**Mail Surveys for Election Forecasting? An Evaluation of the Columbus Dispatch Poll**

Penny S. Visser, Jon A. Krosnick, Jesse Marquette, Michael Curtin

*Public Opinion Quarterly*, Volume 60, Issue 2 (Summer, 1996), 181-227.

Stable URL:
http://links.jstor.org/sici?sici=0033-362X%28199622%2960%3A2%3C181%3AMSFEFA%3E2.0.CO%3B2-G

# MAIL SURVEYS FOR ELECTION FORECASTING?
## AN EVALUATION OF THE *COLUMBUS DISPATCH* POLL

PENNY S. VISSER
JON A. KROSNICK
JESSE MARQUETTE
MICHAEL CURTIN

**Abstract** Because of slow turnaround time and typically low response rates, mail surveys have generally been considered of little value in election forecasting. However, statewide mail surveys conducted by the *Columbus Dispatch* newspaper since 1980 have made remarkably accurate forecasts of Ohio election outcomes. In comparison to statewide surveys by two other organizations employing conventional telephone interview methods, the mail surveys were consistently more accurate and were generally less susceptible to sources of inaccuracy such as high roll-off and low publicity. The mail survey's advantage is attributable at least in part to larger sample sizes, sampling and response procedures that yielded more representative samples of voters, lack of the need to allocate undecided respondents, and superior questionnaire design. These findings suggest that mail surveys not only may be viable alternatives to telephone surveys but may actually be superior to them under some conditions. Further-

PENNY S. VISSER is a doctoral candidate in the department of psychology at Ohio State University and JON A. KROSNICK is an associate professor in the departments of psychology and political science at Ohio State University. JESSE MARQUETTE is a professor of political science and director of the Survey Research Center at the University of Akron. MICHAEL CURTIN is the editor of the *Columbus Dispatch*. This research was supported by grant SBR-9503822 from the National Science Foundation to Jon A. Krosnick and by grant T32-MH19728-03 from the National Institute of Mental Health, which provided a predoctoral fellowship to Penny S. Visser. We thank Andy Smith and Al Tuchfarber from the University of Cincinnati for providing information on their preelection surveys, and Jim Hunter from the *Columbus Dispatch* for archival assistance. We also thank Richard Carson, Robert Cameron Mitchell, and Stanley Presser for helpful comments on an earlier draft. Correspondence regarding this article should be addressed to Penny S. Visser or Jon A. Krosnick, Department of Psychology, Ohio State University, 1885 Neil Avenue, Columbus, OH 43210 (E-mail visser.104@osu.edu or krosnick@osu.edu).

Public Opinion Quarterly Volume 60:181–227  0033-362X/96/6002-0015$02.50

more, these results demonstrate that surveys with low response rates are not necessarily low in validity.

Since their inauspicious beginnings over 150 years ago (Smith 1990), preelection forecasting surveys have undergone tremendous refinement (Field 1983; Frankovic 1992; Gallup 1972; Mann and Orren 1992; Perry 1960, 1979; Rosenstone 1983). Early surveys, such as the straw polls of the *Literary Digest* magazine, were characterized by haphazard sampling methods that precluded consistently accurate forecasts (Converse 1987; Field 1983; Gallup 1972; Rosenstone 1983; Squire 1988). But by the mid-1930s, Gallup, Roper, and Crossley had begun using more scientific sampling techniques to ensure that their respondents accurately represented the electorate. These and other improvements allowed Gallup to forecast the U.S. presidential elections from 1936 to 1956 with an average error of 3.9% (Rosenstone 1983). Further methodological changes in the 1950s and 1960s (Perry 1960, 1979) brought the Gallup forecasts of the national elections from 1960 through 1980 to within an average of 1.6% of the actual election results (Rosenstone 1983).

Perhaps the most sweeping change made by Gallup and other preelection survey outfits has been the shift from face-to-face interviewing to telephone interviewing. Of the 430 preelection surveys examined in one recent study, fully 98% interviewed by telephone (Crespi 1988). Telephone interviews are less expensive to conduct than face-to-face interviews, and the former typically achieve higher response rates than most mail surveys, leaving them less vulnerable to nonresponse bias. Furthermore, telephone interviewing can be completed much more quickly than face-to-face and mail surveys, which permits up-to-the-minute tracking of candidate support.

Despite these presumed improvements in procedures, however, preelection surveys still predict some election results with significant error. In 1980, for example, all four of the major polls underestimated Ronald Reagan's margin of victory over Jimmy Carter, in one case by as much as 9% (Kagay 1992). Similarly, in the 1970 British election, four of the five final polls and 19 out of 20 earlier polls placed the Labor Party in the lead (by an average of 4.3%), whereas the Conservative Party won by 2.4% (Crewe 1992). More recently, last-minute forecasts of the 1993 New Jersey governor's race showed incumbent Jim Florio leading challenger Christine Todd Whitman by as much as 15%, but on Election Day Florio was unseated (Gray 1993). Of course, even in a perfectly executed survey, one would expect some discrepancy between the election forecast and the actual outcome due to sampling error, but the systematic discrepancies cited above cannot be accounted for by sampling error alone. Clearly, there is room for further improvement of preelection survey methodology.

Less clear, though, is how such improvements can best be achieved. In this article, we propose the consideration of self-administered mail surveys for use in election forecasting. We suggest that mail survey methodology, along with scientific sampling techniques and careful questionnaire design, may provide uniquely desirable conditions for accurately forecasting election outcomes.

We begin by detailing the methods of preelection mail surveys conducted by the *Columbus Dispatch* newspaper in Ohio since 1980 and preelection telephone surveys by the University of Cincinnati and the University of Akron. We then compare the accuracy of the polls in forecasting the national and statewide elections every even year from 1980 to 1994 and explore the correlates of forecast accuracy. Finally, we test several possible explanations for the observed differences between the two survey modes.

## Methods

### *COLUMBUS DISPATCH*

The *Dispatch* has drawn its samples from lists of registered Ohio voters obtained from the state board of elections roughly 3 months before each election. From 1980 to 1994, county election boards in Ohio updated the statewide list of registered voters every 6 months, and the various county boards did so at different times. There was virtually no lag between the time the county boards submitted their updates and the time the updates were incorporated into the statewide list. So at the time the *Dispatch* purchased a particular statewide list, some portions of the list could have been as much as 6 months old, whereas other portions may have been updated more recently. In addition, eligible adults can register to vote for 2 months after the *Dispatch* has purchased the list for a particular election. As a result, people who registered to vote in the months immediately preceding a given election were not included in the *Dispatch* sampling frame, nor were people who moved without registering their new addresses with the county election board.

For sampling, the lists of registered voters were stratified into six regions. Prior to 1992, the regional subgroups were (1) the city of Cleveland, (2) the suburbs of Cleveland, (3) Franklin County, including Columbus, (4) Hamilton County, including Cincinnati, (5) the nine next largest metropolitan counties, Butler, Lake, Lorain, Lucas, Mahoning, Montgomery, Stark, Summit, and Trumbull, and (6) the 76 remaining, mostly rural counties. Registered voters were selected from each of these regions in proportions corresponding to the expected turnout in

each region. Expected regional turnout was estimated on the basis of two factors: the number of people registered to vote in each region and the percent of the total statewide vote to have come from each region in the last two statewide elections.

To permit describing candidate preferences for each of the five dominant urban media markets in Ohio, the *Dispatch* changed its stratification technique in 1992. The 1992 and 1994 lists of registered voters were stratified into those five regions (Cincinnati, Cleveland, Columbus, Dayton, and Toledo) plus a southeastern rural Appalachian region. As in past years, the proportion of registered voters selected from each region corresponded to the percent of the total statewide vote to have come from that region in the two previous statewide elections and the number of registered voters in the region.

Questionnaire recipients were selected sequentially from the list of registered voters in each region with varying frequency (e.g., every twenty-fifth voter) depending on the proportion of the overall sample to be drawn from that region. Predominantly black wards of Cleveland were oversampled by a factor of 3, and predominantly black wards of Columbus were oversampled by a factor of 2 in an effort to offset higher nonresponse rates in those wards. These ratios were based on early *Dispatch* experiences that suggested the rate of oversampling needed to produce final numbers of respondents from these wards comparable to the number of respondents in other wards. Oversampling reduced sampling error by increasing the number of respondents from predominantly black wards, but it did not eliminate any bias due to systematic tendencies for some types of people not to participate.

Each selected person was mailed three items: a cover letter, a questionnaire, and a postage-paid return envelope. The cover letter briefly described the survey and encouraged recipients to complete and return the anonymous questionnaire. The letter was printed on *Dispatch* letterhead and was signed by the public affairs editor.

The questionnaire (for an example, see fig. 1) was designed to simulate the actual presentation of voting options on Election Day in terms of race order, party affiliation information, and language. Respondents were asked to indicate their choices by placing X's in the appropriate boxes. The candidates' names were presented to *Dispatch* poll respondents in alphabetical order. All participants received the identical questionnaire printed on paper of six different colors to represent the six regions.

In addition to questions about candidate preferences in the current election, respondents were asked about their intent to vote in the election, party affiliation, voting behavior in prior elections, education, age, gender, race, religion, union membership, and annual income.

The packets were sent out approximately 12 days prior to Election Day without advance notice or follow-up mailings. They were usually mailed on a Thursday, and completed questionnaires were accepted until the Saturday prior to the election. The results were published in the *Columbus Dispatch* the next day.

The number of packets mailed out varied from year to year, but the average for each election between 1980 and 1994 was approximately 8,000. Roughly 12% of the packets were undeliverable each year. Presumably, some of these undeliverable packets were intended for people who were no longer eligible to participate in the election (e.g., those who had moved out of state) and were thus correctly excluded from the sampling frame. Others, however, may have been intended for people who were incorrectly excluded from the sampling frame (e.g., people who moved within Ohio and registered to vote in their new precincts). When undeliverable packets are subtracted from the total sent out, the average number of packets presumed to have been delivered each year was a little over 7,000.

Between 1980 and 1994, the *Dispatch* response rates ranged from 21% to 28%, with an average of 25%. Consequently, the *Dispatch* forecasts were based on roughly 1,600 responses for each election. All returned ballots were used in the *Dispatch* forecasts, including the very small proportion (roughly 1%) that indicated that the respondent did not plan to participate in the upcoming election.

Responses were weighted before the final forecasts were generated. Although the initial stratified sampling ensured that the questionnaires were distributed to each region in roughly accurate proportions, questionnaires were often returned in slightly disproportionate percentages. Weighting the data by region allowed the *Dispatch* to reestablish the desired proportions. Weighting was also sometimes done based upon the demographic characteristics and party affiliations of respondents if the sample did not appropriately reflect what the *Dispatch* expected to be the characteristics of participating voters.[1] Such weighting was minimal, however, and did not substantially alter the *Dispatch*'s election forecasts.

1. This weighting was done when a sizable discrepancy appeared between the demographic characteristics of the survey respondents and the demographic characteristics of the residents of Ohio (as determined by the U.S. Census). When such discrepancies occurred and the U.S. Census Bureau's biannual publication on voting and registration indicated that the demographic groups involved were likely to vote at different rates, weighting was done to compensate. A similar procedure was implemented by comparing the reported party affiliations of respondents to data on the party affiliations of Ohio voters as indicated by their voter registration status. More detail on the weighting procedures used is available from us; however, as we shall describe shortly, this weighting does not seem to have helped the surveys' accuracy.

**The Columbus Dispatch Election Poll**

Please indicate your choice by placing an X in the appropriate box.

**President and vice president of the United States:**

(1)

George Bush and Dan Quayle (Republican)..... ☐

Bill Clinton and Al Gore (Democratic)..... ☐

Ross Perot and James Stockdale..... ☐

**U.S. Senate:**

(2)

Mike DeWine (Republican)..... ☐

John Glenn (Democratic)..... ☐

**State Issue 1: Proposed Constitutional Convention**

Shall there be a convention to revise, alter or amend the Constitution of Ohio?

(3)

Yes..... ☐ No..... ☐

**State Issue 2: Proposed Constitutional Amendment**

Should the Ohio Constitution be amended so that no person shall hold the office of United States Senator from Ohio for a period longer than two successive terms of six years, and that no person shall hold the office of United States Representative from Ohio for a period longer than four successive terms of two years?

(4)

Yes..... ☐ No..... ☐

**Figure 1.** Example of the 1992 *Columbus Dispatch* mail survey questionnaire.

UNIVERSITY OF AKRON

The University of Akron preelection surveys have been conducted by telephone. The samples were drawn by using random combinations of digits added to selected exchanges, so both listed and unlisted phone numbers were included in the sampling frame. Phone numbers were called up to seven times before being retired from the sampling frame. Since 1988, the adult household resident who had the most recent birthday was selected to participate in the survey. In surveys conducted before 1988, participants were selected using the Troldahl-

The following questions will be used for statistical purposes only. Please mark the box that best describes you:

**Do you intend to vote November 3?** (5)

Yes.....
No.....

**Do you consider yourself a:** (6)

Democrat.....
Republican.....
No Party.....

**For whom did you vote in 1988 for president?** (7)

George Bush.....
Michael S. Dukakis.....
Did Not Vote.....

**For whom did you vote in 1990 for governor?** (8)

Anthony J. Celebrezze Jr....
George V. Voinovich.....
Did Not Vote.....

**This ballot must be completed and in the mail by October 29 for your vote to count.**

**Education:** (9)

College graduate.....
Some college.....
High school graduate....
Some high school.....
Did not attend high school.....

**Age:** (10)

18-24..... 45-54....
25-34..... 55-64....
35-44..... 65-74....
75+.....

**Sex:** (11)

Male..... Female...

**Race:** (12)

White..... Asian....
Black..... American Indian..
Hispanic... Other....

**Religion:** (13)

Baptist.... Jewish...
Protestant. None.....
Catholic... Other....

**Does a union member live in your household?** (14)

Yes..... No.....

**Annual Income:** (15)

Less than $10,000.....
$10,001 to $20,000.....
$20,001 to $30,000.....
$30,001 to $40,000.....
$40,001 to $50,000.....
$50,001 to $60,000.....
$60,001 to $70,000.....
$70,001 or more.....

**Figure 1.** (*Continued*)

Carter procedure (Troldahl and Carter 1964). Those who were selected but chose not to take part in the survey were called a second time by a senior interviewer in an effort to persuade them to reconsider. The response rates for the surveys conducted by the University of Akron between 1980 and 1994 have consistently been about 60%. The Akron forecasts for these years were based on interviews with roughly 800 participants.

The surveys were usually conducted over 5–10-day periods concluding 3–6 days prior to the election. Participants were first asked 10–23 preliminary questions on topics such as interest in the election, presidential and gubernatorial evaluations, the economy, respondents' party identification and political ideology, and respondents' past voting behavior. For each race, the candidates' names and party affiliations were presented, and the order in which the candidate names were presented was randomly rotated across respondents to avoid order effects. The wordings of candidate preference questions changed from year to year, asking which candidate in a particular race a respondent would vote for, which candidate he or she supported, or which candidate he or she would like to see win.

Response options also changed from year to year. Prior to 1990, only the candidates in the particular race were explicitly offered as response options. In 1990 and 1992, however, respondents were invited to indicate that they had not yet decided which candidate they would vote for, and in 1994, they were invited to indicate that they would skip a particular race. For 12 of the 32 races examined between 1980 and 1994, respondents who indicated that they were undecided were asked which candidate they leaned toward, and those leaning toward a candidate were recorded as supporters of him or her.

For most elections, the University of Akron published one set of results gathered from respondents who identified themselves as being registered to vote and a separate set of results for those who reported that they were registered to vote, were very interested in the election, and definitely intended to vote (i.e., likely voters). Because the survey results based on responses from likely voters were presumably more accurate, we focused our evaluations on these results.

### UNIVERSITY OF CINCINNATI

The University of Cincinnati has conducted its preelection surveys in much the same way as the University of Akron. Interviews were completed by telephone, and households were selected through random-digit dialing. Unanswered calls were retried at least six times and sometimes as many as 10 times. Within households, the University of Cincinnati used the most-recent-birthday method to select survey

participants. Individuals who refused to participate were contacted by a senior interviewer who tried to persuade them to reconsider. The response rate for the surveys conducted in the 1990s was about 66%; although exact response rates are unavailable for the 1980s, the University of Cincinnati estimates them to be close to 70% (A. Smith, personal communication, November 1993). The final number of completed interviews ranged from 500 to 900 across the years.

The University of Cincinnati interviews were generally conducted over 4–5-day periods concluding the Sunday before the election. Respondents were asked about their registration status, their interest in the election, and their intent to vote before being asked about their candidate preferences.

Like the University of Akron, the University of Cincinnati varied the wording of its candidate preference questions over the years. In some years, respondents were asked to indicate which candidate they would vote for in each race, but in other years, they were asked which candidate they were more likely to vote for, or which candidate they would like to see win. Respondents were not explicitly offered an "undecided" response. Prior to 1994, respondents who said they were undecided were asked which candidate they leaned toward, and those leaning toward one candidate or another were recorded as supporters of that candidate. In 1994, respondents who volunteered that they were undecided were asked, "Suppose you had to vote right this minute, do you think you would decide *not* to vote for (*race*), that you would vote for (*candidate A*), or that you would vote for (*candidate B*)?" As in the Akron surveys, the order in which the candidates' names were presented was rotated randomly across respondents.

The University of Cincinnati's election forecasts were based only on respondents who reported that they were registered to vote and would definitely vote or would vote unless an emergency arose. In some cases, additional factors, such as interest in the current election or past voting behavior, were also taken into consideration in the selection of likely voters.

## Results

### ACCURACY

Between 1980 and 1994, election forecasts based on the *Columbus Dispatch* preelection surveys were remarkably accurate. Forecasts for the 32 statewide candidate races deviated from the actual election results by an average of 1.6% (see table 1, col. 1), ranging from an

**Table 1.** *Columbus Dispatch* Forecast Accuracy Relative to Election Results, Other Forecasts, and Correlates of Accuracy

| Year and Race | Average Error between *Dispatch* Forecast and Actual Results | Average Error between Akron Forecast and Actual Results | Average Error between Cincinnati Forecast and Actual Results | Characteristics of Races, Elections, and Surveys | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Roll-Off % | Average Number of Newspaper Stories | Race Order on Ballot | Margin of Victory | Voter Turnout | Days between Akron Survey and Election | Days between Cincinnati Survey and Election |
| 1980: | | | | | | | | | | |
| President | 1.4 | 1.8 | · · · | .02 | 84.5 | 3 | 10.6 | .74 | 5 | · · · |
| 1982: | | | | | | | | | | |
| Governor | .6 | 3.0 | · · · | .05 | 41.5 | 4 | 20.0 | .62 | 5 | · · · |
| U.S. senator | .7 | 5.0 | · · · | .04 | 17.5 | 9 | 15.6 | .62 | 5 | · · · |
| Attorney general | 1.1 | 3.6 | · · · | .06 | 5.0 | 5 | 25.1 | .62 | 5 | · · · |
| Auditor | .8 | 7.0 | · · · | .09 | 3.5 | 6 | 4.4 | .62 | 5 | · · · |
| Secretary of state | .7 | 2.9 | · · · | .09 | 4.0 | 7 | 11.6 | .62 | 5 | · · · |
| Treasurer | .8 | 4.0 | · · · | .10 | 7.0 | 8 | 2.3 | .62 | 5 | · · · |
| 1984: | | | | | | | | | | |
| President | .7 | .8 | .7 | .03 | 104.0 | 1 | 18.8 | .74 | 3 | 4 |
| 1986: | | | | | | | | | | |
| Governor | .1 | 10.0 | 6.4 | .06 | 82.0 | 1 | 21.2 | .54 | 3 | 5 |
| U.S. senator | 2.1 | 2.6 | 6.5 | .04 | 21.0 | 6 | 25.0 | .54 | 3 | 5 |
| Attorney general | 2.2 | 4.7 | · · · | .07 | 5.0 | 2 | 19.6 | .54 | 3 | · · · |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Auditor | 1.6 | 7.2 | . . . | .07 | 3.5 | 3 | 33.0 | .54 | 3 | . . . |
| Secretary of state | .6 | 8.9 | . . . | .07 | 6.5 | 4 | 19.4 | .54 | 3 | . . . |
| Treasurer | 1.9 | 1.9 | . . . | .08 | 6.5 | 5 | 9.8 | .54 | 3 | . . . |
| Chief justice | 1.0 | 5.9 | 16.8 | .11 | 26.0 | 7 | 7.6 | .54 | 3 | 5 |
| Justice | 1.0 | 22.3 | . . . | .20 | 7.0 | 8 | 1.2 | .54 | 3 | . . . |
| Justice | 4.4 | 6.0 | . . . | .19 | 5.0 | 9 | .8 | .54 | 3 | . . . |
| 1988: | | | | | | | | | | |
| President | 1.8 | 2.2 | .9 | .02 | 142.0 | 1 | 10.9 | .72 | 5 | 2 |
| U.S. senator | .8 | 1.9 | 7.5 | .03 | 27.0 | 2 | 14.0 | .72 | 5 | 2 |
| 1990: | | | | | | | | | | |
| Governor | .8 | 1.8 | 7.5 | .04 | 77.0 | 4 | 11.4 | .61 | 6 | 5 |
| Attorney general | 3.2 | 7.1 | 1.4 | .07 | 13.0 | 5 | .1 | .61 | 6 | 5 |
| Auditor | 4.5 | 14.3 | 12.6 | .08 | 6.5 | 6 | 5.6 | .61 | 6 | 5 |
| Secretary of state | .3 | 5.8 | .5 | .06 | 20.0 | 7 | 6.0 | .61 | 6 | 5 |
| Treasurer | 2.3 | 9.6 | 5.4 | .09 | 6.0 | 8 | 19.0 | .61 | 6 | 5 |
| 1992: | | | | | | | | | | |
| President | 1.5 | 2.2 | 3.8 | .02 | 146.0 | 5 | 1.9 | .79 | 3 | 2 |
| U.S. senator | 5.2 | 4.5 | 4.2 | .05 | 28.0 | 6 | 8.7 | .79 | 3 | 2 |
| 1994: | | | | | | | | | | |
| Governor | .8 | 7.9 | 2.5 | .06 | 33.5 | 1 | 46.8 | .57 | 1 | 3 |
| U.S. senator | 1.0 | 2.6 | 3.1 | .04 | 57.5 | 6 | 14.2 | .57 | 1 | 3 |
| Attorney general | 3.4 | 4.9 | 2.6 | .06 | 26.0 | 2 | 2.8 | .57 | 1 | 3 |
| Auditor | 2.5 | 2.8 | 5.5 | .10 | 14.0 | 3 | 17.0 | .57 | 1 | 3 |
| Secretary of state | .2 | 8.0 | 2.4 | .09 | 4.0 | 4 | 29.6 | .57 | 1 | 3 |
| Treasurer | .2 | 1.1 | 2.5 | .09 | 10.0 | 5 | 11.4 | .57 | 1 | 3 |
| Average | 1.6 | 5.4 | 4.9 | | | | | | | |

average error of 0.1% in the gubernatorial race in 1986 to 5.2% in the U.S. Senate race in 1992.[2] Forecasts based on the University of Akron preelection surveys for the same races were considerably less accurate (see table 1, col. 2). The Akron surveys had an average error of 5.4%, ranging from a low of 0.6% in the 1994 race for auditor to a high of 22.3% in the 1986 race for justice. The University of Cincinnati conducted surveys for 19 of the races from 1984 to 1994, with an overall average error of 4.9% (see table 1, col. 3).[3] The accuracy of the Cincinnati surveys ranged from a 0.5% error in the secretary of state race in 1990 to a 16.8% error in the chief justice race in 1986.

The *Dispatch* and the University of Akron conducted preelection surveys for the same 32 statewide candidate races between 1980 and 1994. Across the 32 races, the *Dispatch* forecasts were significantly more accurate than the Akron forecasts ($t(31) = 4.98, p < .01$).[4] The University of Cincinnati conducted surveys for 19 of these races, and again, the accuracy of the *Dispatch* poll across these 19 races was significantly higher ($t(18) = 3.29, p < .01$). For these 19 races, we can compare the overall accuracy of the forecasts of all three outfits: the average error of the *Dispatch* forecasts was 1.7%, compared to a 5.1% average error for the University of Akron forecasts and a 4.9% average error for the University of Cincinnati forecasts. Thus, the two telephone surveys seem to be roughly comparable and were not significantly different in terms of accuracy ($t(18) = 0.17, p = .87$).[5]

2. We calculated the average error for each race by summing the absolute value of the actual proportion of the vote each candidate received minus the proportion of the vote each was predicted to receive and dividing by the number of candidates in the race.

3. In 1994, Cincinnati used a new strategy for allocating undecided respondents, which we will describe later. The analyses presented here reflect the accuracy of the University of Cincinnati surveys when "undecided" responses were dropped from the 1994 final forecasts, as was done in previous years. This provides continuity across election years in the handling of "undecided" responses and thus allows a meaningful comparison of the differences in accuracy across years between mail and telephone forecasts.

4. *t*-tests of this sort should be considered heuristic tools for assessing the reliability of the differences between the three surveys, rather than formal tests of these differences. *t*-tests are based on the assumption that observations are independent of one another. In our data, this assumption is violated. Within each preelection survey, the same respondents contributed data to the forecasts of several races. Furthermore, absolute values of the errors are bounded at zero, so the distribution of these measures of forecast error is not normal, which violates another assumption underlying the *t*-tests. To correct for nonindependence, we averaged the accuracy of the forecasts of races in each year and gauged the significance of the differences in accuracy across the eight election years rather than across the 32 races. Using this conservative test, we found that the *Dispatch* forecast error was significantly smaller than the Akron error ($t(7) = 2.90, p < .02$) and the Cincinnati error ($t(5) = 2.16, p < .05$). This is reassuring about the validity of our conclusions.

5. In 1994, the *Dispatch* hired the Gallup Organization to conduct a statewide preelection telephone survey, which permits an examination of the methodology and forecast accuracy of a fourth survey outfit. Gallup conducted random-digit dialing telephone interviews with a statewide sample of 805 Ohio adults between November 2 and November

CORRELATES OF ACCURACY

We also explored the susceptibility of the three polls to several factors that may affect forecast accuracy, indicated by publicity, roll-off, the order in which the race appeared on the ballot, the margin of victory, overall voter turnout, and the length of time between completion of the survey and Election Day.

*Publicity.* We anticipated a negative correlation between the amount of publicity a race received and the degree of prediction error for that race. If a race is highly publicized, voters are more likely to be informed about the race, to have thought about it, and to have crystallized opinions about it. These factors should not only result in more accurate measurement of voter preferences but they should lead to less shifting of opinion between the time a survey is conducted and Election Day.

To test this prediction, we assessed the publicity of each race by content analyzing the headlines of the *Columbus Dispatch* and the *Cleveland Plain Dealer,* two of the largest circulation newspapers in Ohio, from October 1 through Election Day for each of the election years. The correlation between the two newspapers in terms of the number of times each race was mentioned was .90, so we averaged the two figures for each race. As expected, we found a negative correlation between publicity and degree of forecast error for all three surveys (see table 2, row 1). However, this relation was stronger in the Akron and Cincinnati surveys ($r = -.35$ and $r = -.23$, respectively) than in the *Dispatch* surveys ($r = -.15$). This indicates a greater consistency in the accuracy of the *Dispatch* surveys across variations in the amount of publicity received by races.[6]

---

4, 1994, using methods quite similar to those used by the universities of Akron and Cincinnati. However, the Gallup methodology had some unique advantages and disadvantages. Unlike Akron and Cincinnati, Gallup weighted the data from the full sample to match statistics from the U.S. Census on the age, gender, education, race, and geographic location of Ohio residents. However, unanswered calls were retried only three times, and a less optimal respondent selection strategy was used: within each household, the youngest male, 18 years or older, who was at home at the time of the call was selected to be interviewed; if no such male was available, the oldest female, 18 years or older, who was at home at the time was interviewed. The *Dispatch* poll was substantially more accurate (average error = 1.4%) than the Gallup poll (average error = 5.2%) in forecasting the 1994 elections. In fact, of the three statewide telephone surveys conducted in 1994 (by the University of Akron, the University of Cincinnati, and Gallup), Gallup generated the least accurate forecasts.

6. We also assessed these associations and all those discussed below with unstandardized regression coefficients, and they revealed the same patterns as the correlations reported in the text. One might imagine subjecting all these correlations to tests of statistical significance, as well as conducting multivariate analysis with all predictors of accuracy. Unfortunately, however, the small number of races available here makes such tests and analyses uninformative; even very large correlations are not significant.

**Table 2.** Correlations between Forecast Error and Election Factors

| Source of Error | *Columbus Dispatch* Error | University of Akron Error | University of Cincinnati Error |
|---|---|---|---|
| Average number of newspaper stories | −.15 | −.35 | −.23 |
| Roll-off percent | .17 | .61 | .40 |
| Race order on ballot | .15 | .26 | .33 |
| Margin of victory | −.33 | −.06 | −.13 |
| Voter turnout | .13 | −.38 | −.26 |
| Days between survey and election | N.A. | .04 | .37 |

NOTE.—These correlations are based upon the 32 races listed in table 1.

*Roll-off.* Our content analysis of campaign publicity reflects the amount of press attention paid to particular races, but it does not necessarily gauge the amount of such attention that was translated into knowledge stored in voters' memories. A more direct measure of knowledge levels may be roll-off, the percentage of voters who participated in an election but did not vote in a particular race. Roll-off rates for high-visibility races such as president or governor are typically very low, as few voters who participate in an election neglect to vote in these races. Roll-off is considerably higher for low-visibility races, presumably reflecting lack of knowledge regarding such races (Burnham 1965; Robinson and Standing 1960; Vanderleeuw and Engstrom 1987).

And indeed, we found positive correlations between roll-off rate and degree of forecast error for all three surveys (see table 2, row 2). However, the Akron and Cincinnati surveys were considerably more susceptible to increased error in high roll-off races than were the *Dispatch* surveys: the correlations were .61 and .40 for the Akron and Cincinnati surveys, respectively, and only .17 for the *Dispatch* surveys.

*Race order.* Yet another way to operationalize voter engagement in a race is its placement on the ballot. Because of lower levels of knowledge and interest in races lower on the ballot and perhaps higher levels of voter fatigue, we expected that forecasts for such races may be less accurate than forecasts for races higher on the ballot. As expected, we found positive correlations between race order and degree of error for all three surveys (see table 2, row 3). Again, this correlation was

stronger for the Akron and Cincinnati surveys ($r = .26$ and $r = .33$, respectively) than for the *Dispatch* surveys ($r = .15$), indicating that the former were more susceptible than the latter to error in races at the bottom of the ticket.

*Margin of victory*. We expected to find a positive relation between the margin of victory in a particular race and the forecast error for that race. When preelection polls forecast a landslide election victory, a proportion of the leading candidate's supporters do not bother to vote, which can increase forecast error (Crespi 1988). Additionally, when the margin of victory is forecast to be large, survey participants may be reluctant to admit that they intend to vote for the less popular candidate and may falsely report their candidate preferences, which would also lead to greater inaccuracy in such elections (Noelle-Neumann 1984). In fact, we found very weak negative relations between the margin of victory and forecast error in the Akron and Cincinnati surveys ($r = -.06$ and $r = -.13$, respectively) and a stronger negative relation between margin of victory and forecast error in the *Dispatch* polls ($r = -.33$; see table 2, row 4). Margin of victory, then, appears to be one source of forecast error to which the *Dispatch* surveys were more susceptible than were the telephone surveys.

*Turnout*. We predicted a negative correlation between voter turnout and forecast error. As turnout decreases, the challenge of determining who will actually participate in an election increases, and the potential for forecast error presumably increases as well (Crespi 1988). For the Akron and Cincinnati surveys, we did find negative correlations between turnout and forecast error: $-.38$ and $-.26$, respectively (see table 2, row 5). But for the *Dispatch* surveys, we found essentially no correlation between turnout and forecast error ($r = .13$), suggesting that the *Dispatch*'s accuracy was impervious to fluctuations in levels of voter participation.

*Timing*. We expected that surveys conducted immediately preceding an election would be more accurate than those conducted far in advance. As Election Day draws closer, survey participants are more likely to have established candidate preferences, and these preferences will presumably be strongly correlated with voting behavior (Crespi 1988).

It was not possible to test this hypothesis for the *Dispatch* surveys, because the length of time between the completion of the surveys and the elections did not vary. The Akron and Cincinnati surveys did vary in this respect from 1980 to 1994, although not by great amounts. The Akron surveys were completed as close as 3 days before the election and as far in advance as 6 days. The Cincinnati surveys were conducted 2–5 days in advance. The variation in advance time was essen-

tially uncorrelated with forecast error for the Akron surveys ($r = .04$), but the expected positive relation ($r = .37$) was found for the Cincinnati surveys (see table 2, row 6).

## Possible Explanations for the *Columbus Dispatch* Forecast Accuracy

The success of the *Dispatch* surveys is, in many ways, counterintuitive. Like many mail surveys, the *Dispatch* polls have had very low response rates. With roughly 75% of the samples failing to participate, there was obviously tremendous potential for nonresponse bias to distort the results. Furthermore, at the time of publication, 2 days before each election, the results of the *Dispatch* surveys were as much as a week old. This slow turnaround time precluded the detection of late shifts in candidate support and may therefore have compromised forecast accuracy.

There are, however, a number of possible explanations for the surprising discrepancy in accuracy between the mail and telephone surveys. They include sample size, better identification of likely voters, weighting by region and demographics, interviewer effects, probing "undecided" responses, allocation of undecided respondents, questionnaire design, privacy, anonymity, certainty of survey source, personalized contact, pacing, a visible record of votes, and voting order. We will discuss each of these possibilities below and report some empirical tests of them.

### SAMPLE SIZE

*Hypothesis.* The Dispatch has based its forecasts on almost twice as many participants as were included in the telephone surveys. Conventional polling wisdom suggests that within the range of typical sample sizes, larger samples do not notably improve forecast accuracy (e.g., Rosenstone 1983). And in fact, in studies of hundreds of preelection polls, Crespi (1988) and Lau (1994) concluded that increases in sample size from 100 people to 6,000 people (in Crespi's case) and from 575 to 2,100 (in Lau's case) did not increase forecast accuracy. This conclusion might seem to suggest that differences in sample sizes between the *Dispatch* and the telephone surveys did not contribute to the observed differences in forecast precision.

This conclusion flies in the face of conventional sampling theory, which indicates that standard errors of estimates do decrease as a result of such sample size increases (Kish 1965). Of course, sampling error is only one source of survey inaccuracy. If forecast error was

mostly due to sources other than sampling error, then one might not expect increases in sample size to improve accuracy. However, both Crespi (1988) and Lau (1994) did find the expected negative correlation between sample size and forecast error, though it was not statistically significant in either study. Sample size differences may, therefore, have played at least some role in accounting for accuracy differences.

*Evidence.* Calculating just how much of the difference in accuracy between the *Dispatch* surveys and the telephone polls is attributable to sample size can be done for two-candidate races using the following formula:

$$\text{sampling error} \approx \sqrt{[p(1-p)/N]}\,\sqrt{(2/\pi)},$$

where $p$ represents the proportion of the vote received by candidate A in the actual election, and $N$ represents sample size (see, e.g., Hogg and Craig 1995, p. 145). According to this formula, the expected absolute deviation for a sample of 1,700 respondents (the average size of the *Dispatch* samples) is .5% smaller than the expected absolute deviation for a sample of 700 respondents (the average size of the telephone surveys).

An empirical exploration of the impact of sample size on forecast accuracy using data from the 1992 and 1994 *Dispatch* surveys yielded similar results. We randomly selected 15 subsamples of 700 respondents from both the 1992 and 1994 *Dispatch* data sets and generated election forecasts from each. The overall average error of the smaller subsamples was 2.1%, compared to an average error of 1.7% for the full samples, a difference of about half of a percentage point. Although this difference was not large, one-third of the smaller samples incorrectly predicted that Bush would defeat Clinton in the 1992 presidential race. This suggests that the *Dispatch*'s forecast accuracy is partly attributable to larger samples.

### BETTER IDENTIFICATION OF LIKELY VOTERS

*Hypotheses.* There are several reasons to believe that the *Dispatch* mail survey methods may have yielded samples that were more representative of the voters who actually participated in each election than were the telephone survey samples. First, the *Dispatch* samples were drawn from lists of registered voters, whereas the telephone samples were drawn from the population of households equipped with telephones. As a result, the Akron and Cincinnati researchers had to identify respondents who were not in fact registered to vote (via self-reports) in order to drop them from their final forecast samples. Unfortunately, though, some people who are not registered typically claim to be so (Katosh and Traugott 1981). Because of overreporting

of registration status, it is likely that candidate preferences expressed by some ineligible voters were not dropped from the telephone survey forecasts, thus contributing error.

An important step in election forecasting is dropping respondents who are registered but are unlikely to turn out. The telephone survey outfits have done so using self-reports of attributes such as interest in the election, intent to vote, and past voting behavior. As with registration status, however, the social desirability of being an active participant in the democratic process often leads to overreporting of these attributes (Katosh and Traugott 1981; Sigelman 1982). Furthermore, reports of interest in an election may be distorted by prior relevant questions in a survey (Bishop, Oldendick, and Tuchfarber 1984a, 1984b). Consequently, use of such self-reports to identify people unlikely to vote may have introduced error into the telephone survey forecasts.

In contrast, although mail survey response rates are typically very low, the people who do respond tend to be highly interested in their topics (Ferness 1951; Jobber 1984). And people highly interested in an election are most likely to vote in it (Campbell et al. 1964). As a result, the self-selected samples of the *Dispatch* mail surveys may have been especially likely to turn out. The very nature of the mail survey response procedure may have effectively eliminated nonvoters from the obtained samples.

Finally, telephone surveys may underrepresent people who are difficult to reach at home during usual interviewing times, such as those who work outside of the home during evenings and on weekends and those who travel often. Unlike telephone interviews, mail questionnaires can be completed at whatever times are convenient for the respondents. If there are systematic differences between people who work nights and weekends and those who travel as compared to others, mail surveys may achieve more representative samples as a result.

*Evidence.* To test these hypotheses, we compared the demographics and party identification of the 1992 *Dispatch* sample and the likely voters from the 1992 University of Akron sample to data from the 1992 Voter Research and Surveys General Election Exit Poll.[7] For the latter, self-administered questionnaires were completed by a representative sample of 1,640 Ohio voters. If the *Dispatch* samples did better represent actual voters, they should have corresponded more closely to the exit poll assessments of voter demographics and party identification.

7. Data from the University of Cincinnati and the Gallup Organization preelection polls were not available to us, so the analyses outlined in this section were conducted with telephone survey data collected by the University of Akron.

Demographic characteristics of the electorate derived from exit polls are subject to all of the same sources of error as other survey data and thus provide imperfect benchmarks against which to measure the representativeness of telephone and mail survey samples. However, Voter Research and Surveys (1992) went to great lengths to ensure that their estimates accurately reflected the characteristics of the 1992 voters. They began by selecting a stratified sample of Ohio precincts proportionate to the number of votes cast in each precinct in the 1988 presidential election. This subset of precincts was selected such that every voter in the state had an equal probability of being selected.

Within each selected precinct, voters from whom to gather data were systematically selected (i.e., every *N*th voter) as they left the polls on Election Day. Participating respondents reported their race, age, and gender and interviewers recorded the race, approximate age, and gender of selected voters who declined to participate. The final data were weighted to mirror the demographics of all selected voters and the final vote totals in each precinct as closely as possible.[8] All of these factors presumably helped the exit poll data to provide a reasonably accurate portrait of the 1992 Ohio voters, especially in terms of demographics.

As expected, the *Dispatch* sample matched the exit poll sample more closely than did the Akron sample in terms of four criteria: party identification, gender, religion, and union membership (see table 3). There was no significant difference between the *Dispatch* and exit poll samples in terms of party identification ($\chi^2(1) = 1.7$, $p = .38$), whereas the Akron sample significantly overrepresented Independents ($\chi^2(2) = 20.1$, $p < .01$). There was also no significant difference between the *Dispatch* and exit poll samples in terms of gender ($\chi^2(1) = .03$, $p = .89$), whereas the Akron sample significantly overrepresented women ($\chi^2(1) = 7.3$, $p < .01$).[9] Similarly, there was no significant difference between the *Dispatch* sample and the exit poll sample in terms of religious affiliation ($\chi^2(3) = 1.6$, $p = .67$), but the difference between the Akron sample and the exit poll sample was significant ($\chi^2(3) = 8.6$, $p = .03$), with an underrepresentation of Protestant respondents. Although both the *Dispatch* and Akron samples differed significantly from the exit poll sample in terms of union membership ($\chi^2(1) = 19.5$,

8. Although the nonresponse rate of the exit poll was relatively high (an average, across all selected precincts nationwide, of 42%), rate of nonresponse was not correlated with the discrepancy between the exit poll measure of voting and the actual election outcome across precincts.

9. There is some evidence that the most-recent-birthday method of respondent selection produces samples that overrepresent women (Salmon and Nichols 1983), which may partially explain the overrepresentation of women in the Akron sample.

**Table 3.** Comparison between the 1992 Ohio Electorate and the 1992 *Columbus Dispatch* Sample and the University of Akron Subsample of Likely Voters

| Variable | Exit Polling Sample (%) | *Dispatch* Sample (%) | University of Akron Likely Voters |
|---|---|---|---|
| Gender: | | | |
| Male | 47.6 | 47.8 | 41.6 |
| Female | 52.4 | 52.2 | 58.4 |
| Average error | | .2 | 6.0 |
| $\chi^2(1)$ | | .03 | 7.3 |
| Significance ($p$) | | N.S. | <.01 |
| Race: | | | |
| White | 88.7 | 92.8 | 92.6 |
| Nonwhite | 11.3 | 7.2 | 7.4 |
| Average error | | 4.1 | 3.9 |
| $\chi^2(1)$ | | 18.5 | 9.6 |
| Significance ($p$) | | <.01 | <.01 |
| Education: | | | |
| Not a high school graduate | 6.8 | 8.8 | 5.9 |
| High school graduate | 34.0 | 30.2 | 33.4 |
| Some college | 27.9 | 28.2 | 28.7 |
| College graduate | 31.3 | 32.7 | 32.0 |
| Average error | | 1.9 | .8 |
| $\chi^2(3)$ | | 9.3 | 1.3 |
| Significance ($p$) | | <.03 | N.S. |
| Religion: | | | |
| Protestant | 58.2 | 57.8 | 52.7 |
| Catholic | 29.0 | 27.9 | 30.6 |
| Jewish | 1.3 | 1.4 | 1.5 |
| None/other | 11.5 | 12.8 | 15.3 |
| Average error | | .7 | 2.8 |
| $\chi^2(3)$ | | 1.6 | 8.6 |
| Significance ($p$) | | N.S. | <.05 |
| Union member in household: | | | |
| Yes | 23.8 | 30.4 | 14.8 |
| No | 76.2 | 69.6 | 85.2 |
| Average error | | 6.6 | 9.0 |
| $\chi^2(1)$ | | 19.5 | 25.6 |
| Significance ($p$) | | <.01 | <.01 |
| Party ID: | | | |
| Democrat | 37.6 | 39.3 | 35.2 |
| Republican | 35.7 | 37.5 | 30.8 |
| Independent | 26.8 | 23.2 | 34.1 |
| Average error | | 2.4 | 4.9 |
| $\chi^2(2)$ | | 1.7 | 20.1 |
| Significance ($p$) | | N.S. | <.01 |
| Overall average error | | 2.7 | 4.6 |

NOTE.—The $\chi^2$ statistics reflect the significance of the difference between the exit poll sample and each survey sample in the distribution of respondent characteristics across each variable. The *Dispatch* and Akron data are unweighted.

$p < .01$, and $\chi^2(1) = 25.6$, $p < .01$, respectively), the *Dispatch* more closely corresponded to the exit poll sample (average discrepancies of 6.6% and 9.0% for the *Dispatch* and Akron samples, respectively).

No *Dispatch* advantage appeared for the remaining two criterion variables. There were significant differences between the exit poll sample and both the *Dispatch* and Akron samples in terms of race ($\chi^2(1) = 18.5$, $p < .01$, and $\chi^2(1) = 9.6$, $p < .01$, respectively), but the two survey samples did not differ significantly from one another on this variable ($\chi^2(1) = 2.05$, $p = .15$). And the *Dispatch* sample differed significantly from the exit poll sample in terms of educational background ($\chi^2(3) = 9.3$, $p < .03$), whereas the Akron sample did not ($\chi^2(3) = 1.3$, $p = .65$).

Across the five demographic variables and party identification, the overall average discrepancy between the *Dispatch* sample and the exit poll sample was 2.7%, compared to an overall average discrepancy of 4.6% for the Akron sample.[10] And whereas the Akron sample was significantly different from the exit poll sample on all but one of the six criteria, the *Dispatch* sample was significantly different for only three of the six. Similar results were obtained when the 1988 and 1990 *Dispatch* and Akron samples were compared to Ohio exit poll samples for those years: the *Dispatch* samples corresponded more closely to the exit poll samples than did the Akron samples (average discrepancies of 2.3 vs. 4.3 in 1988 and 4.3 vs. 4.9 in 1990). These results suggest that the accuracy of the *Dispatch* forecasts may be at least partially due to samples that were more representative of Ohio voters than were the telephone survey samples.

One might expect, then, that weighting the data from the Akron telephone surveys to match the demographics and party identification of the *Dispatch* respondents would improve Akron's forecast accuracy. And, indeed, this was the case: forecasts of the 1994 statewide races based on unweighted data from likely voters had an overall average error of 5.6%, compared to an average error of 4.5% for forecasts based on data from likely voters that had been weighted to match the demographics and party identification of 1994 *Dispatch* respondents (see table 4). These results provide further evidence that the *Dispatch* samples better represented Ohio voters than did the telephone survey samples.

As we have suggested thus far, the *Dispatch*'s advantage in better

10. To calculate the average discrepancy between the survey samples and the exit poll sample on each demographic variable, we subtracted the proportion of exit poll respondents at each level of the demographic variable from the proportion of survey respondents at that level. We then averaged the absolute values of these discrepancies across levels of each demographic variable.

**Table 4.** 1994 University of Akron Election Forecasts Based on Unweighted Data from Likely Voters and on Data from Likely Voters, Weighted to Match the Demographics of the 1994 *Columbus Dispatch* Sample

| Race and Candidates | Actual Election Outcomes | Forecasts Based on Unweighted Data | Forecasts Based on Weighted Data |
|---|---|---|---|
| Governor: | | | |
| Burch | 25.0 | 13.5 | 16.7 |
| Voinovich | 71.8 | 85.7 | 82.6 |
| Inmon | 3.2 | .7 | .7 |
| Average error | | 9.3 | 7.2 |
| Attorney general: | | | |
| Fisher | 48.6 | 52.5 | 51.8 |
| Montgomery | 51.4 | 47.5 | 48.2 |
| Average error | | 3.9 | 3.2 |
| Auditor: | | | |
| Petro | 58.5 | 52.4 | 56.7 |
| Sweeney | 41.5 | 47.6 | 43.4 |
| Average error | | 6.1 | 1.9 |
| Treasurer: | | | |
| Blackwell | 53.7 | 54.7 | 55.5 |
| Licht | 3.9 | . . . | . . . |
| Sykes | 42.3 | 45.3 | 44.5 |
| Average error | | 2.6 | 2.6 |
| U.S. senator: | | | |
| Hyatt | 39.2 | 37.3 | 36.0 |
| DeWine | 53.4 | 57.7 | 60.2 |
| Slovenec | 7.3 | 5.0 | 3.8 |
| Average error | | 2.8 | 4.5 |
| Secretary of state: | | | |
| Brady | 35.2 | 26.4 | 27.4 |
| Taft | 64.8 | 73.6 | 72.6 |
| Average error | | 8.8 | 7.8 |
| Overall average error | | 5.6 | 4.5 |

correspondence to the exit polls could be due to a better suited sampling frame, elimination of people unlikely to vote, and/or better access to hard-to-reach respondents. However, it is also possible that this correspondence is attributable to the *Dispatch*'s larger samples. Just as the *Dispatch* vote forecasts were apparently more accurate because of larger samples, so may the demographics have been as well. To assess the viability of this alternative explanation, we randomly selected five subsamples ($N = 700$) from the full 1992 *Dispatch* sample and compared the demographic characteristics and party identification of the subsamples to those of the exit poll sample and the Akron sample.

Just as was the case with the full sample, the smaller *Dispatch* samples were never significantly different from the exit poll sample in terms of three criteria (party identification, gender, and religion), and the directions of the differences on the remaining three criteria (race, education, and union membership) were the same as occurred for the full sample. Similarly, just as with the full sample, the smaller *Dispatch* samples were never significantly different from the Akron sample in terms of one criterion (race) and the directions of the differences on the remaining five criteria (gender, race, education, union membership, and party identification) were again the same as occurred for the full sample. Across all the criteria, the smaller *Dispatch* samples were an average of only .2% farther from the exit poll results than was the full *Dispatch* sample. Because we obtained equivalent results whether we used the full sample or the smaller subsamples, the superior representativeness of the full *Dispatch* sample apparently cannot be accounted for by its size and is therefore presumably attributable to superior sample selection procedures.

### WEIGHTING

*Hypothesis.* For the *Dispatch* final projections, samples were weighted in order to assure proper balancing of region and demographic characteristics of respondents. In contrast, the Akron and Cincinnati telephone surveys did no such weighting. If weighting enhances accuracy, this may partly account for the *Dispatch*'s advantage.

*Evidence.* We explored this possibility by computing the average error for predictions made from the unweighted and the weighted *Dispatch* data for the elections between 1986 and 1994 (the only years for which the appropriate data were available). The former was 1.6%, whereas the latter was 1.8%. So weighting decreased the *Dispatch* forecasts' accuracy a bit and therefore seems not to have been responsible for any of its advantage.

INTERVIEWER EFFECTS

*Hypothesis.* Interviewer effects are well-documented sources of survey error (Dohrenwend, Colombotos, and Dohrenwend 1968; Hanson and Marks 1958; Maccoby and Maccoby 1954), and the absence of interviewers eliminated the potential for such error in the *Dispatch* surveys.

*Evidence.* To test whether interviewer effects reduced the accuracy of the telephone surveys, we used the 1994 University of Akron preelection survey data to gauge differences in results across the 29 interviewers, who were randomly assigned to respondents.[11] In fact, across the six statewide races, no significant effect of interviewer on candidate choice emerged (governor, $\chi^2(56) = 48.37$, $p = .76$; attorney general, $\chi^2(28) = 21.21$, $p = .82$; auditor, $\chi^2(28) = 18.64$, $p = .91$; secretary of state, $\chi^2(28) = 24.14$, $p = .87$; treasurer, $\chi^2(28) = 29.65$, $p = .38$; U.S. senator, $\chi^2(56) = 57.88$, $p = .41$). Thus, we found no support for the notion that interviewer effects detracted from telephone survey accuracy.

PROBING "UNDECIDED" RESPONSES

*Hypothesis.* Although very few *Dispatch* respondents failed to express a candidate preference for any given race, the telephone survey respondents were quite a bit more likely to do so. For all of the races before 1994, University of Cincinnati telephone respondents who said they were undecided were asked whether they leaned toward one candidate or another. This "leaning" question was also posed to University of Akron respondents for 12 of the 32 races between 1980 and 1994. Some individuals may have provided top-of-the-head responses to such questions that were less predictive of actual voting behavior than more thoughtful responses. Consequently, this unique aspect of some telephone surveys may have contributed error to their forecasts.

*Evidence.* To assess the impact of probing "undecided" telephone responses, we generated two forecasts for the five races in which the University of Akron had done such probing in a comparable way across races.[12] The first forecasts were based only on responses from

11. Telephone numbers were randomly assigned to interviewers, and respondents were randomly selected within households. People who initially refused to participate in the survey, however, were recontacted by senior interviewers who tried to persuade them to reconsider. These respondents, therefore, were not randomly assigned to interviewers. The number of such respondents, though, was extremely small (less than 2%), so our tests of interviewer effects were presumably not clouded by the nonrandom assignment of these respondents to interviewers.

12. In 1994, for the first time, the University of Akron explicitly offered respondents the opportunity to say they would skip a particular race. Presumably, a portion of

participants who initially indicated a candidate preference. The second forecasts also included responses from participants who initially indicated that they were undecided but, after being probed, indicated that they were leaning toward one of the candidates (called leaners).

Including the leaners seemed to have improved forecast accuracy slightly (see table 5). For three races (Bush vs. Dukakis in 1988, Voinovich vs. Metzenbaum in 1988, and Clinton vs. Bush vs. Perot in 1992), average error was smaller when the leaners were included than when they were not. For the remaining two races (Glenn vs. DeWine vs. Grevatt in 1992 and Celebrezze vs. Voinovich in 1990), average error was virtually the same whether leaners were included or not. Across the five races, the average error was slightly smaller when leaners were included (2.9%) than when they were not (3.3%). These results are inconsistent with the notion that probing "undecided" responses significantly reduced forecast accuracy, so this seems not to be a shortcoming of the telephone survey approach.

### ALLOCATION OF "UNDECIDED" RESPONSES

*Hypothesis.* Between 1980 and 1994, many telephone survey respondents were classified as not yet having decided which candidate they would vote for in a particular race (see the figures reported in the appendix). And except for the University of Cincinnati's 1992 and 1994 surveys, the surveys' final forecasts were based only on individuals who expressed preferences (i.e., "undecided" responses were dropped). This approach presumes that undecided individuals would refrain from voting in the particular race or would ultimately vote for each candidate in proportions that matched the split among decided respondents. These assumptions may not have been valid, so this approach may have created unique error in the telephone survey forecasts (for evidence consistent with this reasoning, see Kim [1995]).

*Evidence.* To see whether dropping undecided respondents from forecasts hurt the accuracy of the telephone surveys, we assessed whether an alternative approach yielded more accurate results. For their 1994 survey, the University of Cincinnati developed and implemented a new strategy for allocating undecided respondents using three rules. First, undecided "core" voters in the Democratic or Re-

---

respondents who would otherwise have indicated that they were undecided about the race chose this option instead. Because probing the remaining undecided respondents is not comparable to the probing that was done in earlier years, the 1994 data are not informative regarding the impact of probing in those earlier years. Data from the 1984 Akron survey were not available for analysis of the impact of probing undecided respondents, so we were only able to examine its impact in five of the six races in which the probes were comparably used.

**Table 5.** University of Akron Election Forecasts Based on Initial Candidate Preferences and on Initial Preferences Plus Leaners (Percentages)

| Race and Candidates | Election Outcomes | Forecasts Based on Initial Responses | Forecasts Including Leaners |
|---|---|---|---|
| 1988: | | | |
| President: | | | |
| Bush | 55.0 | 59.8 | 59.1 |
| Dukakis | 44.1 | 40.2 | 40.9 |
| Average error | | 4.8 | 4.1 |
| *N* | | 946 | 1,011 |
| U.S. senator: | | | |
| Voinovich | 43.0 | 38.7 | 39.3 |
| Metzenbaum | 57.0 | 61.3 | 60.7 |
| Average error | | 4.3 | 3.7 |
| *N* | | 941 | 1,009 |
| 1990: | | | |
| Governor: | | | |
| Celebrezze | 44.3 | 43.7 | 45.0 |
| Voinovich | 55.7 | 56.3 | 55.0 |
| Average error | | .6 | .7 |
| *N* | | 446 | 535 |
| 1992: | | | |
| President: | | | |
| Clinton | 40.2 | 43.2 | 42.3 |
| Bush | 38.3 | 38.9 | 38.7 |
| Perot | 21.5 | 17.8 | 19.0 |
| Average error | | 2.4 | 1.7 |
| *N* | | 701 | 789 |
| U.S. senator: | | | |
| Glenn | 51.0 | 51.4 | 52.4 |
| DeWine | 42.3 | 48.6 | 47.6 |
| Grevatt | 6.7 | N.A. | N.A. |
| Average error | | 4.5 | 4.5 |
| *N* | | 646 | 750 |
| Overall average error | | 3.3 | 2.9 |

publican Party (identified on the basis of strength of party identification and consistency of partisan voting over the last 5 years) were allocated to their party's candidate.[13] Second, in races with an incumbent, undecided "swing" voters (i.e., those who were not "core" voters in either party) were allocated two to one to the challenger. Finally, in races with no incumbent, undecided "swing" voters were allocated to each candidate in proportions that matched the split among swing voters who did indicate a candidate preference.

This strategy substantially improved the accuracy of the University of Cincinnati forecasts. As reflected in table 6, forecasts for five of the six statewide races were more accurate when undecided respondents were allocated to a candidate rather than dropped, and the overall average error of the Cincinnati forecasts decreased by 1.1% as a result of the allocation strategy. Furthermore, when undecided respondents were allocated, the University of Cincinnati was the only survey organization to correctly forecast all six of the 1994 statewide races. Thus, the handling of undecided respondents can have an impact on forecast accuracy, and the absence of undecided responses in the *Dispatch* surveys eliminated this potential source of error.

## QUESTIONNAIRE DESIGN

The *Dispatch* questionnaire was designed to match an actual election ballot as much as possible, whereas the telephone questionnaires were

13. Three items were used to identify "core" voters. First, strength of party identification was assessed using a 7-point scale, where 1 represented "strong Republican" and 7 represented "strong Democrat." Similarly, partisan voting behavior over the last 5 years was assessed on a 7-point scale, where 1 represented "only voted for Republicans," 2 represented "voted for mostly Republicans," 3 represented "voted for slightly more Republicans than Democrats," 4 represented "voted for an equal number of Republicans and Democrats," 5 represented "voted for slightly more Democrats than Republicans," 6 represented "voted for mostly Democrats," and 7 represented "only voted for Democrats." Finally, intention to vote consistently for candidates of one political party was assessed. For each of the eight races (six statewide races, plus regional races for U.S. House of Representatives and Ohio House of Representatives), each vote for a Democrat was scored as +1 and each vote for a Republican was scored as −1. Next, the three scales were converted to range from zero to 33.3, where zero corresponded to the strongest Republican response and 33.3 corresponded to the strongest Democratic response. The scales were then summed, resulting in a single 101-point scale, where, in principle, zero would represent respondents who identified themselves as strong Republicans, voted for only Republicans over the last 5 years, and intended to vote for all Republicans in the upcoming election. A score of 100 would, in principle, indicate that the respondent was a strong Democrat, voted for only Democrats over the last 5 years, and intended to vote for all Democrats in the upcoming election. Finally, the 101-point scale was divided into three equal segments. Respondents who scored in the top one-third of the scale were identified as core voters in the Democratic Party, and those who scored in the bottom one-third of the scale were identified as core voters in the Republican Party. Respondents who scored in the middle one-third of the scale were identified as swing voters.

**Table 6.** Accuracy of 1994 University of Cincinnati Forecasts When Undecided Respondents are Dropped and When Undecided Respondents are Allocated to a Candidate (Percentages)

| Race and Candidates | Election Outcomes | Undecided Respondents Dropped | Undecided Respondents Allocated |
|---|---|---|---|
| Governor: | | | |
| Burch | 25.0 | 21.4 | 23.9 |
| Inmon | 3.2 | 3.0 | 2.8 |
| Voinovich | 71.8 | 75.6 | 73.3 |
| Average error | | 2.5 | 1.0 |
| U.S. senator: | | | |
| DeWine | 53.4 | 58.1 | 58.5 |
| Hyatt | 39.2 | 36.4 | 35.8 |
| Slovenec | 7.3 | 5.5 | 5.6 |
| Average error | | 3.1 | 3.4 |
| Secretary of state: | | | |
| Brady | 35.2 | 32.8 | 35.6 |
| Taft | 64.8 | 67.2 | 64.4 |
| Average error | | 2.4 | .4 |
| Attorney general: | | | |
| Fisher | 48.6 | 51.2 | 48.7 |
| Montgomery | 51.4 | 48.8 | 51.3 |
| Average error | | 2.6 | .1 |
| Treasurer: | | | |
| Blackwell | 53.7 | 50.2 | 50.6 |
| Licht | 3.9 | 4.1 | 4.5 |
| Sykes | 42.3 | 45.7 | 44.9 |
| Average error | | 2.5 | 2.1 |
| Auditor: | | | |
| Petro | 58.5 | 53.0 | 53.8 |
| Sweeney | 41.5 | 47.0 | 46.2 |
| Average error | | 5.5 | 4.7 |
| Overall average error | | 3.1 | 2.0 |

quite different. For example, the *Dispatch* questionnaires asked respondents simply to indicate their choices. In contrast, the telephone survey questionnaires asked this question in a variety of different, more complex ways over the years. Because question wording can shape responses (Schuman and Presser 1981), the lack of correspondence between the telephone survey candidate preference items and

the format of an actual election ballot may have reduced forecast accuracy.

The accuracy of telephone survey forecasts may also have been compromised by question order effects. No questions preceded the candidate preference items in the *Dispatch* questionnaires, but various preliminary questions have been included in the telephone questionnaires. In fact, all of the Akron surveys between 1980 and 1994 included at least 10 and as many as 23 substantive political questions prior to the candidate preference items. Answers to the preliminary questions may have made certain issues or considerations more accessible and thus more likely to be used when respondents were later asked about which candidates they favored (Judd et al. 1991).

Furthermore, answers to certain preliminary questions may have constrained later responses. For instance, the 1994 Akron survey included a question about party identification early in the interview. Having identified themselves as Democrats or Republicans, some respondents may have been motivated to maintain images of being loyal to their parties. Consequently, they may have been less likely to reveal any intentions to cross party lines when subsequently asked about their vote choices.

The University of Cincinnati surveys did not ask substantive political questions early in the interview, but they did assess registration status, interest in the election, and intent to vote. As we have discussed, these factors are likely to be overreported (Katosh and Traugott 1981; Sigelman 1982). And having just indicated that they are eligible to vote, are interested in the election, and intend to vote, some respondents may have been less likely to admit that, just days before the election, they had not decided for whom to vote or that they would probably not cast votes in certain races. Instead, they may have provided top-of-the-head responses, which may have distorted forecast accuracy.

To see whether questionnaire design differences might have accounted for the accuracy differences, we collected survey data just before the 1994 election using both the *Dispatch* and the University of Akron questionnaire formats via telephone interviews and self-administered questionnaires.

*Self-administered questionnaire experiment.* In one experiment, 400 adults were approached by the Opinion Centers of America between November 4 and November 6, 1994, at a shopping mall in Cleveland and were asked to complete a self-administered questionnaire concerning the upcoming election. The sample was made up of 37% Democrats, 32% Republicans, and 30% who indicated no party affiliation. The respondents were 90% white, 50% women, and predominantly Protestant or Catholic (77%). The majority of respondents were either

college graduates (43%) or had some college education (33%), whereas 19% of the respondents completed only high school and 5% did not graduate from high school.

Respondents answered one of two different questionnaires: the 1994 *Columbus Dispatch* questionnaire or a questionnaire designed to parallel the 1994 telephone survey questionnaire used by the University of Akron.[14] The questionnaires were arranged in a random order and distributed to participants by interviewers who were unaware of which format each participant was receiving. The two samples did not differ significantly on five demographic variables or party identification (race, $\chi^2(1) = 1.93$, $p = .19$; gender, $\chi^2(1) = 1.16$, $p = .55$; religion, $\chi^2(4) = 3.69$, $p = .44$; education, $\chi^2(4) = 4.66$, $p = .32$; party identification, $\chi^2(2) = 1.88$, $p = .39$), which suggests that random assignment to questionnaire format was successful.

As expected, the formats yielded significantly different distributions of candidate preference responses for three of five statewide races (the race for auditor, $\chi^2(1) = 5.62$, $p < .02$; attorney general, $\chi^2(1) = 9.18$, $p < .01$; and secretary of state, $\chi^2(1) = 7.48$, $p < .01$; see table 7).[15] Thus, questionnaire format did make a difference in the self-administered mode and may have contributed to the accuracy of the *Dispatch* surveys.

*Telephone interview experiment.* To assess which format yielded more accurate results, we conducted a second experiment. For this study, a randomly selected statewide sample of 1,090 Ohio adults were contacted by telephone between November 1 and November 7, 1994. Each day trained interviewers from the University of Akron Survey Research Center were randomly assigned to administer one of two questionnaire formats: the standard University of Akron questionnaire or a questionnaire paralleling the *Dispatch*'s. As in previous University of Akron surveys, unanswered calls were retried seven times, and people who were selected but who chose not to participate in the survey were recontacted by a senior interviewer in an effort to persuade them to reconsider. Within each selected household, the adult resident who had the most recent birthday was selected to participate. The response rate was 56% for the *Dispatch* format and 59% for the Akron format.

14. Copies of the instruments used in the self-administered and telephone interview experiments can be obtained from us.

15. Although there were six statewide Ohio races in 1994, only five of the races were compared in these analyses because of differences between the two questionnaire formats in the response options provided for one race. For the state treasurer race, the *Dispatch* provided three candidate response options: Blackwell (the Republican), Sykes (the Democrat), and Licht (an independent). The University of Akron did not mention Licht.

**Table 7.** Results from 1994 Preelection Experiments (Percentages)

| Race and Candidates | Election Outcomes | Survey Results: Self-Administered Mode: *Dispatch* Format | Survey Results: Self-Administered Mode: University of Akron Format | Survey Results: Telephone Mode: *Dispatch* Format | Survey Results: Telephone Mode: University of Akron Format |
|---|---|---|---|---|---|
| Governor: | | | | | |
| Burch | 25.0 | 25.3 | 18.9 | 20.3 | 13.6 |
| Inmon | 3.2 | 3.5 | 4.2 | 2.2 | 1.9 |
| Voinovitch | 71.8 | 71.2 | 76.8 | 77.4 | 84.4 |
| Average error | | | | 3.7 | 8.4 |
| Attorney general: | | | | | |
| Fisher | 48.6 | 63.8 | 48.4 | 53.5 | 53.4 |
| Montgomery | 51.4 | 36.2 | 51.6 | 46.5 | 46.6 |
| Average error | | | | 4.9 | 4.8 |
| Auditor: | | | | | |
| Petro | 58.5 | 54.5 | 66.7 | 51.7 | 51.0 |
| Sweeney | 41.5 | 45.5 | 33.3 | 48.3 | 49.0 |
| Average error | | | | 6.8 | 7.5 |
| Secretary of state: | | | | | |
| Brady | 35.2 | 46.1 | 32.0 | 31.2 | 27.4 |
| Taft | 64.8 | 53.9 | 68.0 | 68.8 | 72.6 |
| Average error | | | | 4.0 | 7.8 |
| U.S. senator: | | | | | |
| DeWine | 53.4 | 49.5 | 56.9 | 52.6 | 55.7 |
| Hyatt | 39.2 | 40.6 | 34.6 | 39.9 | 40.0 |
| Slovenec | 7.3 | 9.9 | 8.5 | 7.6 | 4.8 |
| Average error | | | | .6 | 2.3 |
| Overall average error | | | | 4.0 | 6.2 |

The 1994 Akron survey included 14 political questions before the candidate preference items. Respondents were not explicitly offered a "don't know" response option, but they were invited to indicate that they would probably not cast a vote in particular races. Respondents who volunteered that they were undecided were asked if they leaned toward one candidate or another. As in prior years, respondents who indicated that they were leaning toward one of the candidates were considered supporters of that candidate. To permit a direct comparison of the questionnaire formats (rather than likely voter selection strategies), respondents were not screened for likelihood of voting.

Roughly half of the sample received the Akron format ($N$ = 541) and half received the *Dispatch* format ($N$ = 549). As in the self-

administered experiment, the two telephone survey subsamples did not differ significantly on any demographic variables or party identification (race, $\chi^2(4) = 1.94$, $p = .75$; gender, $\chi^2(1) = .11$, $p = .74$; religion, $\chi^2(5) = 1.93$, $p = .74$; education, $\chi^2(5) = 2.56$, $p = .76$; party identification, $\chi^2(2) = 2.49$, $p = .15$), suggesting that random assignment to questionnaire format was successful.

Again, the two questionnaires yielded substantially different forecasts (see table 7). The average error of the Akron forecasts was roughly double the average error of the *Dispatch* formats for the governor, secretary of state, and U.S. Senate races. On average, the *Dispatch* format predicted the election outcomes with an average error of 4.0%, compared to an average error of 6.2% for the Akron format. Thus, the *Dispatch* questionnaire design apparently contributed to its accuracy.

### OTHER POSSIBLE FACTORS

Several additional factors may have contributed to the superior accuracy of the *Dispatch* forecasts. Although the available data do not permit tests of these explanations, they warrant consideration.

*Privacy*. The process of completing the mail survey afforded *Dispatch* respondents complete privacy, just as the process of voting does. In contrast, telephone respondents expressed their candidate preferences aloud to interviewers, which may have instigated some editing of responses. In addition, telephone respondents might have answered questions within earshot of other household members, which could have produced additional social pressure atypical of actual voting. And, indeed, this notion is consistent with a number of studies suggesting that mail/self-administered questionnaires are less subject to social desirability biases in responses than are telephone interviews (Aquilino 1994; Martin and Nagao 1989; Wiseman 1972–73).

*Anonymity*. The *Dispatch* cover letters assured respondents of their complete anonymity, and the enclosed ballots revealed no signs of being traceable to the respondents after they were completed and returned. This assurance that their survey responses could not be linked to them personally may have made mail survey respondents especially inclined to respond truthfully. Telephone survey respondents, on the other hand, may have recognized that their responses could be linked, if not to them personally, at least to their households. As a result, they may have tailored their responses to conform to perceived social norms. Again, the evidence on mail/self-administered questionnaires suggests that reduced social desirability biases may have advantaged the *Dispatch* (Aquilino 1994; Martin and Nagao 1989; Wiseman 1972–73).

*Certainty of survey source.* Because *Dispatch* respondents mailed their completed questionnaires directly to the newspaper's offices, they could be certain as to the source of the survey. Telephone survey respondents, on the other hand, had grounds to be more skeptical about who was actually conducting the survey. This skepticism may have induced socially desirable or nonsubstantive (e.g., "don't know") responses in the telephone survey data, thus compromising accuracy.

*Personalized contact.* Although the *Dispatch* ballots were anonymous, the accompanying cover letters were addressed specifically to respondents and informed them that they had been selected from among all registered Ohio voters to participate in the survey. This personalized contact may have increased respondents' sense that they were uniquely valuable to the survey and that they were not replaceable in the sample. This may have elicited more meaningful responses.

*Pacing.* Mail survey respondents could complete their questionnaires at their own pace and at a time that was convenient for them. In contrast, telephone survey respondents were required to make judgments at a faster pace at whatever time they happened to be contacted. This may have led the telephone respondents to make less thoughtful judgments than they would have in the voting booth, thus compromising accuracy.

*Visible record of votes.* As mail survey respondents progressed through the questionnaire, they could refer back to previous vote choices they had recorded in order to monitor their overall vote pattern across races. No such visible record of previous choices was present for telephone respondents. If voters sometimes monitor their pattern of voting before finalizing their choices in the voting booth, the lack of this opportunity in the telephone surveys might have disadvantaged them.

*Voting order.* Finally, as in the actual voting booth, mail survey respondents could control the order in which they considered the races. Telephone survey respondents were required to complete the candidate preference items in an order determined by the questionnaire. If order has impact on voting, this, too, could put the telephone approach at a disadvantage.

## Discussion

Despite the well-documented shortcomings of mail surveys, *Columbus Dispatch* election forecasts for the statewide Ohio races from 1980 to 1994 were remarkably accurate. In fact, the *Dispatch* forecasts were considerably more accurate than forecasts from the two other Ohio

survey outfits, both of which employed conventional telephone interview methods. The *Dispatch* forecasts were also less susceptible to error caused by low publicity, high roll-off, later placement of a race on the ballot, and low voter turnout.

This advantage seems to be attributable to at least four factors. First, the *Dispatch* samples were considerably larger. Second, the *Dispatch* survey procedures were more effective at identifying actual voters, as evidenced by the fact that the *Dispatch* respondents were more representative of the voting electorate overall than were the telephone samples' "likely voters." Third, virtually all of the *Dispatch* respondents completed all candidate preference items, whereas many telephone respondents said they were undecided. Decisions about whether and how to allocate these undecided respondents were potential sources of forecast error to which the *Dispatch* was not susceptible. And finally, the *Dispatch*'s questionnaire design appears also to have contributed to its accuracy.

We did not find support for the notion that the accuracy of telephone surveys was distorted by interviewer effects, nor was there evidence that probing "undecided" responses detracted from telephone survey accuracy. These findings are encouraging for telephone survey methodology. In addition, weighting the *Dispatch* data by region and demographics did not improve forecast accuracy.

Other factors may also have contributed to the success of the mail survey methodology in forecasting election outcomes. Future investigations involving manipulations of preelection survey features in side-by-side survey mode comparisons will undoubtedly add to our understanding and provide additional insight into the strengths and liabilities of various survey modes. Our investigation does, however, take some steps toward full understanding of these counterintuitive survey mode effects.

Our results also have some implications regarding ways in which to maximize the validity of telephone survey forecasts. First, it seems helpful to probe "undecided" responses to assess whether respondents lean one way or another in a race. Second, allocating "undecided" responses using some formula for inferring latent candidate preferences seems to be constructive. And finally, matching the questionnaire to the actual election ballot as closely as possible seems desirable.

It may be tempting to presume that the results we have reported generalize to other preelection polling circumstances, but doing so may be premature. It is not clear, for example, whether mail surveys for smaller, regional races, or for election results nationwide, can achieve high accuracy. Even in statewide election forecasting, it may be pre-

mature to assume that similar results would be obtained through comparable operations in other states. Unique conditions in the political or social climate of Ohio may have contributed to the success of the *Dispatch* preelection mail surveys. It is also unclear whether another, less known organization would obtain equal levels of accuracy using the same procedures. Future research addressing these questions will allow greater specificity in the evaluation of mail survey methodology for election forecasting.

One additional interesting implication of our findings involves last-minute shifts during the final days of an election campaign. This often has been cited as the reason for discrepancies between preelection polls and actual election outcomes (Mosteller et al. 1949; Rosenstone 1983). And indeed, candidates often step up their campaign advertising just before Election Day, presumably reflecting the belief that undecided voters can be swayed at that time (West et al. 1995).

There is some evidence to document such last-minute shifts. For example, Crespi (1980) found that even when the impact of other election factors was controlled, the number of days between the completion of a preelection survey and Election Day was a strong predictor of forecast accuracy, such that closer surveys were more accurate. These results have been interpreted as evidence that surveys conducted shortly before an election detect late shifts in voter sentiment, which permits them to accurately predict election outcomes.

The *Dispatch* data have been collected about a week before Election Day, so the success of those data in forecasting outcomes has apparently not been hampered by such last-minute shifts. How could this be? One possibility is that last-minute shifts in preferences expressed in conventional telephone polls are indeed often sizable and are responsible for errors in those polls but do not hamper the accuracy of the mail polls. Gelman and King (1993) recently argued that the ups and downs of candidate popularity during the course of an election campaign may be wholly irrelevant to the final outcome. When asked about their candidate preferences during a campaign, voters may take the opportunity to express their reactions to recent, salient campaign events, no matter how much those events will actually affect their ultimate votes. But the closer voters get to entering the booth on Election Day, the more likely they may be to discard much of the day-to-day hoopla of the campaign and focus on a few considerations that they consider genuinely important, rooted in stable partisan attachments and policy preferences. If mail survey response procedures do indeed more closely match actual voting, they may better elicit choices based on the stable underpinnings that will ultimately be voters' focus. Thus, the last-minute shifts in telephone poll results may reflect their

respondents' finally coming to employ the voting calculi that all voters ultimately implement on Election Day and that mail survey respondents have been using all along.

## Conclusion

The consistently accurate results of the *Columbus Dispatch* mail surveys clearly stand at odds with conventional wisdom. Not long ago, Bradburn and Sudman asserted: "There are . . . many examples of mail surveys with cooperation rates in the 10–20 percent range, or even lower. The biases in such studies are so great as to make the results almost meaningless" (1988, p. 104). Certainly, the *Dispatch* results are far from meaningless. Indeed, as this study indicates, the mail survey methodology used by the *Columbus Dispatch* can render election forecasts that are more accurate than those of more conventional telephone surveys. Of course, mail surveys have limitations, and low response rates are clearly one of them, but it is also clear that these limitations are not necessarily as debilitating as some assume.

In light of this conclusion, it may be wise to reconsider a recommendation made recently by Lau (1994). He suggested that election forecasters abandon the common practice of reporting "margins of error" based solely upon sample size, a notion with which we strongly agree. However, Lau (1994) suggested that, instead, assessments of likely error should be based more on response rates, a notion that is clearly at variance with the findings we have reported. And as Brehm (1993) demonstrated, even surveys with relatively high response rates can significantly underrepresent some groups in the population. Yet adjusting survey results to compensate for underrepresentation of some groups usually has little or no impact on the substantive conclusions of social science research. So to view a high response rate as a necessary condition for accuracy is not necessarily sensible, nor is the notion that a low response rate necessarily means low accuracy.

## Appendix

Below are the actual election results for each of the 32 races between 1980 and 1994 for which the *Columbus Dispatch* conducted surveys, as well as the results of each of the preelection surveys. Whenever possible, the raw data have been included in the table, along with the data that have been weighted by region and demographic information and the data that reflect the dropping of undecided respondents. The University of Akron and the University of Cincinnati results reflect data collected only from respondents identified as likely voters.

**Table A1.**

| Year and Candidate | Actual Election Results | *Columbus Dispatch* Results | | | Comparison Surveys | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | University of Akron Results for Likely Voters | | University of Cincinnati Results for Likely Voters | |
| | | Raw | Raw—Undecided Dropped | Weighted by Region and Demographics | Raw—Likely Voters | Raw—Undecided Dropped | Raw—Likely Voters | Raw—Undecided Dropped |
| 1980 | (11/4/80) | (10/23/80–10/28/80) | | $N$ = 1,796 | (10/26/80–10/30/80) $N$ = 854 | | | |
| President: | | | | | | | | |
| Reagan | 51.5 | 46 | 49.5 | | 43 | 47.8 | | |
| Carter | 40.9 | 38 | 40.8 | | 39 | 43.3 | | |
| Anderson | 5.9 | 8 | 8.6 | | 6 | 6.7 | | |
| Other | 1.7 | 1 | 1.1 | | 2 | 2.2 | | |
| Undecided | | 7 | | | 10 | | | |
| Average error | | | 1.4 | | | 1.9 | | |
| 1982 | (11/2/82) | (10/25/82–10/27/82) | | $N$ = 1,591 | (10/22/82–10/28/82) $N$ = 762 | | | |
| Governor: | | | | | | | | |
| Celeste | 59.0 | | | 61 | 56 | 66.7 | | |
| Brown | 39.0 | | | 38 | 27 | 32.1 | | |
| Goetz | 1.0 | | | 1 | 1 | 1.2 | | |
| Landefeld | .0 | | | 0 | 0 | | | |
| Reupert | .0 | | | 0 | 0 | | | |
| Undecided | | | | | 16 | | | |
| Average error | | | | .6 | | 3.0 | | |

**Table A1.** (*Continued*)

| Year and Candidate | Actual Election Results | *Columbus Dispatch* Results | | | Comparison Surveys | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | University of Akron Results for Likely Voters | | University of Cincinnati Results for Likely Voters | |
| | | Raw | Raw—Undecided Dropped | Weighted by Region and Demographics | Raw—Likely Voters | Raw—Undecided Dropped | Raw—Likely Voters | Raw—Undecided Dropped |
| Attorney general: | | | | | | | | |
| Celebrezze | 61.3 | | | 63 | 50 | 66.7 | | |
| Saxbe | 36.2 | | | 35 | 25 | 33.3 | | |
| Schuller | 2.5 | | | 2 | 0 | | | |
| Undecided | | | | | 25 | | | |
| Average error | | | | 1.1 | | 3.6 | | |
| Auditor: | | | | | | | | |
| Ferguson | 52.2 | | | 53 | 45 | 59.2 | | |
| Campanella | 47.8 | | | 47 | 31 | 40.8 | | |
| Undecided | | | | | 24 | | | |
| Average error | | | | .8 | | 7.0 | | |
| Treasurer: | | | | | | | | |
| Withrow | 48.1 | | | 48 | 30 | 47.6 | | |
| Rinehart | 45.8 | | | 47 | 33 | 52.4 | | |
| Brown | 6.1 | | | 5 | 0 | | | |
| Undecided | | | | | 37 | | | |
| Average error | | | | .8 | | 4.0 | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| U.S. senator: | | | | | | | |
| Metzenbaum | 56.7 | | 58 | 57 | 65.5 | | |
| Pfeifer | 41.1 | | 40 | 28 | 32.2 | | |
| Herzing | 1.1 | | 1 | 2 | 2.3 | | |
| Merel | 1.1 | | 1 | 0 | | | |
| Undecided | | | | 13 | | | |
| Average error | | | .7 | | 5.0 | | |
| Secretary of state: | | | | | | | |
| S. Brown | 53.6 | | 53 | 28 | 54.9 | | |
| V. Brown | 42.0 | | 43 | 23 | 45.1 | | |
| Leech | 4.4 | | 4 | 0 | | | |
| Undecided | | | | 49 | | | |
| Average error | | | .7 | | 2.9 | | |
| 1984 | (11/6/84) | (10/30/84–11/3/84) | $N$ = 1,617 | (10/27/84–11/1/84) $N$ = 776 | | (10/22/84–11/2/84) $N$ = 544 | |
| President: | | | | | | | |
| Reagan | 58.9 | | 60 | 59 | 60.2 | 55 | 60 |
| Mondale | 40.1 | | 40 | 39 | 39.8 | 36 | 40 |
| Other | 1.0 | | 0 | 0 | 0 | 0 | |
| Undecided | | | | 2 | | 9 | |
| Average error | | | .7 | | .8 | | .7 |
| 1986 | (11/4/86) | (10/27/86–10/30/86) | $N$ = 1,560 | (10/27/86–11/1/86) $N$ = 876 | | (10/20/86–10/30/86) $N$ = 509 | |
| Governor: | | | | | | | |
| Celeste | 60.6 | 60.1 | 60.5 | 60 | 70.6 | 59.6 | 67 |
| Rhodes | 39.4 | 39.9 | 39.5 | 25 | 29.4 | 29.5 | 33 |
| Undecided | | | | 15 | | 10.9 | |
| Average error | | .5 | .1 | | 10.0 | | 6.4 |

**Table A1.** (*Continued*)

| Year and Candidate | Actual Election Results | *Columbus Dispatch* Results | | | Comparison Surveys | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | University of Akron Results for Likely Voters | | University of Cincinnati Results for Likely Voters | |
| | | Raw | Raw—Undecided Dropped | Weighted by Region and Demographics | Raw—Likely Voters | Raw—Undecided Dropped | Raw—Likely Voters | Raw—Undecided Dropped |
| Attorney general: | | | | | | | | |
| Celebrezze | 59.8 | 61.0 | | 62.0 | 49 | 64.5 | | |
| Levey | 40.2 | 39.0 | | 38.0 | 27 | 35.5 | | |
| Undecided | | | | | 24 | | | |
| Average error | | 1.2 | | 2.2 | | 4.7 | | |
| Auditor: | | | | | | | | |
| Ferguson | 66.5 | 67.8 | | 68.1 | 59 | 73.7 | | |
| Rose | 33.5 | 32.2 | | 32.9 | 21 | 26.3 | | |
| Undecided | | | | | 20 | | | |
| Average error | | 1.3 | | 1.6 | | 7.2 | | |
| Secretary of state: | | | | | | | | |
| S. Brown | 59.7 | 60.2 | | 60.3 | 48 | 68.6 | | |
| Campanella | 40.3 | 39.8 | | 39.7 | 22 | 31.4 | | |
| Undecided | | | | | 30 | | | |
| Average error | | .5 | | .6 | | 8.9 | | |
| Treasurer: | | | | | | | | |
| Withrow | 54.9 | 56.3 | | 56.8 | 42 | 56.8 | | |
| Jacobs | 45.1 | 43.7 | | 43.2 | 32 | 43.2 | | |
| Undecided | | | | | 26 | | | |
| Average error | | 1.4 | | 1.9 | | 1.9 | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| U.S. senator: | | | | | | | |
| Glenn | 62.5 | 59.3 | 60.4 | 56 | 65.1 | 61.1 | 69 |
| Kindness | 37.5 | 40.7 | 39.6 | 30 | 34.9 | 27.8 | 31 |
| Undecided | | | | 14 | | 11.2 | |
| Average error | | 3.2 | 2.1 | | 2.6 | | 6.5 |
| Chief justice: | | | | | | | |
| Moyer | 53.8 | 53.4 | 52.8 | 35 | 47.9 | 31.6 | 37 |
| Celebrezze | 46.2 | 46.6 | 47.2 | 38 | 52.1 | 54.1 | 63 |
| Undecided | | | | 27 | | 13 | |
| Average error | | .4 | 1.0 | | 5.9 | | 16.8 |
| Justice: | | | | | | | |
| Holmes | 50.6 | 50.0 | 49.6 | 13 | 28.3 | | |
| Sweeney | 49.4 | 50.0 | 50.4 | 33 | 71.7 | | |
| Undecided | | | | 54 | | | |
| Average error | | .6 | 1.0 | | 22.3 | | |
| Justice: | | | | | | | |
| Brown | 50.4 | 54.9 | 54.8 | 31 | 56.4 | | |
| George | 49.6 | 45.1 | 45.2 | 24 | 43.6 | | |
| Undecided | | | | 45 | | | |
| Average error | | 4.5 | 4.4 | | 6.0 | | |
| 1988 | (11/8/88) | (10/31/88–11/3/88) | $N$ = 1,485 | (10/24/88–11/3/88) $N$ = 1,432 | | (11/3/88–11/6/88) $N$ = 516 | |
| President: | | | | | | | |
| Bush | 55.0 | 54.9 | 53.2 | 56 | 58.3 | 54.2 | 54.2 |
| Dukakis | 44.1 | 45.1 | 46.8 | 40 | 41.7 | 45.4 | 45.4 |
| Other | .9 | 0 | 0 | 0 | 0 | .4 | .4 |
| Undecided | | | | 4 | | | |
| Average error | | .7 | 1.8 | | 2.2 | | .9 |

**Table A1.** (*Continued*)

| Year and Candidate | Actual Election Results | *Columbus Dispatch* Results | | | Comparison Surveys: University of Akron Results for Likely Voters | | Comparison Surveys: University of Cincinnati Results for Likely Voters | |
|---|---|---|---|---|---|---|---|---|
| | | Raw | Raw—Undecided Dropped | Weighted by Region and Demographics | Raw—Likely Voters | Raw—Undecided Dropped | Raw—Likely Voters | Raw—Undecided Dropped |
| U.S. senator: | | | | | | | | |
| Metzenbaum | 57.0 | 56.5 | | 57.8 | 53 | 58.9 | 63.7 | 64.5 |
| Voinovich | 43.0 | 43.5 | | 42.2 | 37 | 41.1 | 35.1 | 35.5 |
| Undecided | | | | | 10 | | 1.2 | |
| Average error | | .5 | | .8 | | 1.9 | | 7.5 |
| 1990 | (11/6/90) | (10/27/90–11/1/90) | | $N$ = 1,562 | (10/23/90–10/31/90) $N$ = 801 | | (10/24/90–11/1/90) $N$ = 806 | |
| Governor: | | | | | | | | |
| Voinovich | 55.7 | 55.9 | | 54.9 | 51.6 | 53.9 | 52.3 | 57.4 |
| Celebrezze | 44.3 | 44.1 | | 45.1 | 44.2 | 46.1 | 38.8 | 42.6 |
| Undecided | | | | | 4.2 | | 8.9 | |
| Average error | | .2 | | .8 | | 1.8 | | 7.5 |
| Attorney general: | | | | | | | | |
| Fisher | 50.02 | 45.5 | | 46.8 | 31.2 | 42.9 | 37.5 | 48.6 |
| Pfeifer | 49.98 | 54.5 | | 53.2 | 41.6 | 57.1 | 39.7 | 51.4 |
| Undecided | | | | | 27.2 | | 22.8 | |
| Average error | | 4.5 | | 3.2 | | 7.1 | | 1.4 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Auditor: | | | | | | | |
| Ferguson | 52.8 | 56.5 | 57.3 | 49.3 | 67.3 | 56.9 | 65.4 |
| Petro | 47.2 | 43.5 | 42.7 | 23.9 | 32.7 | 30.1 | 34.6 |
| Undecided | | | | 26.8 | | 13 | |
| Average error | | 3.7 | 4.5 | | 14.5 | | 12.6 |
| Secretary of state: | | | | | | | |
| Taft | 53.0 | 53.9 | 52.7 | 41.5 | 47.2 | 46.4 | 52.5 |
| Brown | 47.0 | 46.1 | 47.3 | 46.5 | 52.8 | 41.9 | 47.5 |
| Undecided | | | | 12 | | 11.7 | |
| Average error | | .9 | .3 | | 5.8 | | .5 |
| Treasurer: | | | | | | | |
| Withrow | 59.5 | 61.1 | 61.8 | 45.1 | 69.1 | 50.6 | 64.9 |
| Brachman | 40.5 | 38.9 | 38.2 | 20.2 | 30.9 | 27.4 | 35.1 |
| Undecided | | | | 34.7 | | 22 | |
| Average error | | 1.6 | 2.3 | | 9.6 | | 5.4 |
| 1992 | (11/3/92) | (10/26/92–10/29/92) | $N$ = 1,980 | (10/26/92–10/31/92) $N$ = 833 | | (10/28/92–11/1/92) $N$ = 931 | |
| President: | | | | | | | |
| Clinton | 40.2 | 41.0 | 41.1 | 42.0 | 43.8 | 38.7 | 41.2 |
| Bush | 38.3 | 39.7 | 39.6 | 36.0 | 37.5 | 36.5 | 42.9 |
| Perot | 21.5 | 19.3 | 19.3 | 18.0 | 18.7 | 15.2 | 15.3 |
| Undecided | | | | 4.0 | | 7.8 | |
| Average error | | 1.5 | 1.5 | | 2.2 | | 3.8 |
| U.S. senator: | | | | | | | |
| Glenn | 51.0 | 49.7 | 49.9 | 50.0 | 54.3 | 49.5 | 52.8 |
| DeWine | 42.3 | 50.3 | 50.1 | 42.0 | 45.7 | 40.2 | 46.8 |
| Grevatt | 6.7 | | | | | .3 | .4 |
| Undecided | | | | 9.0 | | 7.7 | |
| Average error | | 5.3 | 5.2 | | 4.5 | | 4.2 |

**Table A1.** (*Continued*)

| Year and Candidate | Actual Election Results | *Columbus Dispatch* Results | | | Comparison Surveys | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | University of Akron Results for Likely Voters | | University of Cincinnati Results for Likely Voters | |
| | | Raw | Raw—Undecided Dropped | Weighted by Region and Demographics | Raw—Likely Voters | Raw—Undecided Dropped | Raw—Likely Voters | Raw—Undecided Dropped |
| 1994 | (11/8/94) | (10/27/94–11/5/94) | | $N$ = 2,046 | (11/1/94–11/7/94) $N$ = 373 | | (11/2/94–11/5/94) $N$ = 838 | |
| Governor: | | | | | | | | |
| Burch | 25.0 | 25.0 | | 26.0 | 12.6 | 15.4 | 19.9 | 21.4 |
| Inmon | 3.2 | 2.3 | | 2.0 | .8 | 1.0 | 2.8 | 3.0 |
| Voinovich | 71.8 | 72.7 | | 72.0 | 68.6 | 83.7 | 70.3 | 75.6 |
| Undecided | | | | | 8.3 | | 12.2 | |
| Skip race | | | | | 8.3 | | | |
| Average error | | .6 | | .8 | | 7.9 | | 2.5 |
| Attorney general: | | | | | | | | |
| Fisher | 48.6 | 51.2 | | 52.0 | 45.0 | 53.5 | 45.3 | 51.2 |
| Montgomery | 51.4 | 48.8 | | 48.0 | 39.1 | 46.5 | 43.1 | 48.8 |
| Undecided | | | | | 7.5 | | 11.7 | |
| Skip race | | | | | 6.2 | | | |
| Average error | | 2.6 | | 3.4 | | 4.9 | | 2.6 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Auditor: | | | | | | | |
| Petro | 58.5 | 59.0 | 56.0 | 37.0 | 55.7 | 45.3 | 53.0 |
| Sweeney | 41.5 | 41.0 | 44.0 | 29.5 | 44.4 | 40.1 | 47.0 |
| Undecided | | | | 7.5 | | 14.5 | |
| Skip race | | | | 6.2 | | | |
| Average error | | .5 | 2.5 | | 2.8 | | 5.5 |
| Secretary of state: | | | | | | | |
| Brady | 35.2 | 34.3 | 35.0 | 20.4 | 27.2 | 28.9 | 32.8 |
| Taft | 64.8 | 65.7 | 65.0 | 54.4 | 72.8 | 59.3 | 67.2 |
| Undecided | | | | 10.7 | | 11.7 | |
| Skip race | | | | 12.1 | | | |
| Average error | | .9 | .2 | | 8.0 | | 2.4 |
| Treasurer: | | | | | | | |
| Blackwell | 53.7 | 55.4 | 54.0 | 32.4 | 54.8 | 42.7 | 50.2 |
| Licht | 3.9 | 4.5 | 4.0 | | | 3.5 | 4.1 |
| Sykes | 42.3 | 40.1 | 42.0 | 26.8 | 45.3 | 38.9 | 45.7 |
| Undecided | | | | 14.8 | | 14.9 | |
| Skip race | | | | 22.8 | | | |
| Average error | | 1.5 | .2 | | 1.1 | | 2.5 |
| U.S. senator: | | | | | | | |
| DeWine | 53.4 | 53.7 | 52.0 | 46.9 | 57.4 | 51.0 | 58.1 |
| Hyatt | 39.2 | 38.4 | 40.0 | 31.1 | 38.0 | 32.0 | 36.4 |
| Slovenec | 7.3 | 7.9 | 8.0 | 3.8 | | 4.8 | 5.5 |
| Undecided | | | | 9.7 | | 12.2 | |
| Skip race | | | | 7.2 | | | |
| Average error | | .6 | 1.0 | | 2.6 | | 3.1 |
| Overall average error | | 1.8 | 1.6 | | 5.4 | | 4.9 |

## References


Aquilino, William S. 1994. "Interview Mode Effects in Surveys of Drug and Alcohol Use." *Public Opinion Quarterly* 58:210–40.

Bishop, George F., Robert W. Oldendick, and Alfred Tuchfarber. 1984a. "Interest in Political Campaigns: The Influence of Question Order and Electoral Context." *Political Behavior* 6:159–69.

———. 1984b. "What Must My Interest in Politics Be If I Just Told You 'I Don't Know'?" *Public Opinion Quarterly* 48:510–19.

Bradburn, Norman M., and Seymour Sudman. 1988. *Polls and Surveys*. San Francisco and London: Jossey-Bass.

Brehm, J. 1993. *The Phantom Respondents*. Ann Arbor: University of Michigan Press.

Burnham, Walter D. 1965. "The Changing Shape of the American Political Universe." *American Political Science Review* 59:7–28.

Campbell, Angus, Philip E. Converse, Warren E. Miller, and Donald E. Stokes. 1964. *The American Voter*. New York, London, and Sydney: Wiley.

Converse, Jean M. 1987. *Survey Research in the United States*. Berkeley, Los Angeles, and London: University of California Press.

Crespi, Irving. 1980. "Polls as Journalism." *Public Opinion Quarterly* 44:462–76.

———. 1988. *Pre-election Polling: Sources of Accuracy and Error*. New York: Sage.

Crewe, Ivor. 1992. "A Nation of Liars? Opinion Polls and the 1992 Election." *Parliamentary Affairs* 45:475–95.

Dohrenwend, D. A., J. Colombotos, and B. P. Dohrenwend. 1968. "Social Distance and Interviewer Effects." *Public Opinion Quarterly* 32:410–22.

Ferness, Abbott L. 1951. "A Note on Stimulating Response to Questionnaires." *American Sociological Review* 16:247–53.

Field, Mervin D. 1983. "Political Opinion Polling in the United States of America." In *Political Opinion Polling*, ed. Robert M. Worcester. London: Macmillan.

Frankovic, Kathleen A. 1992. "Technology and the Changing Landscape of Media Polls." In Mann and Orren 1992.

Gallup, George. 1972. *The Sophisticated Poll Watcher's Guide*. Princeton, NJ: Princeton Opinion Press.

Gelman, Andrew, and Gary King. 1993. "Why Are American Presidential Election Campaign Polls So Variable When Votes Are So Predictable?" *British Journal of Political Science* 23:409–51.

Gray, Jerry. 1993. "Whitman's Tactics Could Still Decide Race in New Jersey." *New York Times*, October 31, p. B1.

Hanson, R. H., and E. S. Marks. 1958. "Influence of the Interviewer on the Accuracy of Survey Results." *Journal of the American Statistical Association* 53: 635–55.

Hogg, R. V., and A. T. Craig. 1995. *Introduction to Mathematical Statistics*. Englewood Cliffs, NJ: Prentice Hall.

Jobber, David. 1984. "Response Bias in Mail Surveys: Further Evidence." *Psychological Reports* 54:891–94.

Judd, C. M., J. W. Downing, R. A. Drake, and J. A. Krosnick. 1991. "Some Dynamic Properties of Attitude Structure: Context-Induced Response Facilitation and Polarization." *Journal of Personality and Social Psychology* 60:193–202.

Kagay, Michael R. 1992. "Variability without Fault: Why Even Well-Designed Polls Can Disagree." In Mann and Orren 1992.

Katosh, John P., and Michael W. Traugott. 1981. "The Consequences of Validated and Self-Reported Voting Measures." *Public Opinion Quarterly* 45:519–35.

Kim, Taeyong. 1995. "Discriminant Analysis as a Prediction Tool for Uncommitted Voters in Pre-Election Polls." *International Journal of Public Opinion Research* 7: 110–27.

Kish, Leslie. 1965. *Survey Sampling*. New York: Wiley.

Lau, Richard R. 1994. "An Analysis of the Accuracy of 'Trial Heat' Polls during the 1992 Presidential Election." *Public Opinion Quarterly* 58:2–20.

Maccoby, Eleanor E., and Nathan Maccoby. 1954. "The Interview: A Tool of Social Science." In *Handbook of Social Psychology,* ed. Gardner Lindzey. Reading, MA: Addison-Wesley.

Mann, Thomas E., and Gary R. Orren, eds. 1992. *Media Polls in American Politics.* Washington, DC: Brookings Institute.

Martin, Christopher L., and Dennis H. Nagao. 1989. "Some Effects of Computerized Interviewing on Job Application Responses." *Journal of Applied Psychology* 74: 72–80.

Mosteller, Frederick, Herbert Hyman, Philip J. McCarthy, Eli S. Marks, and David B. Truman. 1949. *The Pre-election Polls of 1948.* New York: Social Science Research Council.

Noelle-Neumann, Elizabeth. 1984. *The Spiral of Silence: Public Opinion—Our Social Skin.* Chicago: University of Chicago Press.

Perry, Paul. 1960. "Election Survey Procedures of the Gallup Poll." *Public Opinion Quarterly* 24:531–42.

———. 1979. "Certain Problems in Election Survey Methodology." *Public Opinion Quarterly* 43:312–25.

Robinson, James A., and William H. Standing. 1960. "Some Correlates of Voter Participation: The Case of Indiana." *Journal of Politics* 22:96–111.

Rosenstone, Steven J. 1983. *Forecasting Presidential Elections.* New Haven, CT: Yale University Press.

Salmon, Charles T., and John Spicer Nichols. 1983. "The Next-Birthday Method of Respondent Selection." *Public Opinion Quarterly* 47:270–76.

Schuman, Howard, and Stanley Presser. 1981. *Questions and Answers in Attitude Surveys: Experiments on Question Form, Wording, and Context.* New York: Academic Press.

Sigelman, Lee. 1982. "The Nonvoting Voter in Voting Research." *American Journal of Political Science* 26:47–56.

Smith, Tom W. 1990. "The First Straw? A Study of the Origins of Election Polls." *Public Opinion Quarterly* 54:21–36.

Squire, Peverill. 1988. "Why the 1936 Literary Digest Poll Failed." *Public Opinion Quarterly* 52:125–33.

Troldahl, V. C., and R. E. Carter. 1964. "Random Selecting of Respondents within Households in Telephone Surveys." *Journal of Marketing Research* 1:71–76.

Vanderleeuw, James W., and Richard L. Engstrom. 1987. "Race, Referendums, and Roll-Off." *Journal of Politics* 49:1081–92.

Voter Research and Surveys. 1992. "Voter Research and Surveys General Election Polls, 1992." New York: Voter Research and Surveys [producer]; 2d release: Ann Arbor, MI: Inter-University Consortium for Political and Social Research [distributor], 1993.

West, Darrel M., Montague Kern, Dean Alger, and Janice M. Goggin. 1995. "Ad Buys in Presidential Campaigns: The Strategies of Electoral Appeal." *Political Communication* 12:275–90.

Wiseman, Frederick. 1972–73. "Methodological Bias in Public Opinion Surveys." *Public Opinion Quarterly* 36:105–8.