare the survey results with the “truth,” involving measures of two-party vote choice<sup>10</sup> and turnout.<sup>11</sup> As shown in Table 10, for the case of actual turnout, the face-to-face ANES data were more accurate than the Internet data for both the weighted and unweighted data. The ANES data produced turnout rates between 15 and 22 percentage points greater than actual turnout, but the Internet data produced turnout reports over 34 percentage points greater. For candidate choice, the ANES data more closely matched the 2000 two-party vote result than did the Harris data (unweighted). Whereas the ANES showed that Gore received 4 percentage points more than his actual two-party vote share, the Harris survey showed Bush receiving an excess of 6 percentage points. In all other cases, the modes/sampling methods were similarly accurate. Furthermore, the demographic characteristics of the face-to-face samples matched those of the population more closely than did those of the Internet samples. Of the 16 comparisons between variables shown in Tables 1 and 10, the face-to-face data were more accurate in 14 instances, or 88%. This superior accuracy of the face-to-face data suggests that perhaps the associations between variables in those data may have been more accurate as well. But we have no direct evidence of this because we have no correlational benchmarks against which to compare.

It is important to point out that the mode/sampling comparisons reported here are not optimally designed. It would have been preferable to collect the Internet data during exactly the same time period as the face-to-face data, to eliminate history confounds with mode/sampling. And it would also have been preferable to collect the Internet data using exactly the same questionnaires as were used in the face-to-face interviews. However, even in instances where exactly the same question was asked in the two modes/sampling methods and the response categories were identical, we found mode/sampling differences. For example, party identification and presidential approval were measured identically in the two 2004 surveys, but we found significant mode/sampling differences in the equations predicting vote choice (see Table 5). Similarly, party identification was measured identically in the two 2000 surveys, and the mode/sampling interaction was statistically significant for the weighted data. This suggests that the observed mode/sampling differences are not due to questionnaire differences alone, though we cannot be sure of this with this design.

We hope that future research will test for mode/sampling effects with identical questionnaires focused on politics. This could be accomplished by coordinating administration of the ANES with simultaneous Internet administration of the same questionnaire. Future research can also attempt to explain the differences observed between the modes/sampling methods that we saw here. We found evidence suggesting that these differences cannot be attributed to variation in the political interest of the samples. So the search for the mediators at work seems worth continuing in other directions.

<sup>10</sup>When these analyses were conducted comparing the truth with each candidate’s share of all votes (rather than just of the two-party vote) and the vote margin, comparable results were obtained.

<sup>11</sup>Our measure of “true” turnout was taken from [http://elections.gmu.edu/voter\\_turnout.htm](http://elections.gmu.edu/voter_turnout.htm), where the numerator is the number of voters and the denominator is the voting-eligible population, which excludes noncitizens and ineligible felons and includes eligible citizens living overseas or living in institutional settings in the United States (e.g., college dorms or group homes). This causes some unavoidable noncomparability with the surveys. For example, the Internet samples could have included noncitizens and convicted felons who were not in jail; the face-to-face samples could also have included felons but did not include American citizens who were overseas or who were living in institutional settings in the United States; and none of the survey samples included people who did not speak/read English. Therefore, we should not expect exact matches of the survey samples with true turnout rates.

Another interesting path for future research would be to determine whether the differences documented here are attributable to mode (face-to-face interviewing versus Internet self-completion) or to sampling method (area probability sample versus volunteers) or both. Some past studies suggest that reporting precision varies depending upon mode of data collection, with computer self-completion acquiring more accurate self-descriptions than telephone interviewing does (e.g., Chang and Krosnick 2002). Similar experiments could be conducted comparing face-to-face interviewing with computer self-completion to see what difference this mode change makes and whether it might be partly responsible for the results we report here. Likewise, studies could be conducted in which mode of data collection is held constant (e.g., computer self-completion only) and sampling method (area probability versus nonprobability volunteer) is varied, to isolate the effect of the latter. There are considerable practical challenges in doing this sort of study, but it may be possible and would certainly be informative.

## **7 The Value of Volunteer Samples**

Is there value in data collected from volunteer, nonprobability general public samples via the Internet? Absolutely. A great deal of valuable social science theory building has been done with nonprobability samples, especially in psychology and sociology (see, e.g., Sears 1986). But the goal of that work has almost always been to conduct experiments and to document the impact of experimental manipulations. Thus, compromises in external validity have been made regularly in the service of maximizing internal validity of causal inference. Much can be done to test hypotheses and move social science forward in this way and that should certainly continue. But such results should be generalized to the general public with confidence only after those findings have been replicated with representative general public samples.

## **8 Conclusion**

The Internet offers wonderfully attractive potential for collecting survey data quickly and inexpensively. And the cost can be reduced even more if volunteer samples are used instead of probability samples. But saving money in these ways when conducting election studies seems most worthwhile if substantive conclusions about distributions of variables, associations between variables, and trends over time would not be distorted, not only regarding voters but also for the population as a whole (as is required by academic scholars studying the predictors of turnout).

The evidence reported here suggests that results may differ considerably depending upon the mode/sampling method employed. So shifts in mode/sampling method, and the use of volunteer Internet samples in particular, should be done cautiously, and researchers should be sure to acknowledge clearly and unambiguously in their research reports their mode of data collection and that they are analyzing nonprobability samples when this is true (in keeping with the Professional Standards of the American Association for Public Opinion Research; see [http://www.aapor.org/pdfs/AAPOR\\_Code\\_2005.pdf](http://www.aapor.org/pdfs/AAPOR_Code_2005.pdf)). In addition, researchers should be sure to clearly acknowledge any mode and sampling method differences when comparing the results of two different surveys.

## **Appendix: Items and Coding**

This appendix presents the question wordings and response options for all the measures analyzed in this paper. The coding of each response option is shown in parentheses after the response option wording. "Removed" means that a response was treated as missing.

## Appendix

2004 ANES

2004 YouGov

## Predicted Vote Choice:

Who do you think you will vote for in the election for President?/If you were going to vote, who do you think you would vote for in the election for President?

- John Kerry (0)
- George W. Bush (1)
- Other (Removed)

## Actual Vote Choice:

Who did you vote for?

- John Kerry (0)
- George W. Bush (0)
- Ralph Nader (Removed)
- Other/DK/Refused (Removed)

## Actual Turnout:

In talking to people about elections, we often find that a lot of people were not able to vote because they weren't registered, they were sick, or they just didn't have time. How about you—did you vote in the elections this November?

- Yes, voted (1)
- No, didn't vote (0)

## Interest in Campaigns:

Some people don't pay much attention to political campaigns. How about you? Would you say that you have been very much interested, somewhat interested or not much interested in the political campaigns so far this year?

- Very much interested (1)
- Somewhat interested (.5)
- Not much interested (0)

## Party ID:

Generally speaking, do you usually think of yourself as a Republican, a Democrat, an independent, or what?

- Republican
- Democrat
- Independent
- Other party
- (If Democrat) Would you call yourself a strong Democrat or a not very strong Democrat?
- Strong (0)
- Not very strong (.167)

## Predicted Vote Choice:

If you were to vote, who do you think you are most likely to vote for?

- John Kerry (0)
- George W. Bush (1)
- Other (Removed)

## Actual Vote Choice:

Who did you vote for?

- John Kerry (0)
- George W. Bush (0)
- Ralph Nader (Removed)
- Someone else/DK (Removed)

## Actual Turnout:

Many people don't vote in presidential elections either because they don't want to, because they weren't registered, they were sick, or they just didn't have time. How about you—did you vote in the elections this November?

- Yes (1)
- No (0)

## Interest in Campaigns:

Some people don't pay much attention to political campaigns. How about you? Would you say that you have been very much interested, somewhat interested or not much interested in the political campaigns so far this year?

- Very much interested (1)
- Somewhat interested (.5)
- Not much interested (0)

## Party ID:

Generally speaking, do you think of yourself as a . . .

- Republican
- Democrat
- Independent
- Something else
- (If Democrat) Would you call yourself a strong Democrat or a not very strong Democrat?
- Strong (0)
- Not very strong (.167)

Continued

## Appendix (continued)

2004 ANES

2004 YouGov

(If Republican) Would you call yourself a strong Republican or a not very strong Republican?  
 –Strong (1)  
 –Not very strong (.833)  
 (If Independent/Other) Do you think of yourself as closer to the Republican Party or to the Democratic Party?  
 –Closer to Republican (.667)  
 –Neither (.5)  
 –Closer to Democratic (.333)  
 Presidential Job Approval:  
 Do you approve or disapprove of the way George W. Bush is handling his job as president?  
 –Approve  
 –Disapprove  
 (If approve) Do you approve strongly or not strongly?  
 –Strongly (1)  
 –Not strongly (.67)  
 (If disapprove) Do you disapprove strongly or not strongly?  
 –Strongly (0)  
 –Not strongly (.33)  
 Iraq Worth It:  
 Taking everything into account, do you think the war in Iraq has been worth the cost or not?  
 –Worth it (1)  
 –Not worth it (0)  
 Sociotropic Retrospective:  
 Now thinking about the economy in the country as a whole, would you say that over the past year the nation's economy has gotten better, stayed about the same, or gotten worse?  
 –Gotten better  
 –Stayed about the same (.5)  
 –Gotten worse  
 (If better) Much better or somewhat better?  
 –Much (1)  
 –Somewhat (.75)  
 (If worse) Much worse or somewhat worse?  
 –Much (0)  
 –Somewhat (.25)

(If Republican) Would you call yourself a strong Republican or a not very strong Republican?  
 –Strong (1)  
 –Not very strong (.833)  
 (If Independent/Something else) Do you think of yourself as closer to the Republican Party or to the Democratic Party?  
 –Closer to Republican (.667)  
 –Not closer to either one (.5)  
 –Closer to Democratic (.333)  
 Presidential Job Approval:  
 Do you approve or disapprove of the way George W. Bush is handling his job as president?  
 –Approve  
 –Disapprove  
 (If approve) Do you approve strongly or just somewhat?  
 –Strongly (1)  
 –Somewhat (.67)  
 (If disapprove) Do you disapprove strongly or just somewhat?  
 –Strongly (0)  
 –Somewhat (.33)  
 Iraq Worth It:  
 Overall, do you think it was worth going to war in Iraq, or not?  
 –Worth it (1)  
 –Not worth it (0)

Sociotropic Retrospective:  
 Now thinking about the economy in the country as a whole, would you say that overall it has. . .  
 –Gotten a lot better (1)  
 –Gotten a little better (.75)  
 –Stayed about the same (.5)  
 –Gotten a little worse (.25)  
 –Gotten a lot worse (0)

Continued

## Appendix (continued)

<i>2004 ANES</i>	<i>2004 YouGov</i>
<p>Sociotropic Prospective:            What about the next 12 months? Do you expect the economy, in the country as a whole, to get better, stay about the same, or get worse?            –Get better            –Stay about the same (.5)            –Get worse            (If better) Much better or somewhat better?            –Much (1)            –Somewhat (.75)            (If worse) Much worse or somewhat worse?            –Much (0)            –Somewhat (.25)</p> <p>Pocketbook Retrospective:            We are interested in how people are getting along financially these days. Would you say that you (and your family living here) are better off or worse off than you were a year ago?            –Better            –Worse            –The same (.5)            (If better) Much better or somewhat better?            –Much (1)            –Somewhat (.75)            (If worse) Much worse or somewhat worse?            –Much (0)            –Somewhat (.25)</p> <p>Pocketbook Prospective:            Now looking ahead, do you think that a year from now you (and your family living here) will be better off financially, worse off, or just about the same as now?            –Better            –Worse            –The same (.5)            (If better) Much better or somewhat better?            –Much (1)            –Somewhat (.75)            (If worse) Much worse or somewhat worse?            –Much (0)            –Somewhat (.25)</p> <p>Gender:            Interviewer coded            –Male (1)            –Female (0)</p>	<p>Sociotropic Prospective:            What about the next 12 months? Do you expect the economy in the country as a whole to . . .            –Get a lot better (1)            –Get a little better (.75)            –Stay about the same (.5)            –Get a little worse (.25)            –Get a lot worse (0)</p> <p>Pocketbook Retrospective:            We are interested in how people are getting along financially these days. Would you say that you (and your family living here) are . . .            –A lot better off (1)            –A little better off (.75)            –About the same (.5)            –A little worse off (.25)            –A lot worse off (0)</p> <p>Pocketbook Prospective:            Now looking ahead and thinking about the next few years, do you expect your financial situation to be . . .            –A lot better (1)            –A little better (.75)            –About the same (.5)            –A little worse (.25)            –A lot worse (0)</p> <p>Gender:            Are you . . .            –Male (1)            –Female (0)</p>

*Continued*

## Appendix (continued)

2004 ANES

2004 YouGov

## Race:

What racial or ethnic group or groups best describes you?

- Black (1)
- Asian (0)
- Native American (0)
- Hispanic or Latino (0)
- White (0)
- Other (0)

## Education:

What is the highest grade of school or year of college you have completed?

- 0-12 grades
- 13-16 grades
- 17+ grades
- (If 0-12) Did you get a high school diploma or pass a high school equivalency test?
- Yes (High School = 1)
- No (Baseline Category)
- (If 13+) What is the highest degree that you have earned?
- Bachelor's Degree (College = 1)
- Master's Degree (Graduate = 1)
- PhD, LIT, SCD, DFA, DLIT, DPH, DPHIL, JSC, SJD (Graduate = 1)
- LLB, JD (Graduate = 1)
- MD, DDS, DVM, MVSA, DSC, DO (Graduate = 1)
- JDC, STD, THD (Graduate = 1)
- Associate degree (AA) (College = 1)
- No degree earned (Some College = 1)

## Age:

What is the month, day and year of your birth?

Coded into dummies representing four age groups: 30-39, 40-49, 50-64, and 65+. Respondents between 18-29 years old represent baseline category.

## Race:

What is your race?

- White (0)
- Black/African-American (1)
- American Indian or Alaskan Native (0)
- Asian/Pacific Islander (0)
- Other (0)

## Education:

What is the highest degree or level of education you have completed?

- Less than high school (Baseline Category)
- Completed some high school (Baseline Category)
- High school graduate or equivalent (High School = 1)
- Completed some college (Some College = 1)
- College graduate (College = 1)
- Completed some graduate school, no degree (College = 1)
- Completed graduate school (eg MS, MD, PhD) (Graduate = 1)

## Age:

Which year were you born in?

Coded into dummies representing four age groups: 30-39, 40-49, 50-64, and 65+. Respondents between 18-29 years old represent baseline category.

Continued

## Appendix (continued)

2000 ANES

2000 Harris

## Predicted Vote Choice:

Who do you think you will vote for in the election for President?/If you were going to vote, who do you think you would vote for in the election for President?

- Al Gore (0)
- George W. Bush (1)
- Other (Removed)

## Actual Vote Choice:

Who did you vote for?

- Al Gore (0)
- George W. Bush (1)
- Pat Buchanan (Removed)
- Ralph Nader (Removed)
- Other/DK/Refused (Removed)

## Actual Turnout:

In talking to people about elections, we often find that a lot of people were not able to vote because they weren't registered, they were sick, or they just didn't have time. Which of the following statements best describes you: One, I did not vote (in the election this November); Two, I thought about voting this time—but didn't; Three, I usually vote, but didn't this time; or Four, I am sure I voted?

- I am sure I voted (1)
- I did not vote/I thought about voting this time, but didn't/I usually vote, but didn't this time (0)

## Party ID:

Generally speaking, do you think of yourself as a Republican, a Democrat, an Independent, or what?

## Predicted Vote Choice:

Many things may happen between now and then that may influence your views of the candidates. And of course you can't know now what those things will be. But we'd like to ask for your best guess about who you will vote for in the election for President in November?/Many things may happen between now and then that may influence your views of the candidates. And of course you can't know now what those things will be. But knowing what you know today, we'd like to ask for your best guess. If you were to end up voting in the election for President in November, who would you be most likely to vote for?

- Al Gore (the Democrat) (0)
- George W. Bush (the Republican) (1)
- Other (Removed)

## Actual Vote Choice:

Who did you vote for in the election for President, Al Gore, the Democrat, George W. Bush, the Republican, or someone else?

- Al Gore (0)
- George W. Bush (1)
- Someone Else (Removed)

## Actual Turnout:

Things often happen that make it difficult for people to vote in elections, even if they want to. In the past, have you usually voted in elections, or have you usually not voted?

- Usually voted
- Usually not voted

What about in the election for President of the United States that just took place—did you vote in that election, or did you not vote this time?

- Did vote (1)
- Did not vote (0)

## Party ID:

Generally speaking, do you think of yourself as:

- Republican
- Democrat

Continued



## Appendix (continued)

<i>2000 ANES</i>	<i>2000 Harris</i>
-Republican	-Independent
-Democrat	-Other
-Independent	(If Democrat): Would you call yourself a strong Democrat or a not very strong Democrat?
-Other party	-Strong (0)
(If Democrat): Would you call yourself a strong Democrat or a not very strong Democrat?	-Not Very Strong (.167)
-Strong (0)	(If Republican): Would you call yourself a strong Republican or a not very strong Republican?
-Not Very Strong (.167)	-Strong (1)
(If Republican): Would you call yourself a strong Republican or a not very strong Republican?	-Not very strong (.833)
-Strong (1)	(If Independent or Other): Do you think of yourself as:
-Not very strong (.833)	-Closer to Republican (.667)
(If Independent, No Preference, Other): Do you think of yourself as closer to the Republican Party or to the Democratic Party?	-Closer to Democratic (.333)
-Closer to Republican (.667)	-Equally Close to Both (.5)
-Neither (.5)	
-Closer to Democratic (.333)	
Gun control:	Gun control:
Do you think the federal government should make it more difficult for people to buy a gun than it is now, make it easier for people to buy a gun, or keep these rules about the same as they are now?	Do you think the federal government should make it more difficult for people to buy a gun than it is now, make it easier for people to buy a gun, or keep these rules about the same as they are now?
-More difficult	-A lot more difficult (0)
-Make it easier	-A little more difficult (.25)
-Keep these rules about the same (.5)	-About the same (.5)
(If more difficult): A lot more difficult or somewhat more difficult?	-A little easier (.75)
-A lot more difficult (0)	-A lot easier (1)
-Somewhat more difficult (.25)	
(If easier): A lot easier or somewhat easier?	
-A lot easier (1)	
-Somewhat easier (.75)	
Clinton Economy Retrospective:	Clinton Economy Retrospective:
Since 1992, would you say President Clinton has made the nation's economy better, made the economy worse, or had no effect on the economy one way or the other?	Next, we'd like to ask whether you believe some things in the country now are different from how they were when Bill Clinton became President in 1993, or whether each of these things is pretty much the same now as it was then. Compared to eight years ago, would you say that [the nation's economy] is now...?
-Made the economy better	-Much better (0)
-Made the economy worse	-Somewhat better (.25)
-No effect (.5)	
(If better): Much better or somewhat better?	

Continued

## Appendix (continued)

<i>2000 ANES</i>	<i>2000 Harris</i>
-Much better (0)	-About the same (.5)
-Somewhat better (.25)	-Somewhat worse (.75)
(If worse): Much worse or somewhat worse?	-Much worse (1)
-Much worse (1)	
-Somewhat worse (.75)	
Clinton Crime Retrospective: Would you say that the Clinton administration has made the nation's crime rate better, worse, or hasn't it made much difference either way?	Clinton Crime Retrospective: Next, we'd like to ask whether you believe some things in the country now are different from how they were when Bill Clinton became President in 1993, or whether each of these things is pretty much the same now as it was then. Compared to eight years ago, would you say that [the amount of crime] is now...?
-Better	-Much better (0)
-Worse	-Somewhat better (.25)
-No difference (.5)	-About the same (.5)
(If better): Much better or somewhat better?	-Somewhat worse (.75)
-Much better (0)	-Much worse (1)
-Somewhat better (.25)	
(If worse): Much worse or somewhat worse?	
-Much worse (1)	
-Somewhat worse (.75)	
Gore: Moral: Think about Al Gore. In your opinion, does the phrase "he is moral" describe Al Gore extremely well, quite well, not too well, or not well at all?	Gore: Moral: In your opinion, how well does the phrase "moral" describe Al Gore?
-Extremely well (0)	-Extremely well (0)
-Quite well (.33)	-Very well (.33)
-Not too well (.67)	-Somewhat (.67)
-Not well at all (1)	-Not at all (1)
Gore: Care About People: What about really cares about people like you? Does this phrase describe Gore extremely well, quite well, not too well, or not well at all?	Gore: Care About People: In your opinion, how well does the phrase "really cares about people like you" describe Al Gore?
-Extremely well (0)	-Extremely well (0)
-Quite well (.33)	-Very well (.33)
-Not too well (.67)	-Somewhat (.67)
-Not well at all (1)	-Not at all (1)
Gore: Intelligent: What about intelligent? Does this phrase describe Gore extremely well, quite well, not too well, or not well at all?	Gore: Intelligent: In your opinion, how well does the phrase "intelligent" describe Al Gore?
-Extremely well (0)	-Extremely well (0)
-Quite well (.33)	-Very well (.33)
-Not too well (.67)	-Somewhat (.67)
-Not well at all (1)	-Not at all (1)

Continued

## Appendix (continued)

2000 ANES

2000 Harris

## Gore: Strong Leader:

What about provides strong leadership?

Does this phrase describe Gore extremely well, quite well, not too well, or not well at all?

-Extremely well (0)

-Quite well (.33)

-Not too well (.67)

-Not well at all (1)

## Bush: Moral:

Think about George W. Bush. In your opinion, does the phrase "he is moral" describe George W. Bush extremely well, quite well, not too well, or not well at all?

-Extremely well (1)

-Quite well (.67)

-Not too well (.33)

-Not well at all (0)

## Bush: Care About People:

What about really cares about people like you? Does this phrase describe George W. Bush extremely well, quite well, not too well, or not well at all?

-Extremely well (1)

-Quite well (.67)

-Not too well (.33)

-Not well at all (0)

## Bush: Intelligent:

What about intelligent? Does this phrase describe George W. Bush extremely well, quite well, not too well, or not well at all?

-Extremely well (1)

-Quite well (.67)

-Not too well (.33)

-Not well at all (0)

## Bush: Strong Leader:

What about provides strong leadership? Does this phrase describe George W. Bush extremely well, quite well, not too well, or not well at all?

-Extremely well (1)

-Quite well (.67)

-Not too well (.33)

-Not well at all (0)

## Gore: Strong Leader:

In your opinion, how well does the phrase "can provide strong leadership" describe Al Gore?

-Extremely well (0)

-Very well (.33)

-Somewhat (.67)

-Not at all (1)

## Bush: Moral:

In your opinion, how well does the phrase "moral" describe George W. Bush?

-Extremely well (1)

-Very well (.67)

-Somewhat (.33)

-Not at all (0)

## Bush: Care About People:

In your opinion, how well does the phrase "really cares about people like you" describe George W. Bush?

-Extremely well (1)

-Very well (.67)

-Somewhat (.33)

-Not at all (0)

## Bush: Intelligent:

In your opinion, how well does the phrase "intelligent" describe George W. Bush?

-Extremely well (1)

-Very well (.67)

-Somewhat (.33)

-Not at all (0)

## Bush: Strong Leader:

In your opinion, how well does the phrase "can provide strong leadership" describe George W. Bush?

-Extremely well (1)

-Very well (.67)

-Somewhat (.33)

-Not at all (0)

Continued

## Appendix (continued)

2000 ANES

2000 Harris

## Clinton Thermometer:

The feeling thermometer can rate people from 0 to 100 degrees. Ratings between 50 degrees and 100 degrees mean that you feel favorable and warm toward the person. Ratings between 0 degrees and 50 degrees mean that you don't feel favorable toward the person. Rating the person at the midpoint, the 50 degree mark, means you don't feel particularly warm or cold toward the person.

The first person is Bill Clinton. Where on the feeling thermometer would you rate Bill Clinton?

Coded to lie between 0 (100 thermometer score) to 1 (0 thermometer score)

## Gore Thermometer:

The feeling thermometer can rate people from 0 to 100 degrees. Ratings between 50 degrees and 100 degrees mean that you feel favorable and warm toward the person. Ratings between 0 degrees and 50 degrees mean that you don't feel favorable toward the person. Rating the person at the midpoint, the 50 degree mark, means you don't feel particularly warm or cold toward the person.

The next person is Al Gore. Where on the feeling thermometer would you rate him?

Coded to lie between 0 (100 thermometer score) to 1 (0 thermometer score)

## Bush Thermometer:

The feeling thermometer can rate people from 0 to 100 degrees. Ratings between 50 degrees and 100 degrees mean that you feel favorable and warm toward the person. Ratings between 0 degrees and 50 degrees mean that you don't feel favorable toward the person. Rating the person at the midpoint, the 50 degree mark, means you don't feel particularly warm or cold toward the person.

The next person is George W. Bush. Where on the feeling thermometer would you rate him?

Coded to lie between 0 (0 thermometer score) to 1 (100 thermometer score)

## Clinton Thermometer:

We will list the names of several people.

Please rate how favorable or unfavorable you feel toward each person by picking a number between 0 and 100. The larger the number you pick, the more you like the person. Ratings between 50 and 100 mean that you feel favorable toward the person, and ratings between 0 and 50 mean that you feel unfavorable toward the person. You would rate a person at 50 if you don't feel favorable or unfavorable.

Bill Clinton

Coded to lie between 0 (100 thermometer score) to 1 (0 thermometer score)

## Gore Thermometer:

We will list the names of several people.

Please rate how favorable or unfavorable you feel toward each person by picking a number between 0 and 100. The larger the number you pick, the more you like the person. Ratings between 50 and 100 mean that you feel favorable toward the person, and ratings between 0 and 50 mean that you feel unfavorable toward the person. You would rate a person at 50 if you don't feel favorable or unfavorable.

Al Gore.

Coded to lie between 0 (100 thermometer score) to 1 (0 thermometer score)

## Bush Thermometer:

We will list the names of several people.

Please rate how favorable or unfavorable you feel toward each person by picking a number between 0 and 100. The larger the number you pick, the more you like the person. Ratings between 50 and 100 mean that you feel favorable toward the person, and ratings between 0 and 50 mean that you feel unfavorable toward the person. You would rate a person at 50 if you don't feel favorable or unfavorable.

George W. Bush

Coded to lie between 0 (0 thermometer score) to 1 (100 thermometer score)

Continued

## Appendix (continued)

<i>2000 ANES</i>	<i>2000 Harris</i>
Gender: Interviewer coded: –Male (1) –Female (0)	Gender: Are you: –Male (1) –Female (0)
Race: What racial or ethnic group best describes you? –Black (1) –Asian (0) –Native American (0) –Hispanic or Latino (0) –White (0) –Other (0)	Race: Do you consider yourself: –White (0) –Black (1) –African-American (1) –Asian or Pacific Islander (0) –Native American or Alaskan native (0) –Mixed racial background (0) –Other race (0)
Education: What is the highest grade of school or year of college you have completed? –00–12 years –13–16 years –17+ years (If 0–12): Did you get a high school diploma or pass a high school equivalency test? –Yes (High School = 1) –No (Baseline Category) (If 13+): What is the highest degree that you have earned? –Bachelor’s Degree (College = 1) –Master’s Degree (Graduate = 1) –PhD LIT, SCD, DFA, DLIT, DPH, DPHIL, JSC, SJD (Graduate = 1) –LLB, JD (Graduate = 1) –JDC, STD, THD (Graduate = 1) –Associate Degree (AA) (College = 1) –No degree earned (Some College = 1)	Education: What is highest level of education you have completed or the highest degree you have received? –Less than high school (Baseline Category) –Completed some high school (Baseline Category) –High school graduate or equivalent (e.g., GED) (High School = 1) –Completed some college, but no degree (Some College = 1) –College graduate (e.g., B.A., A.B., B.S.) (College = 1) –Completed some graduate school, but no degree (College = 1) –Completed graduate school (e.g., M.S., M.D., Ph.D.) (Graduate = 1)
Age: What is the month, day and year of your birth? Coded into dummies representing four age groups: 25–34, 35–44, 45–54, 55–64, and 65+. Respondents between 18–24 years old represent baseline category.	Age: What is the year of your birth? Coded into dummies representing four age groups: 25–34, 35–44, 45–54, 55–64, and 65+. Respondents between 18–24 years old represent baseline category.

## References

- Chang, LinChiat, and Jon A. Krosnick. 2002. Comparing self-administered computer surveys and auditory interviews: An experiment. Paper presented at the American Association for Public Opinion Research Annual Meeting, St. Petersburg, FL.
- Converse, Philip E. 1964. The nature of belief systems in mass publics. In *Ideology and discontent*, ed. D. E. Apter, 206–61. New York: Free Press.

- Erikson, Robert S., and Kent L. Tedin. 2007. *American public opinion*. 7th ed., updated. New York: Longman.
- Holbrook, Allyson L., Melanie C. Green, and Jon A. Krosnick. 2003. Telephone vs. face-to-face interviewing of national probability samples with long questionnaires: Comparisons of respondent satisficing and social desirability response bias. *Public Opinion Quarterly* 67:79–125.
- Rossi, Peter H., James D. Wright, and Andy B. Anderson. 1983. *Handbook of survey research*. Orlando, FL: Academic Press.
- Saris, Willem. 1998. Ten years of interviewing without interviewers: The Telepanel. In *Computer assisted survey information collection*, ed. Mick P. Couper, Reginald P. Baker, Jelke Bethlehem, Cynthia Z. F. Clark, Jean Martin, William L. Nicholls, II, and James M. O'Reilly, 409–29. New York: Wiley.
- Sears, David O. 1986. College sophomores in the laboratory: Influences of a narrow database on social psychology's view of human nature. *Journal of Personality and Social Psychology* 51(September):515–30.
- Verba, Sidney, Kay Lehman Schlozman, and Henry E. Brady. 1995. *Voice and equality: Civic voluntarism in American politics*. Cambridge, MA: Harvard University Press.
- Wolfinger, Raymond E., and Steven J. Rosenstone. 1980. *Who votes?* New Haven, CT: Yale University Press.