Political Analysis (2005) 13:109–111
doi:10.1093/pan/mpi006

Analyzing the 2000 National Election Study

Jake Bowers
Center for Political Studies 4253 ISR,
University of Michigan, Ann Arbor, Michigan 48106-1248
e-mail: jwbowers@umich.edu

Nancy Burns
Center for Political Studies 4246 ISR,
University of Michigan, Ann Arbor, Michigan 48106-1248
e-mail: nburns@umich.edu
(corresponding author)

Michael J. Ensley
Department of Political Science, 210 Woodburn Hall,
Indiana University, Bloomington, IN 47405
e-mail: ensley@indiana.edu

Donald R. Kinder
Center for Political Studies 4258 ISR,
University of Michigan, Ann Arbor, Michigan 48106-1248
e-mail: drkinder@umich.edu

In an earlier report, two of us (Bowers and Ensley, 2003, National Election Studies Technical Report, www.umich.edu/~nes) provided a general framework for understanding the particular strategy outlined by Fogarty et al. (in this issue). Fogarty et al.'s strategy is to make the face-to-face variables more like the random digit dial (RDD) telephone variables by trimming the ends in order to reduce the variance of the face-to-face (FTF) variables. Perhaps some scholars will want the FTF variables to look like the RDD variables, but that would be a fix for a specific research question. Given the significant differences in the representativeness of the samples, the processes of survey nonresponse, and the quality and character of the responses between data taken from a National Area Probability sample in person and data taken from an RDD telephone sample, research questions involving comparisons with other years in the 50-year time series will require different remedies.

As readers know, the 2000 National Election Study (NES) was carried out at a time when it looked like cost constraints would force the NES away from the face-to-face (FTF) area probability sample it had used for the preceding 50 years and to a random digit dial (RDD) sample with interviews conducted by telephone. Because so little was known about the

Authors' note: We would like to thank the Center for Political Studies and the Provost of the University of Michigan for funding NES Fellowships for Bowers and Ensley. We are grateful to Brian J. Fogarty, Nathan J. Kelly, and H. Whitt Kilburn for raising the issue of mode and the 2000 National Election Study.

comparative quality and character of data across these kinds of samples and modes, NES carried out two side-by-side studies in 2000. Since this study might have to serve as a bridge in the NES time series between the two kinds of studies, it was important to carry each study out using best practices both on the telephone and in person. The two studies differ from one another in three ways: (1) the method of sample selection; (2) the mode of interviewing; and (3) question format, where necessary to improve the quality of the telephone data (for a complete discussion, see Bowers and Ensley 2003).

Best practices on the phone turn out to be very different from best practices in person. For example, seven-point scales impose memory burdens on telephone respondents who do not have access to show cards.[1] Therefore, while the NES time series relies on seven-point scales, we used branching for the telephone study. We made such accommodations to the phone throughout the study, precisely to be able to collect the best data possible on the telephone.

The results from the 2000 mode experiment make clear that there are important differences between data taken in the two modes from these two types of samples, and they underscore the scientific loss to the time series that a move to the phone would cause (Bowers and Ensley 2003; Holbrook et al. 2003).

Given the differences between the data taken from an RDD sample over the telephone and the data taken from a national area probability sample in person, the NES worried about ways to help scholars understand these differences and their implications for analysis. For this reason, the NES commissioned a technical report on analyzing the study (National Election Studies 2000).

In that report, two of us (Bowers and Ensley 2003) provide a general framework for understanding the particular strategy outlined in this issue by Fogarty et al.,[2] whose strategy is to make the face-to-face variables more like the RDD telephone variables by trimming the ends in order to reduce the variance of the FTF variables. Perhaps some scholars will want the FTF variables to look like the RDD variables, but that would be a fix for a specific research question. Research questions involving comparisons with other years in the 50-year time series will require different remedies.

Scholars might be especially leery of trying to force the FTF area probability responses to look like the RDD telephone responses for several reasons: (1) The two kinds of samples (area probability vs. RDD) differ in their representativeness. (2) The two modes entail quite different processes of survey nonresponse—that is, people are differentially likely to be contacted and differentially likely to agree to participate in interviews in the two modes. In 2000, the FTF study had a response rate of 65%, while the phone study's response rate was 56%. (3) There are important differences in the quality and character of responses across the two modes. For example, people are more likely to say "don't know" on the phone than in person; they are more likely to acquiesce; they give more socially desirable responses; and they are more likely to rush through the interview (Bowers and Ensley 2003; Holbrook et al. 2003). Scholars will want to weigh these many differences when deciding whether to combine data from the two modes.

We hope readers interested in analyzing the 2000 NES will look at the discussion of these issues provided in the technical report. Meanwhile, we conclude with two recommendations for analysis taken from the report:

[1]One could imagine mailing such cards to the 60% of respondents whose addresses can typically be matched to their phone numbers, but that 60% of respondents is a systematically biased sample of any telephone sample. And then only a subsample of that 60% would be able to locate the cards when the interviewer called.

[2]In addition to the points we make here, Bowers and Ensley provide a more general discussion of the literature on branching vs. seven-point scales (see, for example, Aldrich et al. 1982, 2002).

> If the researcher seeks to maintain continuity and consistency with research that has utilized studies from previous years, the best solution is to use only the face-to-face respondents. If the researcher seeks to analyze the impact of congressional district level variables on individual behavior and attitudes, it may be better to analyze the respondents that were interviewed over the telephone in the pre- and post-wave of the survey. If there are concerns that necessitate using all of the observations, such as having enough respondents in particular categories (e.g. black, female, age 25–35), the researcher should proceed with caution and check the robustness of any results using the different sets of respondents. In other words, we suggest running the analysis on the face-to-face respondents and the telephone respondents separately, as well as on all of the respondents, in order to assess if there are any significant differences in the results. If this is not possible, the researcher should at least compare the sample statistics (e.g. mean, variance, etc.) of the relevant variables for each mode to detect any significant differences between the modes. (Bowers and Ensley 2003, p. 1)

and

> We do not recommend that analysts use the summary, or combined, version of these scales due to the differences in response options. Instead, we suggest that people see the different scales as a great opportunity to engage in sensitivity analyses—if the same results hold over two different samples (FTF and RDD) and two different variables (Scale and Branch), then researchers can have more confidence in their findings than if the results differ. (Bowers and Ensley 2003, p. 12)

References

Aldrich, John H., John Griffin, and Amy McKay. 2002. Unpublished manuscript, Duke University.

Aldrich, John H., Richard Niemi, George Rabinowitz, and David W. Rohde. 1982. "The Measurement of Public Opinion about Public Policy: A Report on Some New Issue Question Formats." *American Journal of Political Science* 26(2):391–414.

Bowers, Jake, and Michael J. Ensley. 2003. "Issues in Analyzing Data from the Dual Mode 2000 American National Election Study." *National Election Studies Technical Report* (available at www.umich.edu/~nes).

Holbrook, Allyson L., Melanie C. Green, and Jon A. Krosnick. 2003. "Telephone versus Face-to-Face Interviewing of National Probability Samples with Long Questionaires." *Public Opinion Quarterly* 67:79–125.