Advance Access publication December 6, 2006

Political Analysis (2007) 15:257–285
doi:10.1093/pan/mpl010

Does Mode Matter For Modeling Political Choice? Evidence From the 2005 British Election Study

David Sanders
Department of Government, University of Essex, Colchester, England CO4 3SQ
e-mail: sanders@essex.ac.uk (corresponding author)

Harold D. Clarke and Marianne C. Stewart
School of Economic, Political and Policy Sciences, University of Texas at Dallas, Richardson, TX 75083
e-mail: hclarke@utdallas.edu
e-mail: mstewart@utdallas.edu

Paul Whiteley
Department of Government, University of Essex, Colchester, England CO4 3SQ
e-mail: whiteley@essex.ac.uk

Although political scientists have begun to investigate the properties of Internet surveys, much remains to be learned about the utility of the Internet mode for conducting major survey research projects such as national election studies. This paper addresses this topic by presenting the results of an extensive survey comparison experiment conducted as part of the 2005 British Election Study. Analyses show statistically significant, but generally small, differences in distributions of key explanatory variables in models of turnout and party choice. Estimating model parameters reveals that there are few statistically significant differences between coefficients generated using the in-person and Internet data, and the relative explanatory power of rival models is virtually identical for the two types of data. In general, the in-person and Internet data tell very similar stories about what matters for turnout and party preference in Britain. Determining if similar findings obtain in other countries should have high priority on the research agenda for national election studies.

1 Introduction

In recent years, political scientists have recognized the need to determine if well-designed and well-executed Internet surveys produce results comparable to those derived from conventional telephone and face-to-face surveys (Krosnick and Chang 2001; VanBeselaere 2002; Alvarez, Sherman, and VanBeselaere 2003; Berrens et al. 2003; Chang and Krosnick

Authors' note: We thank the U.K. Economics and Social Research Council (ESRC) and Gary Williams, Senior Science Manager at the ESRC, for their generous support of, and interest in, this project. We also thank the editor and anonymous reviewers for helpful comments and suggestions.

© The Author 2006. Published by Oxford University Press on behalf of the Society for Political Methodology.
All rights reserved. For Permissions, please email: journals.permissions@oxfordjournals.org

https://doi.org/10.1093/pan/mpl010 Published online by Cambridge University Press

2003). Although relevant information has accumulated, much remains to be learned about the comparative performance of Internet surveys in major ongoing data collection projects such as national election studies. This paper addresses this topic by presenting the results of an extensive survey comparison experiment that was conducted as part of the 2005 British Election Study (BES). The core of the experiment involved administering identical survey questions to respondents from two different but simultaneously conducted national panel surveys of the British electorate. The first survey, conducted by the National Centre for Social Research (Natcen), utilized face-to-face interviews and a national probability sample. Computer Assisted Personal Interviewing (CAPI) techniques aside, this is essentially the same methodology that has been employed by successive BES teams since the project began in 1963. The second data collection was an Internet survey conducted by YouGov, the Internet survey market leader in the United Kingdom. YouGov uses a methodology similar to that employed by Harris Interactive in the United States. Respondents to any given survey are selected randomly from a larger pool of people recruited from a wide range of Internet sources, who have already agreed to participate in YouGov surveys.

In this paper, we first locate our study within the context of current concerns and existing research on the efficacy of alternative survey modes. Then, we describe the data collection and weighting procedures employed in the 2005 BES in-person and Internet surveys. Next, we summarize marginal distributions on a number of key variables observed in the pre- and postelection waves of the in-person and Internet surveys. The results show that there are statistically significant, but generally small, differences in these marginal distributions. The next section focuses on the pivotal question of causal inference. To this end, we analyze identical models of *turnout* using the two data sets, and then we conduct an identical exercise with respect to models of *party choice*. Analyses reveal that there are few statistically significant differences between the estimated coefficients of the models across the two data sets. Also, the relative explanatory power of rival models of voting behavior is virtually identical for analyses using face-to-face and Internet data. In essence, the in-person and Internet survey data yield very similar inferences about what matter for the decision whether or not to vote and the sources of party preference. We believe these results constitute strong support for the claim that properly designed Internet surveys can have a useful role to play in research on electoral choice in Britain.

2 National Election Surveys by Internet?

Interest in using Internet surveys to study voting behavior in national elections derives from a number of related considerations. Two important ones are money and time. Traditional national election studies conducted using in-person interviews are expensive (by social science standards) and "slow moving." The former means that adequate funding may be difficult to secure, and various design compromises, including smaller than desirable sample sizes, may have to be accepted. The latter means that data are gathered over lengthy time periods, and precise controls on the timing of interviews are difficult to implement. In the case of traditional national postelection surveys, lengthy fieldwork periods engender additional data quality issues since the onward march of political events and conditions may influence survey responses obtained weeks or even months after an election has occurred.

The initial response to these problems, some two decades ago, was to turn to random digit dialing (RDD) telephone methods. Although increasingly expensive, RDD surveys are and likely will remain considerably cheaper than in-person ones. RDD surveys have the additional advantage of enhanced control of interview timing. Quick turnaround for an

https://doi.org/10.1093/pan/mpl010 Published online by Cambridge University Press

entire survey can be easily achieved. Also, if desired, the scheduling of interviews can be distributed in a disciplined way over the course of an election campaign, thus enabling monitoring of trends in vote intentions and key explanatory variables in the run-up to Election Day (Johnston and Brady 2002).

But the telephone mode is not a panacea. One problem is declining response rates (Smith 1995; Groves and Couper 1998; Berrens et al. 2003; Rivers 2006; see also Steeh 1981). Similar to in-person surveys, substantial proportions of people selected for a telephone interview do not provide one. Large numbers of noncontacts, refusals, and breakoffs in telephone (and in-person) surveys have deleterious consequences for unit and item nonresponse and thereby threaten inferences about quantities of interest (e.g., Groves and Couper 1998; see also Brehm 1993; Alvarez, Sherman, and VanBeselaere 2003).[1] Telephone surveys have additional drawbacks. It is difficult to conduct lengthy interviews of the sort that are typical with national election surveys done with in-person interviews.[2] Attempts to do long telephone interviews risk respondent fatigue, irritation, terminations, and attendant threats to data quality and sample integrity. Respondents' negative reactions to lengthy surveys also pose a threat to panel integrity, should a research design involve multiple interviews. Another problem concerns measurement difficulties and limitations produced by the inability of respondents to access visual stimuli such as "show cards" and CAPI presentations (e.g., Tourangeau, Rips, and Rasinski 2000). And there is evidence of increased social desirability biases in telephone surveys (Holbrook, Green, and Krosnick 2003).

Internet surveys may effectively address a number of these problems. Internet surveys are relatively inexpensive, and as Berrens et al. (2003) emphasize, marginal costs are very low. Very large *N* samples become realistic possibilities. Also, Internet surveys are "quick and agile." Huge amounts of data can be gathered within hours, and tight controls can be implemented on the timing of interviews for rolling campaign data collections. There are other possible advantages as well. For example, since Internet surveys are self-administered, interviewer effects are eliminated, and social desirability biases may be minimized. And since respondents proceed at their own pace once they start an interview, lengthy interviews may be seen as less of a burden than is the case for telephone surveys. Yet another set of advantages are inherent in Internet technology, which facilitates the use of visual and audio stimuli and thereby enables sophisticated experiments (Couper, Tourangeau, and Kenyon 2004).[3]

These desirable features notwithstanding, proposals to use the Internet to conduct major national surveys have met with skepticism (e.g., Baker, Curtice, and Sparrow 2002;

[1]Some studies have concluded that declining and low response rates do not pose serious threats (see, e.g., Curtin, Presser, and Singer 2000, 2005; Keeter et al. 2000). The effects of declining and low response rates likely vary across different areas of inquiry, and much remains to be learned about consequences for national election studies in various countries.

[2]Problems attendant upon lengthy telephone surveys vary. Some studies, e.g., the CNES, have been able to do quite long RDD surveys (nearly 30 min on average) and achieve impressive pre-post election panel retention rates, e.g., 78% in the 2000 Canadian election study (CES) (Institute for Social Research n.d.). Self-completion mail-back questionnaires, such as those used by the BES and CES, provide a means of augmenting the total amount of data collected. For the 2005 BES, 3226 of the 4161 (77.5%) postelection respondents completed a self-completion questionnaire distributed at the end of the in-person interview. Using supplementary Internet surveys are another way of increasing data collections initiated with in-person or RDD surveys. On response rates in Internet and mail surveys, see Kaplowitz, Hadlock, and Levine (2004).

[3]For example, the 2005 BES included a "feedback-to-respondent" experiment in the precampaign Internet survey. The aim of the experiment was to investigate how respondent self-placements in a two-dimensional issue space responded to party and party leader cues. Respondents indicated their positions on the issue scales and later in the survey were presented with a two-dimensional graphic display showing their position and those of varying combinations of parties and party leaders. Respondents were told they could use their computer pointing device to change their position, if they so desired.

https://doi.org/10.1093/pan/mpl010 Published online by Cambridge University Press

Schoen and Faas 2005; see also Couper 2000; Dillman 2000). A principal reason concerns sampling. Since there is no master list of Internet addresses comparable to an electoral register or a postcode address file, it is impossible to draw probability samples without first using in-person, telephone, or mail surveys to construct a sampling frame. It is argued that absent being recruited via probability methods, Internet panels are likely to be unrepresentative of populations of interest, and "Literary Digest" types of threats to inference are ever present (Dillman 2000). Other sampling problems include unit nonresponse and limited coverage. Similar to other modes, responses rates for Internet surveys are far from perfect. Also, although Internet access is widespread and continues to expand, even in wealthy mature democracies, substantial numbers of people are not online. For example, although Internet access has grown substantially in Britain, Canada, and the United States, circa 2005, 30%–35% of the populations of these countries had zero probability of self-initiated recruitment into an Internet survey.[4] Also, Internet access is not distributed randomly; rather, disproportionate numbers in groups such as the disabled, elderly, poor, and minorities remain off-line.

Internet survey firms have responded to these criticisms in two ways. In the United States, Knowledge Networks uses telephone recruitment surveys to construct and refresh probability-generated Internet panels (Berrens et al. 2003; see also Alvarez, Sherman, and VanBeselaere 2003). Other companies, such as Harris Interactive and the British company YouGov, use sophisticated recruitment and weighting schemes in efforts to offset sampling biases. And some firms, e.g., YouGov, try to bolster response rates by offering (modest) financial incentives. Over the past half-decade, these three survey houses have established strong records for accurately predicting vote shares in various national and subnational elections, thereby suggesting that the Internet may prove useful for doing electoral research in these countries (e.g., Taylor et al. 2001; Berrens et al. 2003). The analyses presented below are designed to investigate this possibility in the context of the 2005 British general election. First, however, we present basic information regarding the 2005 BES in-person and Internet surveys.

3 The In-Person and Internet Surveys

The design of the 2005 BES mode comparison study is summarized in Fig. 1. As indicated, the 2005 BES conducted two parallel panel surveys. The core study was a two-wave face-to-face national probability panel survey, with the first wave conducted in February–March 2005 and the second wave conducted in May–July 2005, starting right after the May 5 general election. The face-to-face study was complemented by a three-wave Internet panel survey. The first Internet wave was conducted in March 2005, the second wave was implemented during the official campaign in April 2005, and the third went into the field in May 2005, immediately after the election.[5] The preelection wave questionnaires in both the face-to-face and Internet surveys were identical, insofar as this was possible given that different modes were involved. The Internet postelection survey was quite short, reflecting the fact that the Internet respondents had already been interviewed a second time during the campaign. However, it did include a number of key questions—about turnout and party choice—that were asked in the more extensive postelection face-to-face survey.

[4]For comparative data on Internet usage, see http://www.interworldstats.com. In a recent paper on the Internet surveys in the United States, Rivers (2006) observes that coverage problems associated with Internet surveys are diminishing rapidly: "[t]he Digital Divide has diminished substantially and will largely disappear in the next decade, as the Internet becomes the vehicle for the delivery of home entertainment and phone services."

[5]All respondents to the precampaign survey were recontacted in the postelection survey, regardless of whether they had participated in the campaign wave. The three-wave (precampaign-campaign-postelection) panel $N = 4894$.

https://doi.org/10.1093/pan/mpl010 Published online by Cambridge University Press

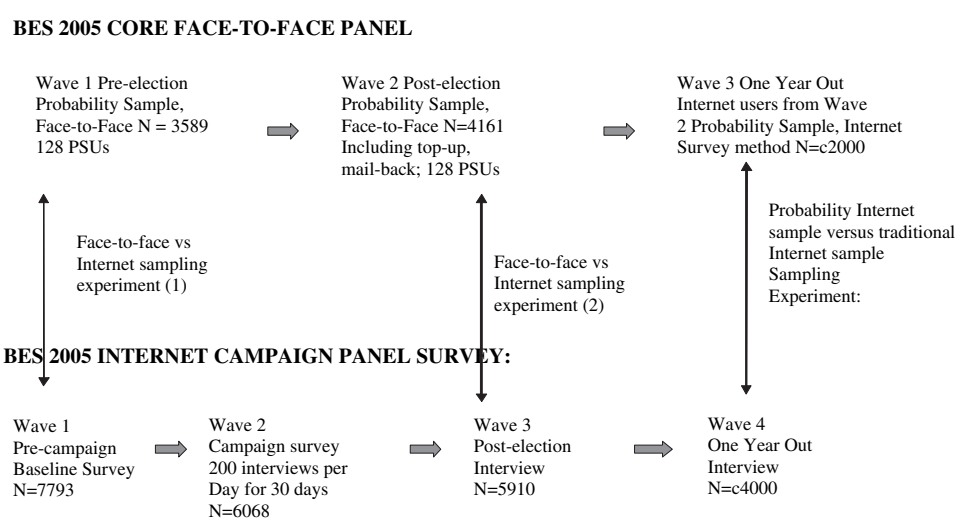


Fig. 1 In-person and internet panel survey design in the 2005 BES.

3.1 *In-Person Surveys*

As noted above, the 2005 BES in-person preelection baseline survey was conducted before the election campaign officially began. The survey was designed to yield a representative sample of "noninstitutionalized" adults aged 18 years and older living in Great Britain (people living in Northern Ireland and Scots living north of the Caledonian canal were excluded). A clustered multistage design was employed.[6] First, 128 constituencies were sampled (77 in England, 29 in Scotland, and 22 in Wales). Constituencies were sampled using three stratification criteria: (1) electoral marginality in the 2001 general election, (2) region in England/Scotland and percent Welsh speakers in Wales, and (3) population density. Within each constituency selected, two wards were randomly chosen, and within each ward, household addresses were selected with equal probability from the national postcode address file. For households with multiple occupants, one person (the potential respondent) was selected at random using a modified Kish grid.

The *N* for the preelection campaign survey was 3589, with a response rate of 60.5%. Beginning immediately after the election, all the preelection respondents were asked to do a second in-person interview. The resulting pre-post panel *N* was 2959 (panel retention rate = 82.4%). To provide a representative national postelection sample, the panel was supplemented by a "top-up" sample ($N = 1202$) chosen using the methods described above. All the postelection top-up respondents were interviewed in-person. The unweighted postelection sample *N* thus was 4161, and altogether, 4791 respondents participated in one or both of the in-person interviews.

The in-person survey data were weighted using a combination of factors designed to correct for unequal selection probabilities arising from deliberate oversampling in Scotland and Wales, deliberate oversampling of marginal constituencies, variation in the number of households at selected addresses, and variation in the number of people

[6]For details, see the *British Election Study 2005 Technical Report* prepared by the Natcen BES team, Mark Johnson, Katarina Thomson, and Shaun Scholes. The report may be downloaded from the BES website http://www.essex.ac.uk/bes.

https://doi.org/10.1093/pan/mpl010 Published online by Cambridge University Press

living in selected households.[7] In addition, a set of poststratification or "calibration" weights for age and gender were employed.

3.2 *Internet Surveys*

Similar to the in-person preelection survey, the first wave of the Internet survey was conducted just before the election campaign formally began. Potential Internet respondents were selected from YouGov's master panel, which included 89,000 people at the time the study was conducted.[8] People join the YouGov master panel in one of three ways: (1) by visiting the YouGov website (http://www.yougov.com) and registering, (2) by being recruited by one of several professional third-party recruiters (e.g., Win4Now) employed by YouGov, (3) through ad hoc alliances between YouGov and partners such as media outlets interested in conducting specific survey research projects. Respondents in such surveys can be invited to join the YouGov master panel.

Potential respondents for the BES preelection baseline Internet survey were randomly selected from subsections of the master panel defined in terms of demographics (age, gender), media consumption (newspaper readership), and a political criterion (reported vote in the preceding [2001] general election). The total (unweighted) *N* for the YouGov precampaign survey was 7793. During the election campaign, 6068 of these respondents participated in a rolling campaign panel survey designed to track the dynamics of public opinion as the campaign unfolded. Immediately after the election, 5910 of the precampaign respondents participated in a postelection survey. The response rate for the initial precampaign survey was 52.0%, and panel retention rates were 77.9% (campaign survey) and 75.8% (postelection survey).

After the three waves of the Internet survey were completed, poststratification weights for the data were developed using demographic criteria (gender, age within gender, region, and social class), as well as newspaper readership and vote in the 2001 general election. Similar to the in-person surveys, information from the 2001 U.K. census was used to develop the demographic weighting factors for the Internet surveys. Data from the National Readership Survey (an annual random probability in-person survey with 34,000 respondents) were used to construct the newspaper readership weighting factor, and the past vote weighting factor was developed based on the results of a large in-house analysis of false-memory effects.[9]

4 Comparing Marginal Distributions

When comparing the pattern of responses to these two different panel surveys, we employ the weights supplied by the companies that collected the data (see above). It should also be noted that, when comparing the patterns of response, we cannot distinguish between

[7]Oversampling in Scotland and Wales is typical in BES surveys and is done to provide sufficient cases for analyses of electoral choice in those regions. Oversampling marginal constituencies was done to provide a sufficient number of such constituencies for analyses of factors that affected election outcomes in various types of "tight races" across Britain.

[8]Technical information regarding YouGov's sampling procedures were supplied by YouGov BES project director, Joe Twyman, and is posted on the BES website (see note 6).

[9]Inspection of distributions of several demographic variables (age, education, ethnicity, gender, income, region) shows that the impact of weighting on these distributions is generally quite mild (see Appendix 1). Across 31 categories of these variables, the average absolute deviation between the unweighted and weighted versions is 1.2% and the range is 0.0%–4.2%. This is also true for the face-to-face surveys, with the exception of region, where the study design involved deliberate oversampling of Scotland and Wales. Excepting region, the average absolute deviation between the unweighted and weighted versions of the face-to-face sample is 2.1% and the range is 0.1%–5.4%.

https://doi.org/10.1093/pan/mpl010 Published online by Cambridge University Press

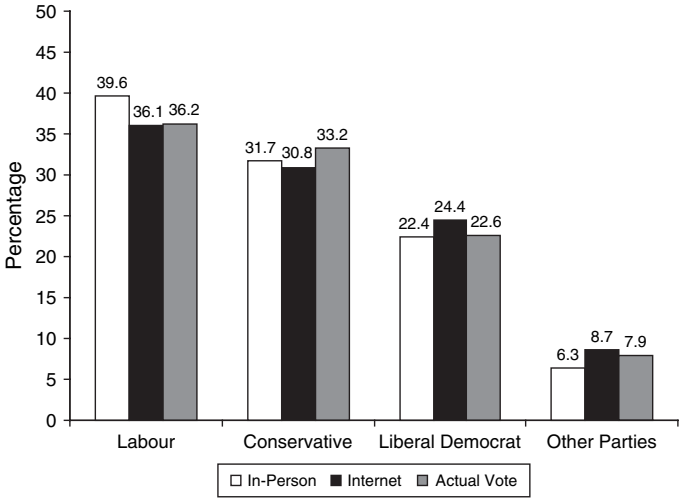


Fig. 2 Reported vote in in-person and Internet postelection surveys and actual vote (Great Britain). Survey comparison: chi square = 23.96, df = 3, $p < .001$, $V = .06$. Source: 2005 BES and Kavanagh and Butler (2005, Appendix 1).

sampling frame effects (probability sample versus sample drawn from a master Internet panel) and survey mode effects (in-person interview in the respondent's home versus impersonal computer interview). We therefore address the overall comparability of the results of the two surveys.[10]

Figure 2 reports the vote shares various parties recorded in the two surveys and compares them with the actual results of the 2005 general election. The in-person survey overestimated the governing Labour Party's share by 3.4%, whereas the Internet survey underestimated it by only 0.1%. As for Conservative support, the in-person survey missed the mark by 1.7% and the Internet survey missed by 2.4%. Comparable figures for Liberal Democrat support are 0.2% (in-person) and 1.8% (Internet). For minor parties, the discrepancies are 2.4% (in-person) and 0.8% (Internet). Overall, the Internet survey was marginally more accurate in its vote share estimates (mean absolute error = 1.3%) than the in-person survey (mean absolute error = 1.7%).

Figure 3 shows reported and actual 2005 turnout. Since their inception, the BES in-person surveys have consistently overreported voting turnout—by an average of 9.9% for the 11 British general elections held between 1964 and 2001.[11] To provide perspective on this pattern, we note that the tendency to exaggerate turnout is also typical of other national election studies. For example, the 2004 American National Election Study (ANES) (in-person survey) overreported turnout by 24.7% among the voting age population using the

[10]The 2005 BES conducted a further set of interviews in late spring 2006. All Internet panel respondents were recontacted and asked to do a follow-up Internet survey. At the same time, all respondents in the postelection wave of the face-to-face probability survey who had Internet access were asked to complete an identical Internet survey. To avoid house effects, both surveys were conducted by YouGov. These two follow-up surveys will enable us to compare the Internet mode responses of respondents selected using two different two sampling frames: a typical Internet sample in which respondents are originally selected from a master Internet panel and a group of Internet users from a traditional BES probability sample. The results of these analyses will be reported in a later paper.

[11]The 2001 BES also included a pre-post election rolling campaign telephone survey. This survey reported turnout at 78.1%, 18.7% above the actual figure.

https://doi.org/10.1093/pan/mpl010 Published online by Cambridge University Press

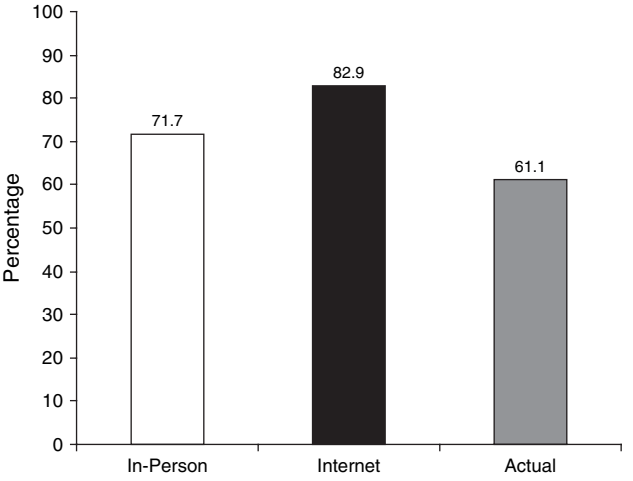


Fig. 3 Reported turnout in in-person and Internet surveys and actual turnout in 2005 BES. Source: 2005 BES and Kavanagh and Butler (2005, Appendix 1).

traditional ANES turnout question and by 17.3% using a revised question designed to minimize the tendency to overreport. Comparable figures for the 2002 ANES (conducted by telephone) were 37.9% for the traditional question and 17.8% for the revised one. Similarly, using the revised question, the 2000 ANES overestimated electoral participation by 20.2% among respondents interviewed in-person and by 27.2% by those interviewed by telephone. Canadian studies show the same tendency. In the 2004 Canadian National Election Study (CNES) (telephone interviews), the reported voting rate was 86.4%, whereas the official figure was 60.9%—a 25.5% difference. The discrepancy in the previous (2000) CNES was also large—21.5%.

Turnout overreporting continued in both of the 2005 BES surveys. Actual electoral participation in 2005 was a modest 61.4%, whereas self-reported voting in the probability sample was 71.7%, and in the Internet survey it was 82.9%. Although there are various possible explanations for these inflated estimates, one is that people who are interested in politics are more likely to agree to participate in political surveys.[12] Indeed, as we show below, reported political interest was significantly higher in the Internet sample than in the probability sample, which may partly explain why the turnout estimate in the former is even higher than in the latter. However, even if we weight the Internet data by the levels of political interest observed in the probability sample, reported turnout in the Internet sample remains above 80%. It is nonetheless evident that in terms of reported turnout neither survey accurately represented the level of British electoral participation in 2005. The in-person survey missed the target by a sizable margin (10.3%), and the Internet survey missed by a considerably larger one (21.5%). Although these British numbers are not atypically large in comparison with the figures for recent U.S. and Canadian

[12]Although the aggregate overreport of turnout in national elections studies may reflect sampling bias, it also may reflect the impact of a preelection interview in pre-post panel surveys. Yet another explanation may involve misreporting due to social desirability biases. Since voting is widely seen to be a civic duty, some respondents may say they have voted when they have not. A vote validation study for the 2001 BES indicated that misreporting and sampling bias played approximately equal roles in accounting for the turnout over report. The 2005 BES vote validation study is currently being conducted for all the in-person survey data and, because of cost considerations, for a sample of the Internet data.

https://doi.org/10.1093/pan/mpl010 Published online by Cambridge University Press

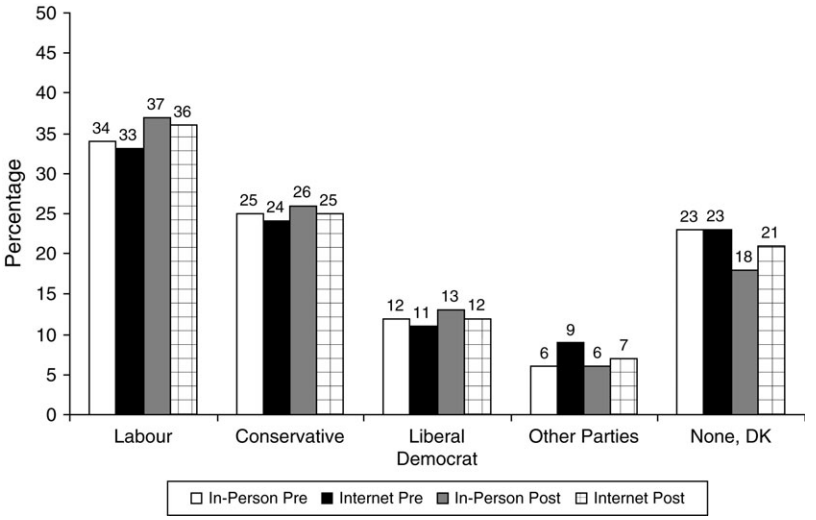


Fig. 4 Party identification in pre- and postelection in-person and Internet surveys. Preelection surveys: chi square = 22.97, df = 4, $p < .001$, Cramer's $V = .06$. Postelection surveys: chi square = 24.67, df = 4, $p < .001$, Cramer's $V = .05$. Source: 2005 BES.

surveys reported above, overestimating turnout represents a continuing measurement problem for British and other national election studies. Below, we investigate whether the in-person and Internet mode differences in reported turnout in the 2005 BES surveys are consequential for inferences regarding forces affecting electoral participation.

Another key variable in studies of electoral choice is party identification. Figure 4 shows the percentages of party identifiers in Britain as measured in the pre- and post-election waves of the in-person and Internet surveys. Unlike party vote shares and turnout, there are no benchmark figures against which one can compare the distribution of partisan attachments. However, we note that the percentages of party identifiers are very similar in the two surveys (mean absolute difference = 1.3%), although there are slightly less Labour, Conservative, and Liberal Democrat identifiers recorded in each wave of the Internet survey than in the in-person one. A chi-square test indicates that, collectively, these several very small differences are statistically significant, but this clearly is a product of the very large sample sizes of the combined in-person and Internet surveys.

This pattern of relatively small but statistically significant differences between the two surveys continues across a range of other theoretically important variables. Table 1 compares the mean scores for variables in the two surveys used in the model of turnout reported below. These variables are the ones we used in articles on the 2005 U.K. general election (Clarke et al. 2006; Whiteley et al. 2006), as well as in a recent book on voting behavior in Britain (Clarke, Sanders, Stewart, and Whiteley 2004). The upper and lower bounds of the scales vary according to the question wordings used. Most vary either from 0 to 10 (e.g., party leader affect scores) or from 1 to 5 (e.g., "cost of voting" variables). However, there is also one 1–4 scale (interest in the general election) and two 0–1 dummy variables (party mobilization variables).

It would be tedious to describe all the differences between the two sets of measures shown in Table 1. We accordingly highlight one variable for illustrative purposes. The first row of the table shows the average scores on a 0–10 scale that asked respondents to rate their influence on politics and public affairs. The mean score in the in-person survey was

https://doi.org/10.1093/pan/mpl010 Published online by Cambridge University Press

Table 1 Mean scores for predictor variables in turnout model, in-person probability and Internet surveys

	In-person		*Internet*			
Predictor variable	*Mean*	*N*	*Mean*	*N*	*p*	*Eta*
Perceived influence (0–10)*	2.7	4136	2.3	5791	.001	.08
Feelings about parties						
Labour (0–10)*	5.2	4122	4.7	5757	.001	.08
Conservative (0–10)*	4.6	4052	3.8	5733	.001	.12
Liberal Democrat (0–10)*	5.2	4015	5.1	5669	.05	.02
SNP, Scotland only (0–10)*	4.8	350	4.5	473	.10	.06
PC, Wales only (0–10)*	4.2	185	3.8	305	.08	.08
Personal benefits						
Political action benefits self (1–5)	3.1	3509	3.2	7260	.01	.03
Feel guilty if not vote (1–5)	2.6	3572	2.5	7647	.001	.04
Satisfaction from voting (1–5)	2.5	3552	2.4	7461	.001	.07
Costs of voting						
Voting—too much time (1–5)	2.8	3532	2.9	7497	.001	.02
Voting—too busy (1–5)	3.4	3549	3.5	7476	.001	.07
Civic duty						
Voting every citizen's duty (1–5)	2.1	3576	2.0	7747	.001	.05
Not voting neglect of duty (1–5)	2.4	3575	2.3	7676	.001	.03
Democracy dissatisfaction (1–5)	2.4	3490	2.6	7319	.001	.14
Election interest (1–4)	2.0	4158	1.8	5905	.001	.16
Party mobilization						
Respondent canvassed (0–1)*	0.21	4161	0.19	5936	.001	.03
Respondent telephoned (0–1)*	0.07	4161	0.08	5936	.001	.01
Relative deprivation						
Government treats R fairly (1–5)	3.0	3565	3.5	7633	.001	.21
Big gap R's expectations and outcomes (1–5)	2.5	3562	2.3	7653	.001	.11
Social norms: friends and family think voting a waste of time (1–5)	3.5	3550	3.4	7478	.01	.03
Social trust—people can be trusted (0–10)	6.1	3570	5.5	7623	.001	.12

Note. All variables except those designated with an asterisk (*) are measured in the preelection surveys. SNP, Scottish National Party; PC, Plaid Cymru; R, Respondent.

2.7, compared with 2.3 in the Internet survey, suggesting that the Internet respondents on average felt somewhat less politically efficacious than their in-person counterparts. As the table indicates, this difference is statistically significant at $p \leq .001$, although with such large *N*'s (4136 for the in-person survey, 5791 for the Internet survey) a high level of significance is not difficult to achieve.

Rather more informative than the significance level in this context is the eta statistic, which measures the strength of the mode (in-person versus Internet) effect. The value of eta = .08 for the perceived influence term suggests that the mode effect is very small. Indeed, if we calculate eta-squared × 100, which represents the variance explained by the mode effect, this gives (.08 × .08 × 100) = 0.64%. In other words, less than 1% of the variance in perceived influence across the two surveys is explained by the survey mode used to measure perceived influence. The explained variance across all the variables shown

https://doi.org/10.1093/pan/mpl010 Published online by Cambridge University Press

in Table 1 ranges from .01% (for the telephone canvass measure) to 4.4% (for the "Government treats the respondent fairly" measure). The average explained variance is small, 0.82%. This suggests that, although the measures recorded using the two survey modes are different from one another (see the uniform pattern of significant p values in Table 1), mode differences are, on average, very modest.

Table 2 reports an equivalent set of mean scores for the predictor variables in the model of party choice reported below. As with the turnout predictors, the measures displayed in Table 2 are all derived from competing models of party choice that we have presented elsewhere (Clarke, Sanders, Stewart, and Whiteley 2004; Clarke et al. 2006; Whiteley et al. 2006). The list of variables covers the "usual suspects" typically included in rival party choice models (leader perceptions, party judged most competent on the most important issue, party identification, party-issue proximities, and economic evaluations), together with variables that are particularly relevant to contemporary U.K. electoral politics (attitudes toward the Iraq War and tactical voting).

The results presented in Table 2 suggest a number of interesting differences and similarities between the in-person and Internet samples. The most noticeable *differences*, as revealed by the eta values, are evident in relation to the leader affect scores and the party-issue proximity scores. The former indicate that Internet respondents were less favorably impressed than in-person respondents with all three major party leaders. Tony Blair's average rating among in-person respondents was 4.9 on the 0–10 scale; among Internet respondents, his rating was only 4.1. Similarly, Michael Howard's average "in-person" rating was 4.4 compared with 3.6 for Internet respondents. The average ratings difference for Charles Kennedy was smaller, but he was also viewed less positively in the Internet survey (5.3) than in the in-person survey (5.5). With regard to the (0–10) party-issue proximity scores, both the tax-spend scale and the European Union (EU) scale responses differ noticeably across the two survey modes. High (low) values on the tax-spend scale represent a preference for increasing (reducing) taxes and higher (lower) spending on public services. High (low) values on the EU scale represent a preference for more (fewer) EU decision-making powers. In both cases, Internet respondents tend to place the parties more toward extremes of the scale. For example, on the EU scale, the average in-person respondent scores the Conservatives at 5.1, whereas the average Internet respondent scores them at 4.3; by the same token, Labour is placed at 7.0 by in-person respondents and at 7.6 by Internet respondents.

The most notable *similarities* in Table 2, as revealed by the nonsignificant p values, relate to the dummy variable measures of party identification and attitudes toward the Iraq War. The mean party identification scores (i.e., the proportions of identifiers) for each of the three main parties are virtually identical across both modes. The mean Labour identification score for the in-person survey is 0.34; for the Internet survey, it is 0.33. The equivalent Conservative figures are 0.25 and 0.24, respectively; the Liberal Democrat scores, 0.11 and 0.11, respectively. A similar pattern is observed in relation to the two Iraq War variables. In both surveys, the mean approval score for the government's handling of the war/occupation is 3.8 on a 1–5 scale. Similarly, in both cases, the mean approval score for British involvement in the conflict is 2.9. These findings represent something of a corrective to the general pattern of significant difference observed previously. In terms of partisanship and attitudes toward the most widely publicized event in British politics between the 2001 and 2005 elections, there are virtually no differences between the Internet and probability sample respondents.

The comparisons presented in Tables 1 and 2 suggest three general conclusions. First, in terms of vote shares and turnout, the in-person and Internet surveys produce two different

https://doi.org/10.1093/pan/mpl010 Published online by Cambridge University Press

Table 2 Mean scores for predictor variables in Labour voting model, in-person probability and Internet surveys

	In-person		*Internet*			
Predictor variable	*Mean*	*N*	*Mean*	*N*	*p*	*Eta*
Party leader affect						
Blair (0–10)*	4.9	4129	4.1	5758	.001	.13
Howard (0–10)*	4.4	4002	3.6	5650	.001	.14
Kennedy (0–10)*	5.5	3969	5.3	5595	.001	.04
Party best, most important issue						
Labour (0–1)*	0.35	4061	0.25	5875	.001	.11
Conservative (0–1)*	0.22	4061	0.24	5875	.01	.03
Liberal Democrat (0–1)*	0.07	4061	0.09	5875	.05	.02
Other party (0–1)*	0.04	4061	0.09	5875	.001	.09
Party identification						
Labour (0–1)	0.34	3568	0.33	3922	.635	.01
Conservative (0–1)	0.25	3568	0.24	3922	.119	.02
Liberal Democrat (0–1)	0.11	3568	0.11	3922	.333	.01
Other party (0–1)	0.06	3568	0.09	3922	.001	.05
Party-issue proximities						
Tax-spend scale						
Self (0–10)	6.2	3491	5.5	7345	.001	.13
Labour (0–10)	6.4	3380	6.6	6609	.001	.05
Conservatives (0–10)	5.3	3275	4.1	6405	.001	.24
Liberal Democrats (0–10)	6.0	3061	6.7	5947	.001	.15
EU scale						
Self (0–10)	5.4	3410	5.2	7228	.001	.03
Labour (0–10)	7.0	3246	7.6	6616	.001	.12
Conservatives (0–10)	5.1	3098	4.3	6275	.001	.14
Liberal Democrats (0–10)	5.9	2810	6.9	5676	.001	.19
Economic evaluations						
Personal retrospective (1–5)	2.9	3560	2.7	7678	.001	.08
National retrospective (1–5)	2.7	3487	2.8	7280	.01	.03
Personal prospective (1–5)	3.0	3484	2.8	7295	.001	.11
National prospective (1–5)	2.9	3389	2.8	6837	.001	.07
Iraq War						
Approve government handling (1–5)	3.8	3517	3.8	7697	.16	.01
Approve U.K. involvement (1–5)	2.9	4067	2.9	5842	.33	.01
Tactical voter/not (0–1)[a]	0.11	2933	0.19	4922	.001	.11

Note. All variables except those designated with an asterisk (*) are measured in the preelection surveys.
[a]Tactical voters are defined as those respondents who indicated that the most important reason for their party choice was that their preferred party could not win in their constituency or those who said that they voted tactically.

sets of marginal distributions. Neither of these distributions conforms precisely to the known vote share and turnout distributions actually observed in the 2005 U.K. general election. That said, the probability sample distribution was closer to the actual turnout result, but both samples, like those gathered in recent ANES and CNES, were well off

https://doi.org/10.1093/pan/mpl010 Published online by Cambridge University Press

target. In contrast, the Internet sample was closer to the actual *vote share* result, but both samples were quite close to the actual result. Insofar as the representativeness of the two samples can be validated against actual outcomes, therefore, it is by no means clear that the probability sample is "superior" to the Internet sample.

Second, the two surveys clearly produce statistically distinct samples on a range of measures relating to perceptions of the costs and benefits of voting, political issues, party leaders, political engagement, and economic evaluations. However, the actual differences between the distributions on key predictors of turnout and party choice, as measured by the relevant eta statistics, are very small.[13] Insofar as there is a discernible overall pattern to the differences between the two samples, the Internet sample tends to be less "left leaning" than the probability sample. This does not necessarily imply distortion in the Internet sample. On the contrary, as the discrepancy between Labour support as measured by the probability sample (39.6%) and Labour's actual support (36.1%) indicates, the British electorate itself also appears to be less left leaning than the probability sample. Third, regarding party identification and attitudes toward the Iraq War, despite their large *N*'s, the two surveys produce marginal distributions that are statistically indistinguishable. This suggests that, in Britain at least, responses to questions about both long-term partisan attachments and highly salient topical political issues may be resistant to mode effects.

5 Different Modes, Different Inferences?

5.1 *Modeling Turnout*

Analyses presented in the previous section indicate that probability and Internet samples produced significantly different levels of reported turnout and significantly different marginal distributions across a range of important variables. The magnitudes of most of the latter differences were quite modest. As observed above, one of the most frequently voiced criticisms of Internet surveys of the sort conducted by Harris and YouGov is that the "samples" are necessarily biased because the respondents initially have to approach the company in order to indicate their preparedness to be interviewed. This self-selection, even if respondents for any given survey are randomly sampled from the company's master panel, necessarily means that Internet respondents are not typical—and the samples thus assembled are not representative—of the general population. As a result, models of voting behavior based on Internet survey data are at risk of telling different stories than those recounted using data generated by traditional probability sampling methods.

We believe that the principal flaw in this argument concerns practice not theory, i.e., the *achieved samples* typically produced by in-person and RDD surveys using probability sampling methods. As discussed above, similar to Internet surveys, the samples achieved using these methods tend to be only modest fractions of target samples, and it cannot be assumed that nonrespondents are completely random subsets of target samples. As a result, even the most carefully conducted in-person and telephone surveys may generate data sets with covariance structures different from what would be observed if one could obtain textbook-level response rates from probability samples. An important implication is that the question of similarities and differences in models of turnout and party choice estimated

[13] Although not reported here, there is no consistent pattern of differences between the standard deviations for the in-person and Internet survey variables, and most differences are quite small. Data are available upon request.

https://doi.org/10.1093/pan/mpl010 Published online by Cambridge University Press

by data gathered by different survey modes is essentially an empirical one. The "gold standards" in survey research are products of theory not realized in practice.

In this section, we address this question directly by estimating parameters in identical models of vote choice and turnout using the BES in-person and Internet samples. We then test if the coefficients of the respective models are significantly different from one another. If there are significant differences in the pattern of coefficients across the samples, then one can conclude that data from the Internet survey would prompt different inferences from those gathered by the in-person probability survey. If one were prepared to make the (unrealistic) assumption that the probability sample (even with a response rate of 60.5%) is unbiased, this would imply bias in the Internet sample. In contrast, an absence of significant differences implies that it is inconsequential whether one uses the in-person or the Internet data to test models of electoral choice.

Drawing on the results of our recent studies in Britain, we estimate a single-equation model of turnout. This model is a composite specification based on the results of analyzing the performance of several rival theoretical models prominent in the literature on electoral participation (Clarke, Sanders, Stewart, and Whiteley 2004, chap. 7, 8). One is the cognitive engagement model that emphasizes the importance of interest in, knowledge of, and engagement with the political process generally (e.g., Dalton 2002). Another is the civic voluntarism model that focuses on politically relevant resources (e.g., education, time, and income), as well as mobilizing activities by parties and other political groups (e.g., Verba, Schlozman, and Brady 1995). Yet another is the social capital model that stresses the importance of interpersonal trust and location in facilitative social networks (e.g., Putnam 1993). An equity-fairness model sees participation as a response to perceptions of the unfairness and injustice of the operation of the political system (e.g., Runciman 1966). Finally, the general incentives model involves a combination of rational calculation, normative conviction, and social norms (Whiteley and Seyd 2002). The model estimated is

$$\begin{aligned}\text{turnout} = \text{fn}(&b_0 + b_1\text{DBEN} + b_2\text{PERS} + b_3\text{COSTS} + b_4\text{DUTY} + b_5\text{DISS} \\ &+ b_6\text{EINT} + b_7\text{PMOB} + b_8\text{RDEP} + b_9\text{SNORM} + b_{10}\text{STRUST} \\ &+ \Sigma b_{11} - b_{22}\text{DEMOG}),\end{aligned} \tag{1}$$

where DBEN is the perceived benefits of the respondent's preferred party being elected multiplied by the respondent's sense of political efficacy; PERS represents the personal utility the individual derives from voting; COSTS represents the perceived costs of voting; DUTY denotes the extent of the respondent's sense of civic duty; DISS is the extent to which the individual is dissatisfied with the operation of democracy in the United Kingdom; EINT is degree of interest in the general election; PMOB is exposure to in-person campaigning by the political parties during the election campaign; RDEP is the sense of not obtaining just desserts; SNORM is the respondent's view of the attitude to voting of family and friends; STRUST is the extent to which respondents believe people in general can be trusted; DEMOG are demographics, i.e., age, education, ethnicity, gender, region, social class; and b_1–b_{22} are effect coefficients. Additional details concerning several predictor variables in the model are provided in Appendix 2.

We estimate parameters in this turnout model using identical measures of the predictor variables for the in-person and Internet samples.[14] Next, to determine if the effect

[14]Since the dependent variable, turnout, is a 0–1 dichotomy, binomial logit (Long 1997) is used for estimation purposes.

https://doi.org/10.1093/pan/mpl010 Published online by Cambridge University Press

Table 3 Turnout models and mode effects, in-person and Internet pre-post election panels

	In-person sample		*Internet sample*		*Interaction effects with pooled data*	
Predictor variables	*B*	*SE*	*B*	*SE*	*B*	*SE*
Efficacy × collective benefits	0.054***	0.009	0.029***	0.006	−0.025*	0.011
Personal benefits	0.017	0.030	0.86***	0.027	0.069*	0.041
Costs	−0.087**	0.036	−0.058*	0.033	0.029	0.049
Civic duty	0.208***	0.034	0.196***	0.028	0.012	0.044
Democracy dissatisfaction	0.043	0.048	0.010	0.40	−0.033	0.063
Election interest	0.740***	0.065	0.787***	0.054	0.047	0.085
Party mobilization	0.322**	0.113	0.224*	0.099	−0.098	0.150
Relative deprivation	0.059	0.037	0.045	0.032	−0.013	0.049
Social norms	0.134**	0.051	0.157***	0.045	0.023	0.068
Social trust	0.048*	0.027	0.044*	0.020	−0.004	0.033
Age	0.033***	0.003	0.025***	0.003	−0.007	0.005
Education	0.164***	0.036	0.069**	0.027	−0.095*	0.045
Ethnicity (white)	0.401*	0.191	0.402*	0.203	−0.001	0.279
Gender (male)	−0.285**	.0102	0.187*	0.087	0.471***	0.134
Social class	−0.019	.119	−.029	0.094	−0.010	0.151
Region						
Southeast	−0.014	0.201	0.149	0.183	0.163	0.272
Southwest	−0.009	0.244	0.180	0.211	0.189	0.323
Midlands	0.496**	0.212	−0.047	0.190	−0.543*	0.285
North	−0.023	0.206	−0.048	0.184	−0.025	0.276
Scotland	0.091	0.245	−0.008	0.221	−0.098	0.330
Wales	0.313	0.287	0.240	0.249	−0.073	0.380
Constant	−6.29***	0.579	−5.59***	0.526	−6.29	0.579
Mode (Internet) dummy					0.704	0.782
McFadden R^2	.29		.28		.30	
McKelvey R^2	.47		.42		.46	
Percent correctly classified	82.0		87.8		85.9	
Lambda	.35		.22		.28	
N	2985		5831		8789	

*$p < .05$; **$p < .01$; ***$p < .001$; one-tailed test.

coefficients vary significantly across the probability and Internet samples, we combine the two samples into a single data set and specify a series of interaction terms. These interaction terms involve multiplying each of the predictor variables in equation (1) by a 0–1 dummy variable that takes the value of unity if a person is an Internet respondent and zero otherwise. The coefficients on these variables measure the deviations from the effects estimated using the in-person survey data. If an interaction coefficient is statistically significant, it means that the effect of the variable involved differs across the two data sets. If an interaction coefficient is nonsignificant, it implies that there are no significant differences.

Table 3 reports the results. The first column of the table shows the results for the in-person survey data; the second shows the results for the Internet data. The third column shows the interaction effects using the pooled dual-sample data set. The results are well determined and theoretically plausible.

https://doi.org/10.1093/pan/mpl010 Published online by Cambridge University Press

Consider the results in columns 1 and 2 of the table. The pattern of significant and nonsignificant coefficients is virtually identical in both columns. In both the in-person and Internet models, *significant and correctly signed* effects obtain for discounted collective benefits, the perceived costs of voting, civic duty, election interest, party mobilization, social norms, and social trust. Significant positive effects are also observed in both models for three of the demographic controls—age, education, and ethnicity. In both the in-person and Internet models, democracy dissatisfaction and relative deprivation are not significant. The only terms where the models produce *clearly different effects* are for the personal benefits of voting (a significant positive effect in the Internet model but nonsignificant in the in-person model), gender (a significant positive effect in the in-person model but a significant negative effect in the Internet model), and the Midlands region (a significant positive effect for the in-person sample but no effect in the Internet sample).

The key findings of Table 3, however, are shown in column 3. The interaction terms in this column test for significant differences in coefficients estimated using the two data sets. The results show that, out of 23 estimated coefficients, six produce significantly different effect estimates. However, four of these differential effects relate to demographic controls and the constant term. The only two substantive terms that produce differential effects are the two benefits variables. The discounted collective benefits variable produces a slightly smaller effect in the Internet sample ($b = .054 - .025 = .029$) than it does in the in-person sample ($b = .054$), and the personal benefits variable, as indicated previously, is significant in the Internet data but not in the in-person sample. Barring these two exceptions, the message of the generally nonsignificant interaction terms in Table 3 is clear. The overall pattern of substantive effects on turnout in the two samples is very similar. Turnout is negatively affected by the costs of voting but positively affected by its perceived benefits, by sense of civic duty, by interest in the election, by mobilization by political parties, by social norms conducive to participation, and by social trust. Using either the probability or the Internet sample yields almost identical inferences about the determinants of turnout. The "stylized facts" are largely invariant.

5.2 *Modeling Party Choice*

Given the wide variety of models of party choice that have been proposed over the years, specifying a testable model for the purposes of comparing data produced by the two survey methodologies inevitably involves selection. As in the previous section on turnout, we use a specification that replicates a composite vote choice model developed in our recent research. This model is informed by analyses of competing sociological, social psychological, and "soft" rational choice perspectives on party choice (Clarke, Saunders, Stewart, Whitely, and Twyman, 2004, chap. 4). In Britain, social class has long been designated as the key politically relevant sociological variable (e.g., Pulzer 1969), with the canonical social psychological account of Butler and Stokes (1969) focusing on how class identities are translated into durable partisan self-images (party identification) that largely govern voting behavior. One strand in the rational choice literature focuses on voters' proximities to competing parties in (possibly multidimensional) issue space (e.g., Heath, Jowell, and Curtice 2001), whereas another follows Fiorina (1981) and reinterprets partisanship as a dynamic "running tally" of party performance evaluations regarding the economy and the delivery of various highly valued public services. According to this latter perspective, valence, not position, issues typically dominate the political agenda (Stokes 1963, 1992). Leader images also play a role in this latter model. Based on work in political psychology (e.g., Sniderman, Brody, and Tetlock 1991), it is hypothesized that party leader images are

https://doi.org/10.1093/pan/mpl010 Published online by Cambridge University Press

important heuristic devices for voters confronted with the task of making electoral choices in a political world where stakes are high and uncertainty abounds.

The specific model used here incorporates the effects of feelings about party leaders, judgments regarding which party is best able to handle most important issues, party identification, perceptions of the ideological distance between respondents and major parties, economic evaluations, attitudes toward the Iraq War, tactical (strategic) voting, as well as social class and several other standard demographics. We focus on voting for the governing Labour Party versus any of the opposition parties in the 2005 general election. In equation form, the specification is

$$\begin{aligned}\text{Labour vote} = \text{fn}(&b_0 + b_1\text{BLAIR} + b_2\text{HOW} + b_3\text{KEN} + b_4\text{LBEST} + b_5\text{CBEST} \\ &+ b_6\text{LDBEST} + b_7\text{OBEST} + b_8\text{LPID} + b_9\text{CPID} + b_{10}\text{LDPID} \\ &+ b_{11}\text{OPID} + b_{12}\text{LDIS} + b_{13}\text{CDIS} + b_{14}\text{LDDIS} + b_{15}\text{ECVAL} \\ &+ b_{16}\text{IRAQ} + b_{17}\text{TACT} + \Sigma b_{18} - b_{27}\text{DEMOG}), \end{aligned} \quad (2)$$

where BLAIR, HOW, and KEN refer to respondents' affective orientations on 0–10 scales of the three major party leaders; LBEST, CBEST, and LDBEST refer, respectively, to Labour, the Conservatives, or the Liberal Democrats being the party best able to handle the issue the respondent considers most important; LDIS, CDIS, and LDDIS are three "distance" variables referring to the average of the absolute distances between the Labour, Conservative, and Liberal Democrat parties, respectively, and the respondent on the 0–10 tax-spend and EU scales described in Table 2; LPID, CPID, LDPID, and OPID are 0–1 dummies, respectively, for whether or not a respondent identifies with Labour, the Conservatives, Liberal Democrats, or one of several minor parties; ECVAL is a composite scale that combines personal and national retrospective and prospective economic judgments; TACT is a 0–1 variable denoting if a respondent reports voting tactically; DEMOG includes measures of age, ethnicity, gender, social class, and region; and b_1 through b_{27} are effect coefficients. Additional details regarding the several predictor variables are provided in Appendix 3.

As in the turnout model, we extend the specification in equation (2) by adding a series of interaction terms—one for each of the predictor variables. Coefficients on these interaction variables capture the extent to which the Internet sample effect of each predictor deviates from the equivalent effect in the in-person probability sample. A 0–1 mode effect variable is also included.

Table 4 reports the results of these analyses. Estimation is again by logistic regression, and the first and second columns of the table, respectively, show the estimated effects for the core party choice model using the in-person and Internet samples. The third column shows the extent to which there are significant differences between the two sets of coefficients. Considering, first, the results in columns 1 and 2 of the table, it is clear that leader assessments, judgments of competence on the most important issue, and party identification all have significant, correctly signed effects in both analyses. Thus, respondents in both surveys are more likely to vote Labour if they like Blair and less likely to do so if they like Howard or Kennedy. Respondents are also more likely to vote Labour if they identify with the party and believe it is best able to handle the issue they deem most important. And if they identify with another party or believe another party is best able to deal with their most important issue, they are likely not to vote Labour.

However, as discussed in relation to Table 3, the crucial feature of Table 4 is the contents of column 3. These coefficients indicate the extent to which the effect of each

https://doi.org/10.1093/pan/mpl010 Published online by Cambridge University Press

Table 4 Labour vote models and mode effects, in-person and Internet pre-post election panels

	In-person sample		*Internet sample*		*Interaction effects with pooled data*	
Predictor variables	*B*	*SE*	*B*	*SE*	*B*	*SE*
Blair affect	0.496***	0.044	0.365***	0.024	−0.132**	0.050
Howard affect	−0.179***	0.039	−0.173***	0.025	0.006	0.046
Kennedy affect	−0.303***	0.046	−0.185***	0.026	0.118*	0.053
Party best, most important issue						
Labour	1.020***	0.177	0.910***	0.123	−0.110	0.216
Conservative	−1.038***	0.270	−1.185***	0.195	−0.147	0.333
Liberal Democrat	−0.646*	0.329	−1.274***	0.217	−0.628	0.394
Other party	−0.082	0.440	−0.668***	0.196	−0.586	0.482
Party identification						
Labour	1.333***	0.203	1.114***	0.123	−0.219	0.237
Conservative	−1.270***	0.257	−1.166***	0.216	0.104	0.336
Liberal Democrat	−1.251***	0.260	−0.949***	0.188	0.302	0.321
Other party	−0.614*	0.327	−0.869***	0.209	−0.255	0.388
Party-issue proximity						
Labour	0.102***	0.030	0.057***	0.020	−0.045	0.036
Conservative	−0.043*	0.026	−0.010	0.014	0.033	0.029
Liberal Democrat	−0.089**	0.033	−0.006	0.021	0.083*	0.039
Economic evaluations	0.071	0.084	0.158**	0.061	0.087	0.104
Attitudes to Iraq War	−0.014	0.041	0.045*	0.026	0.059	0.049
Tactical voting	−0.341	0.224	−0.608***	0.120	−0.267	0.253
Age	−0.016***	0.005	0.006*	0.003	0.022***	0.006
Ethnicity (white)	−0.635*	0.290	−0.444	0.284	0.190	0.407
Gender (male)	−0.395**	0.156	−0.423***	0.105	−0.032	0.188
Social class	−0.334*	0.162	−0.428***	0.110	−0.093	0.196
Region						
Southeast	0.008	0.291	0.299	0.214	0.290	0.361
Southwest	0.160	0.364	−0.624**	0.251	−0.785*	0.442
Midlands	0.281	0.298	0.552**	0.224	0.271	0.373
North	0.219	0.300	0.440*	0.216	0.221	0.370
Scotland	−0.136	0.349	−0.048	0.253	0.087	0.431
Wales	−0.191	0.413	0.468*	0.276	0.659	0.497
Constant	0.728	0.465	−0.300	0.403	0.728	0.465
Mode (Internet) dummy					−1.029*	0.615
McFadden R^2	.58		.59		.58	
McKelvey R^2	.76		.76		.76	
Percent correctly classified	87.3		88.6		88.2	
Lambda	.68		.68		.68	
N	2109		4922		7030	

*$p < .05$; **$p < .01$; ***$p < .001$; one-tailed test.

predictor variable differs across the two samples. Out of a total of 27 estimated effects, only five produce statistically significant differences. Two of these relate to the leader affect terms. The effects of liking for Blair and for Kennedy are smaller in the Internet sample (respectively, $b = .365$ and $b = -.185$) than they are in the in-person sample

https://doi.org/10.1093/pan/mpl010 Published online by Cambridge University Press

($b = .496$ and $b = -.303$). The third difference relates to the impact of party-issue proximity. For the in-person data, the variable measuring the distance between the respondent and the Liberal Democrats on the combined issue proximity scale has a significant negative effect on Labour support, whereas in the Internet sample this variable is not significant. The two remaining differences relate to the demographic correlates of party support. The in-person results suggest that older voters are slightly less likely, ceteris paribus, to support Labour, whereas the Internet results suggest that older voters are slightly more likely to do so. Finally, whereas the in-person results show no evidence of regional effects on Labour voting, the Internet results suggest people in the southwest were less likely to cast a Labour ballot than those residing in Greater London (the reference region category).[15]

The interaction effect analyses also caution that there is no consistent evidence to indicate that Internet survey respondents are, in some sense, more politically sophisticated than their in-person survey counterparts. If this were the case, then one might expect that the Internet respondents would rely more heavily on party-issue proximities when making their electoral choices. After all, such proximities are the core variables in utility maximization models of party support (e.g., Downs 1957; Adams, Merrill, and Grofman 2005). Similarly, Internet respondents should also be more likely to vote tactically, making electoral choices by discounting party-issue proximities by estimates of competing parties' chances of winning. Results reported in Table 4 show that, in fact, the effect of tactical voting on Labour support is statistically significant for the Internet respondents but (just) fails ($p \leq .10$) to achieve significance for the in-person group. However, the interaction effect for this variable is not significant. Also, the effects of the issue proximity variables are different from what the Internet sophistication hypothesis suggests. All three of the party-issue proximity variables are significant for the in-person sample and only one of them is for the Internet sample. And only one of the corresponding interaction terms is significant.[16]

Viewing the results more generally, we note that three of the five significant interaction effects shown in column 3 of Table 4 are significant only at the .05 level, even with the huge combined N (7030 cases) for the two surveys. Given the total number of estimated effects in the column 3 model (27 "core" coefficients plus 27 interactions plus constant plus mode effect = 56 coefficients), we would expect at least two (and possibly three) coefficients to be significant at .05 even if none of the effects was actually significant. Viewed in this light, the implications of the results shown in Table 4 are clear. There are almost no important differences in the directions and magnitudes of the effects exhibited by this large set of theoretically interesting predictor variables. Inferences that one would draw—and the story one would tell—regarding the determinants of Labour voting would be virtually identical using both the probability and Internet samples.

Finally, we observe that there is virtually no overall explanatory purchase gained by including the mode-effect interaction variables in the model. For the core model estimated for the in-person and Internet data, the McFadden R^2 values are .58 and .59, respectively

[15]We also conducted a multinomial logit analysis of opposition party voting—Conservative, Liberal Democrat, and miscellaneous other parties, with government party (Labour) voting as the reference category. Although the combined data set is very large ($N = 7030$), the results (available upon request) again reveal that only a small number of mode interaction effects are statistically significant ($p \leq .05$). In addition, inclusion of mode interaction terms does virtually nothing to improve the fit of the model, increasing the McFadden R^2 from .57 to .58 and the percentage of cases correctly classified from .795 to .801. AIC and BIC model selection criteria testify in favor of the less richly parameterized models that do not include the mode interactions.

[16]The Conservative and Liberal Democrat analyses also fail to provide evidence that the Internet sample is, in some sense, more politically sophisticated than the in-person sample. There are no statistically significant interaction effects for the issue proximity variables or tactical voting.

https://doi.org/10.1093/pan/mpl010 Published online by Cambridge University Press

Table 5 Comparative performance of rival party choice models

	McFadden R^2	*McKelvey* R^2	*AIC*[a]	*BIC*[a]
Models estimated using in-person survey data				
Social class	.01	.02	2794.20	2805.51
All demographics	.03	.06	2753.54	2810.08
Economic evaluations	.07	.13	2633.38	2644.69
Issue proximities	.12	.22	2507.63	2530.25
Most important issue	.27	.40	2079.75	2108.02
Party identification	.37	.48	1794.87	1823.14
Leader images	.40	.65	1692.95	1715.56
Composite model	.58	.76	1256.45	1414.76
Models estimated using internet survey data				
Social class	.01	.01	6409.16	6422.16
All demographics	.02	.04	6328.65	6400.17
Economic evaluations	.14	.24	5564.96	5577.97
Issue proximities	.19	.34	5229.46	5255.52
Most important issue	.33	.48	4299.71	4332.29
Party identification	.36	.50	4163.88	4196.45
Leader images	.44	.64	3617.93	3643.94
Composite model	.59	.76	2715.98	2898.40

[a]Smaller values indicate better model performance.

(see Table 4). For the pooled data with mode interaction effects, this statistic equals .58. The alternative McKelvey R^2 is identical (.76) in all three analyses. The percentage of cases correctly classified also is very similar in the three analyses—87.3% and 88.6% for the core in-person and Internet models and 88.2% for the pooled data with interaction terms.

5.3 *Rival Models*

An additional set of tests can be applied to the party choice models presented in Table 4. These models are "composites," i.e., they include predictor variables from a variety of potentially competing (but also potentially complementary) accounts of why people choose to vote for one party rather than another. In effect, the composite models employed in Table 4 had six main predictor components: leader images, party identifications, perceptions of the party able to handle the most important issue, issue proximities, economic evaluations, and demographics. Given the traditional emphasis in U.K. electoral research on the role of social class, we can also divide the demographics category into social class and "other demographics."

Table 5 shows the consequences of specifying and estimating each of these sets of variables as separate models of Labour voting in 2005. The upper part of the table summaries results for the in-person survey. The first row indicates that McFadden R^2 for a model that predicts Labour voting using only social class is a miniscule .01. Using all demographics (age, gender, ethnicity, social class, and region) increases this R^2 only slightly, to .03. A model based exclusively on economic evaluations yields an R^2 of .07, and one based on issue proximities produces an $R^2 = .12$. The R^2 statistics for the most important issue, party identification, and leader affect models indicate that these models have considerably stronger explanatory power. However, the composite model clearly fares best, with an R^2 of .58. The rank order of the models using the alternative McKelvey R^2 is identical. In addition to the two R^2 measures, the table also reports the

https://doi.org/10.1093/pan/mpl010 Published online by Cambridge University Press

Table 6 Comparative predictive power of rival models of Labour voting using in-person survey and Internet survey data

	In-person Survey		*Internet survey*	
Models	*% Correctly predicted*	*Lambda*	*% Correctly predicted*	*Lambda*
Social class	60.5	.00	63.9	.00
All demographics	63.1	.07	64.3	.01
Economic evaluations	65.3	.12	70.4	.18
Issue proximities	67.9	.19	72.3	.23
Most important issue	78.3	.45	80.8	.47
Party identification	83.5	.59	82.6	.52
Leader mages	82.0	.55	83.7	.55
Composite model	87.3	.68	88.6	.68

Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) statistics (Burnham and Anderson 2002). Recalling that lower AIC and BIC values denote better overall model performance, it is clear that, despite its considerably richer parameterization, the composite model outperforms its rivals in analyses of the in-person survey data.

The lower section of Table 5 shows that analysts using the Internet data would reach exactly the same conclusions about the relative importance of the competing models of party choice. By sizable margins, the composite model produces the highest McFadden and McKelvey R^2 values and the lowest AIC and BIC values. And although the BIC and AIC values are quite different in absolute terms across the two survey modes (recall the differences in sample sizes), the composite model R^2 values across the two modes are virtually identical. And as previously noted, the in-person survey yields a McFadden $R^2 = .58$, compared with a figure of .59 for the Internet survey. Both modes produce a McKelvey $R^2 = .76$. In short, both in terms of coefficient signs and magnitudes (Tables 3 and 4) and relative explanatory power (Table 5), the in-person probability and Internet surveys produce near identical results.

It is also useful to compare the predictive power of the several voting models across the two survey modes. Calculating the probability that respondents in the in-person and Internet surveys will vote for a particular party (here the governing Labour Party) provides a convenient and intuitively plausible way to bring together information about the coefficients in various party choice models and the survey data used to estimate these coefficients.[17] If the in-person and Internet data really can tell the same stories about the absolute and relative efficacy of rival models of electoral choice, then equivalent models should have very similar predictive power across the two survey modes. Data reported in Table 6 indicate that the similarities are impressive. There is only one minor deviation in the rank orders of predictive power of the rival models, the average difference in reduction in prediction errors provided by various models (measured by lambda) is only .031, and the average difference in predictive power is only 2.6%. The correlation between the predictive accuracy of equivalent models estimated by the in-person and Internet data sets is fully .99.

Finally, we take the prediction exercise one step further, using the parameters for the composite model of Labour voting estimated using the in-person survey data to predict Labour voting in the Internet sample. Then, we reverse the procedure, using the parameters for the composite model estimated using the Internet survey to predict Labour voting in the

[17]Since the logit functional form is nonlinear, the impact of any predictor variable is a function of its coefficient and value and the coefficients and values of all other predictor variables (see, e.g., Long 1997).

https://doi.org/10.1093/pan/mpl010 Published online by Cambridge University Press

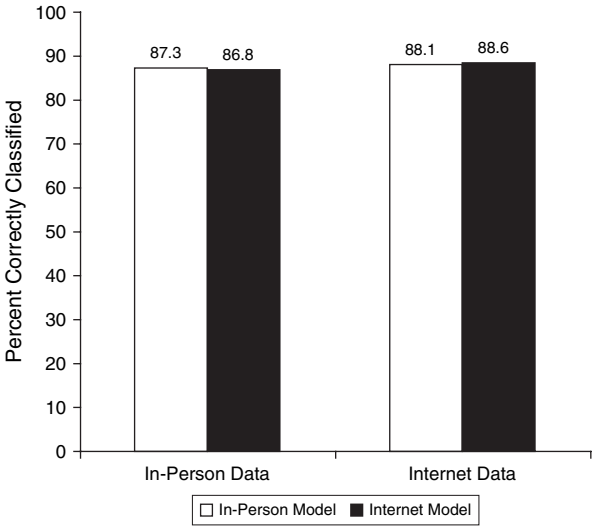


Fig. 5 Cross-predicting Labour voting in the in-person and Internet samples.

in-person sample. Once more, the results clearly testify that the models are telling the same story. As shown in Fig. 5, using the parameters estimated using the in-person data, we can correctly predict 87.3% of the Labour voters in the in-person survey and almost as many, 86.8%, in the Internet survey. Similarly, using parameters generated using the Internet data, we can correctly classify 88.6% of Labour voters in the Internet survey and 88.1% in the in-person survey. In both cases, the cross-survey differences in predictive power verge on rounding error—only 0.5%.

In sum, the analyses presented in Tables 3–6 indicate that the in-person and Internet survey data invite analysts to reach the same conclusions about what mattered most, what mattered less, and what mattered not at all for turnout and party choice in Britain in 2005. Although there are some differences in marginal distributions on key variables across the two data sets, there are strong similarities in the *effects* of most variables and the overall performance of the rival models of electoral participation and party support. These findings corroborate recent work comparing Internet and other survey modes conducted in the United States (Alvarez, Sherman, and VanBeselaere 2003; Berrens et al. 2003).

6 Summary and Conclusions

Concerns about the suitability of Internet surveys for doing national election studies and other major survey research projects are frequently voiced. Like most social scientists educated when in-person surveys were undisputed received wisdom, we instinctively lean toward the face-to-face probability survey as the definitive, most reliable instrument. However, as discussed above, conventional face-to-face surveys with probability samples have a number of drawbacks. Cost and lengthy administration times are often cited as two of them. In addition, it is important to appreciate that these surveys have always suffered potential sampling bias problems because of unit nonresponse, and for political surveys at least, this problem has grown over time. The problem obtains for RDD surveys as well. The more difficult it proves to achieve high response rates in conventional surveys, the more concerned we need to be about the sampling bias problems they may engender. Although cost considerations and other properties of Internet surveys make them an

https://doi.org/10.1093/pan/mpl010 Published online by Cambridge University Press

attractive alternative to in-person and RDD surveys, it is often claimed that sampling is the Internet's Achilles' heel. Intuition suggests that Internet surveys that draw from panels of self-selected willing participants might suffer to an even greater extent from sampling biases than do in-person surveys based on imperfectly implemented probability samples. And similar to in-person and RDD surveys, Internet surveys often encounter high levels of unit nonresponse. But intuitions can mislead, and unfortunately, there is no theorem that can tell us either about the extent of the sampling bias in probability and Internet surveys that derives from nonrandom unit nonresponse or about the bias in Internet surveys of initially self-selected respondents. Questions about similarities and differences in data gathered by various modes can only be resolved by detailed, and repeated, empirical analyses.

This study shows that the marginal distributions on key variables in models of voting behavior differ (although not greatly) across national probability and Internet samples collected in the 2005 BES. In global terms, the Internet sample appeared to be slightly less left leaning than the probability sample. However, it bears emphasis that there is no way of determining whether the probability or the Internet marginal distributions more accurately reflect the "true" views of the British electorate. If we compare the two sets of marginals in terms of vote shares in the 2005 general election, the Internet sample is the more accurate. If we compare actual and reported turnout, the in-person sample yields a closer estimate. However, similar to other election studies conducted in Britain and elsewhere, both surveys are wide of the mark by sizable margins. Pre- and postelection shares of party identifiers in the in-person and Internet surveys are virtually identical.

More important, in our view, the in-person and Internet surveys yield remarkably similar results when it comes to estimating parameters in voting behavior models. We tested explicitly to see if predictors of turnout and of party choice differed in the effects they produced across the in-person and Internet data sets. Overwhelmingly, they did not. With few minor exceptions, the estimated effects of a large number of predictor variables were statistically indistinguishable across the two data sets. And the relative explanatory power of rival models was exactly the same. Analyses of the predictive power of competing models of party choice also yielded impressive similarities. These several findings prompt the conclusion that, by using high-quality Internet surveys, students of British voting behavior are unlikely to be misled about the effects of different variables on turnout and party choice.[18]

That said, it is clear that further research on mode effects in national election studies is necessary. Perhaps most important is learning if the British findings presented here "travel well." Studies conducted in various countries comparing the results of carefully executed Internet surveys with those produced by traditional modes of survey data collection should have high priority. Such studies should focus on inferences about models of turnout and party choice and various (not unrelated) topics concerning the measurement of key variables in such models.

The 2005 BES may provide a useful model for how such parallel surveys might be conducted. At this stage in the development of knowledge about the utility of Internet surveys for conducting national election studies in various countries, we recommend that the Internet be seen as a supplement—not a substitute—for traditional methods of data collection. And as just observed, generalizations should be made with caution. What works well in Britain, may not prove useful elsewhere—or, at least, not yet. That said, when selecting an Internet survey firm, we chose an experienced house that had established a strong track record for getting vote shares very close to target in several elections held

[18]This conclusion is bolstered by the results of an earlier, preliminary, study of mode differences conducted using 2001 BES data (Clarke, Sanders, Stewart, Whiteley, and Twyman 2004).

https://doi.org/10.1093/pan/mpl010 Published online by Cambridge University Press

over the past half-decade. This firm also had the capacity to do the large *N* baseline survey that was crucial for the success of a rolling campaign panel survey design. In addition, the firm had the ability to conduct a variety of novel, technologically sophisticated experiments. Yet another consideration was cost. As in the past, the principal item in the 2005 BES budget was the in-person surveys. These surveys were a mandated priority, and any supplementary Internet surveys—regardless of the value for money they might provide—had to abide by the budget constraint entailed by that priority. Future mode comparison research in national election studies conducted in other countries likely will face similar constraints. Hopefully, this will not deter principal investigators from including Internet components in their research designs.

The 2005 BES results indicate that learning more about how to utilize Internet technologies in the service of national election studies in various political settings is a worthwhile enterprise.

Appendix 1: Selected Demographic Characteristics of 2005 BES Internet and In-Person Respondents—Preelection Surveys

	Unweighted		*Weighted*	
	Internet	*In-person*	*Internet*	*In-person*
Gender				
Male	49.4	44.7	48.2	48.2
Female	50.6	55.3	51.8	51.8
N	7793	3589	7862	3589
Age				
18–25	9.4	8.3	13.1	13.7
26–35	21.9	16.3	19.0	16.9
36–45	19.8	20.1	18.3	19.6
46–55	16.1	17.3	15.5	16.6
56–65	24.9	15.5	26.0	14.3
66 and over	7.9	22.6	8.1	18.8
Mean age	45.4	49.9	45.2	47.2
N	7777	3576	7843	3571
Self-described ethnicity				
White British	95.9	91.5	95.8	88.6
All other	4.1	8.5	4.2	11.4
N	7793	3381	7862	3579
Region				
Southeast	24.3	17.2	26.9	24.9
Southwest	9.9	6.7	11.4	9.0
Midlands	15.8	10.6	17.1	17.3
North	14.5	15.4	13.3	26.0
London	21.1	6.2	16.9	9.0
Scotland[a]	9.0	26.0	8.5	8.8
Wales[a]	5.4	17.9	5.9	5.1
N	7793	3589	7862	3589
Age completed full-time education				
15 or younger	16.0	34.0	17.0	30.3
16	24.8	27.0	25.3	26.4

Continued

https://doi.org/10.1093/pan/mpl010 Published online by Cambridge University Press

Appendix 1: (continued)

	Unweighted		*Weighted*	
	Internet	*In-person*	*Internet*	*In-person*
17	9.5	9.1	9.7	8.7
18	11.9	8.0	12.1	8.8
19 or older	33.1	19.7	29.5	21.5
Still in school	4.7	2.2	6.4	4.4
N	7773	3583	7839	3579
Annual family income (in £)				
0–10,000	11.2	22.0	12.2	17.2
10,000–20,000	27.1	26.9	28.6	24.1
20,000–30,000	24.5	17.5	24.5	18.1
30,000–40,000	16.0	12.8	15.7	15.0
40,000–50,000	9.3	9.0	8.7	11.9
50,000–60,000	4.8	4.8	4.3	4.9
60,000–70,000	2.5	2.8	2.3	3.6
More than 70,000	4.5	4.3	3.8	5.6
N	6925	3118	6929	3056

[a]Deliberate oversampling of Scotland and Wales in in-person survey.

Appendix 2: Question Wordings for Variables in Turnout Model

All starred (*) items are Likert five-point scale questions: Please tell me how far you agree or disagree with ... the following statement.... (Response options: strongly agree, agree, neither agree nor disagree, disagree, strongly disagree, do not know, refused)

Political influence

- On a scale from 0 to 10, where 10 means a great deal of influence and 0 means no influence, how much influence do *you* have on politics and public affairs?

Feelings about parties

- On a scale that runs from 0 to 10, where 0 means strongly dislike and 10 means strongly like, how do feel about the (Labour/Conservative/Liberal Democrat) party?

Personal benefits*

- Being active in politics is a good way to get benefits for me and my family.
- I would feel very guilty if I didn't vote in a general election.
- I feel a sense of satisfaction when I vote.

Costs of voting*

- It takes too much time and effort to be active in politics and public affairs.
- People are so busy that they don't have time to vote.

Civic duty*

- I would be *seriously* neglecting my duty as a citizen if I didn't vote.
- It is every citizen's duty to vote in an election.

https://doi.org/10.1093/pan/mpl010 Published online by Cambridge University Press

Democracy dissatisfaction

- On the whole, are you satisfied or dissatisfied with the way that democracy works in this country? (Response options: very satisfied, fairly satisfied, a little dissatisfied, very dissatisfied, do not know, refused)

Election interest

- How interested were you in the general election that was held on May 5 this year? (Response options: very interested, somewhat interested, not very interested, not at all interested, do not know, refused)

Party mobilization

- Did a canvasser from any party call at your home to talk with you during the election campaign?
- Did anyone from a political party telephone you during the election campaign to ask you how you might vote?

Relative deprivation*

- The Government generally treats people like me fairly.
- There is often a big gap between what people like me expect out of life and what we actually get.

Social norms*

- Most of my family and friends think that voting is a waste of time.

Social trust

- On balance, would you say that most people can't be trusted or that most people can be trusted? Please use the 0 to 10 scale to indicate your view. (Please take your answers from this card.)

Age

- Age in years.

Education

- Six point scale based on formal qualifications.

Ethnicity

- White = 1, not = 0.

Gender

- Male = 1, female = 0.

Social class

- Six point scale based on Registrar General 6-category Occupational scale

Region

- Scotland, Wales, southeast, southwest, Midlands, north

https://doi.org/10.1093/pan/mpl010 Published online by Cambridge University Press

Appendix 3: Question Wordings for Variables in Party Choice Model

Party leader affect

- Now, let's think about party leaders for a moment. Using a scale that runs from 0 to 10, where 0 means strongly dislike and 10 means strongly like, how do you feel about . . . (Tony Blair)?

Party best, most important issue

- Now, I'd like to ask you a few questions about the issues and problems facing Britain today. As far as you're concerned, what is the *single most important issue* facing the country at the present time? (Note: this is an open-ended question.)
- Which party is best able to handle this issue? (Note: this is an open-ended question)

Party identification

- Generally speaking, do you think of yourself as Labour, Conservative, Liberal Democrat (Scottish National/Plaid Cymru), or what?
- (If none, do not know, or refused in previous question) Do you generally think of yourself as a little closer to one of the parties than the others?

Party-issue proximities

Tax-spend scale

- Using the 0 to 10 scale on this card, where the end marked 0 means that government should *cut taxes and spend much less on health and social services* and the end marked 10 means that government should *raise taxes a lot and spend much more on health and social services*, where would you place (yourself, the Labour Party, the Conservative Party, the Liberal Democrat Party) on this scale?

EU scale

- Now, we would like your opinion about Britain's membership in the EU. On the scale shown on this card, 0 means that Britain should definitely get out of the EU and 10 means that Britain should definitely stay in the EU. Where would you place (yourself, the Labour Party, the Conservative Party, the Liberal Democrat Party) on this scale?

Economic evaluations

- How does the *financial situation of your household* now compare with what it was *12 months ago*? (Response options: got a lot worse, got a little worse, stayed the same, got a little better, got a lot better, do not know, refused)
- How do you think the *general economic situation in this country* has changed over the *last 12 months*? (Response options: got a lot worse, got a little worse, stayed the same, got a little better, got a lot better, do not know, refused)
- How do you think the *financial situation of your household* will change over the *next 12 months*? (Response options: get a lot worse, get a little worse, stay the same, get a little better, get a lot better, do not know, refused)
- How do you think the *general economic situation in this country* will develop over the *next 12 months*? (Response options: get a lot worse, get a little worse, stay the same, get a little better, get a lot better, do not know, refused)

https://doi.org/10.1093/pan/mpl010 Published online by Cambridge University Press

Iraq War

- Please tell me whether you strongly approve, approve, disapprove, or strongly disapprove of *Britain's involvement* in Iraq. (Response options: strongly approve, approve, disapprove, strongly disapprove, do not know, refused)
- How well do you think the present government has handled the situation in Iraq? (Response options: very well, fairly well, neither well nor badly, fairly badly, very badly, do not know, refused)

Voted tactically

- People give different reasons for why they vote for one party rather than another. Which of the following best describes your reasons? Please take your answers from this card. Response options:

  (a) The party had the best policies
  (b) The party had the best leader
  (c) I really preferred another party but it stood no chance of winning in my constituency.
  (d) I voted tactically (volunteered)
  (e) Other (write in)
  (f) Do not know
  (g) Refused.

  Tactical voters defined as those who responded to option c or d.

References

Adams, James F., Samuel Merrill III, and Bernard Grofman. 2005. *A unified theory of party competition*. Cambridge: Cambridge University Press.

Alvarez, R. Michael, Robert P. Sherman, and Carla VanBeselaere. 2003. Subject acquisition for web-based surveys. *Political Analysis* 11:23–43.

Baker, Ken, John Curtice, and Nick Sparrow. 2002. Internet poll trial: Research report. ICM Research: Unpublished paper.

Berrens, R. P., Alok K. Bohara, Hank Jenkins-Smith, Carol Silva, and David L. Weimer. 2003. The advent of Internet surveys for political research: A comparison of telephone and Internet samples. *Political Analysis* 11:1–22.

Brehm, John. 1993. *The phantom respondents: Opinion surveys and political representation*. Ann Arbor: University of Michigan Press.

Burnham, Kenneth P., and David R. Anderson. 2002. *Model selection and multimodel inference*. 2nd ed. New York: Springer.

Butler, David, and Donald E. Stokes. 1969. *Political change in Britain*. New York: St Martin's.

Chang, Lin Chiat, and Jon A. Krosnick. 2003. *Comparing oral interviewing with self-administered computerized questionnaires: An experiment*. Columbus, OH: Department of Psychology, The Ohio State University.

Clarke, Harold D., David Sanders, Marianne C. Stewart, and Paul Whiteley. 2004. *Political choice in Britain*. Oxford: Oxford University Press.

———. 2006. Taking the bloom off New Labour's rose: Party choice and voting turnout in Britain, 2005. *Journal of Elections, Public Opinion, and Parties* 16:3–36.

Clarke, Harold D., David Sanders, Marianne C. Stewart, Paul F. Whiteley, and Joe Twyman. 2004. The 2001 British election study Internet poll: A methodological experiment. *Journal of Political Marketing* 3:29–55.

Couper, Mick P. 2000. Web surveys: A review of issues and approaches. *Public Opinion Quarterly* 64:464–94.

Couper, Mick P., Roger Tourangeau, and Kristin Kenyon. 2004. Picture this! Exploring visual effects in web surveys. *Public Opinion Quarterly* 68:255–66.

Curtin, Richard, Stanley Presser, and Eleanor Singer. 2000. The effects of response rate changes on the index of consumer sentiment. *Public Opinion Quarterly* 64:413–28.

———. 2005. Changes in telephone survey nonresponse over the past quarter century. *Public Opinion Quarterly* 69:87–98.

https://doi.org/10.1093/pan/mpl010 Published online by Cambridge University Press

Dalton, Russell. 2002. *Citizen politics: Public opinion and political parties in advanced industrial democracies*. 3rd ed. Chatham, NJ: Chatham House Publishers.

Dillman, Don A. 2000. *Mail and Internet surveys: The tailored design method*. New York: John Wiley.

Downs, Anthony. 1957. *An economic theory of democracy*. New York: Harper and Row.

Fiorina, Morris. 1981. *Retrospective voting in American national elections*. New Haven, CT: Yale University Press.

Groves, Robert M., and Mick P. Couper. 1998. *Nonresponse in household interview surveys*. New York: John Wiley.

Heath, Anthony, Roger Jowell, and John Curtice. 2001. *Party policies and voter choices*. Oxford: Oxford University Press.

Holbrook, Allyson L, Melanie C. Green, and Jon A. Krosnick. 2003. Telephone versus face-to-face interviewing of national probability samples with long questionnaires: Comparisons of respondent satisficing and social desirability. *Public Opinion Quarterly* 67:79–125.

Institute for Social Research. n.d. *The 2000 Canadian election survey: Technical documentation*. Toronto, ON: York University.

Johnston, Richard, and Henry E. Brady. 2002. The rolling cross-section design. In *The future of election studies*, ed. Mark Franklin and Christopher Wlezien, 123–36. Amsterdam, The Netherlands: Pergamon.

Kaplowitz, Michael D., Timothy D. Hadlock, and Ralph Levine. 2004. A comparison of web and mail survey response rates. *Public Opinion Quarterly* 68:94–101.

Kavanagh, Dennis, and David Butler. 2005. *The British general election of 2005*. London: Palgrave Macmillan.

Keeter, Scott, Carolyn Miller, Andrew Kohut, Robert M. Groves, and Stanley Presser. 2000. Consequences of reducing nonresponse on a national telephone survey. *Public Opinion Quarterly* 64:125–48.

Krosnick, Jon A., and Lin Chiat Chang. 2001. A comparison of the RDD telephone survey methodology with Internet survey methodology as implemented by Knowledge Networks and Harris Interactive. Columbus, OH: Department of Psychology, The Ohio State University.

Long, J. Scott. 1997. *Regression models for categorical and limited dependent variables*. Thousand Oaks, CA: Sage Publications.

Pulzer, Peter. 1969. *Political representation and elections in Britain*. London: Allen and Unwin.

Putnam, Robert. 1993. *Making democracy work: Civic traditions in modern Italy*. Princeton, NJ: Princeton University Press.

Rivers, Douglas. 2006. *Sample matching: Representative sampling from Internet panels*. Palo Alto, CA: Polimetrix.

Runciman, W. G. 1966. *Relative deprivation and social justice*. Berkeley, CA: University of California Press.

Schoen, Harald, and Thorsten Faas. 2005. When methodology interferes with substance: The difference of attitudes towards e-campaigning and e-voting in online and offline surveys. *Social Science Computer Review* 23:326–33.

Smith, Tom W. 1995. Trends in nonresponse rates. *International Journal of Public Opinion Research* 7:157–71.

Sniderman, Paul, Richard Brody, and Philip E. Tetlock, eds. 1991. *Reasoning and choice: Explorations in political psychology*. Cambridge: Cambridge University Press.

Steeh, Charlotte. 1981. Trends in nonresponse rates, 1952–1979. *Public Opinion Quarterly* 59:66–77.

Stokes, Donald E. 1963. Spatial models of party competition. *American Political Science Review* 57:368–77.

———. 1992. Valence politics. In *Electoral politics*, ed. Dennis Kavanagh. Oxford: Oxford University Press.

Taylor, Humphrey, John Brenner, Gary Overmeyer, Jonathan W. Siegel, and George Terhanian. 2001. Touchdown! Online polling scores big in November 2000. *Public Perspective* (March/April):38–9.

Tourangeau, Roger, Lance J. Rips, and Kenneth Rasinski. 2000. *The psychology of survey response*. Cambridge: Cambridge University Press.

VanBeselaere, Carla. 2002. *Survey response quality: Shirking behavior in Internet and telephone surveys*. Pasadena, CA: California Institute of Technology, School of Social Sciences.

Verba, Sidney, Kay Lehman Schlozman, and Henry Brady. 1995. *Voice and equality: Civic voluntarism in American politics*. Cambridge, MA: Harvard University Press.

Whiteley, Paul, and Patrick Seyd. 2002. *High intensity participation—The dynamics of party activism in Britain*. Ann Arbor: University of Michigan Press.

Whiteley, Paul, Marianne C. Stewart, David Sanders, and Harold D. Clarke. 2006. The issue agenda and voting in 2005. In *Britain votes 2005*, ed. Pippa Norris and Christopher Wlezien, 146–61. Oxford: Oxford University Press.

https://doi.org/10.1093/pan/mpl010 Published online by Cambridge University Press