Differences in an RDD and List Sample:

An Experimental Comparison

Robert W. Oldendick
Institute for Public Service and Policy Research
University of South Carolina

Dennis N. Lambries
Institute for Public Service and Policy Research
University of South Carolina
Differences in an RDD and List Sample:
An Experimental Comparison

Abstract

One of the underlying reasons for the development of random-digit dialing was the belief that telephone samples based on directories were likely to be biased due to their incompleteness (resulting from unlisted numbers, new listings, and the like). RDD samples were more representative because the frame from which the numbers were drawn was more complete. As response rates to RDD surveys have declined in recent years an interest has developed in using list samples to reduce costs in telephone interviewing. This research compares the results of two samples – one selected by RDD and the other a list sample – in terms of the demographic characteristics of respondents, their responses to substantive questions on a range of issues, and the costs of data collection, including number of completions per hour and number of attempts to complete an interview. Implications for the use of list samples in telephone surveys are discussed.
Differences in an RDD and List Sample: An Experimental Comparison

One of the considerations that led to the development of random-digit dialing (RDD) was the concern among public opinion researchers that samples based on directory listings were potentially biased due to the exclusion of individuals or households with voluntarily unlisted numbers or with numbers that were added or changed after the listing was compiled (Cooper, 1964). As Tuchfarber and Klecka (1976: 16) noted, one of the principal advances of RDD was that “it avoids the sampling problems that have traditionally plagued telephone surveys, particularly the exclusion of households with unlisted numbers.”

The concern over excluding some significant portion of the population from potential selection in a sample based on directory listings is understandable, since such lack of coverage is one of the potential sources of survey error (Groves, 1989: 11). During the period of the development and adoption of RDD as a method for selecting samples for telephone surveys, the research that was done on the issue of undercoverage due to sampling only listed numbers (Brunner and Brunner, 1971; Glasser and Metzger, 1975) concluded that the potential biases were sufficient in most instances to warrant consideration of an alternative sampling method. As Brunner and Brunner (1971: 124) conclude, “Thus one cannot assume that listed telephone subscribers are representative of all telephone subscribers. To ignore the differences which do exist may introduce bias that might be significant in some research situations.” The major research efforts in comparing mode of interviewing effects between telephone and face-to-face interviews employed RDD samples for the telephone component of the comparisons due to RDD’s advantage in reaching households with non-listed numbers (Groves and Kahn, 1979; Tuchfarber and Klecka, 1976).
Since RDD has become widely accepted as a method for selecting representative samples in telephone surveys, relatively little research has been done which directly compares the results of samples selected by RDD with those from list samples. Although some such comparisons have been presented in the work of Traugott and his colleagues (Traugott, Groves and Lepkowski, 1987; Traugott and Goldstein, 1989) the purpose of this research was to assess the use of a dual-frame design to improve response rates in telephone surveys rather than to make a direct comparison of RDD and list samples. Similarly, the investigation of this issue by Landenberger, Groves and Lepkowski (1984) focused more on the role of telephone number lists in sample design for telephone surveys than on a comparison of RDD and list sample results. Other research by Leuthold and Scheele (1971) reported the direction of biases that may be found in using telephone directories as the basis for samples.

Recent declines in response rates for RDD telephone surveys have raised concerns about the representativeness of the resulting data and led survey practitioners to investigate alternate approaches. Some of these methods have been operational, such as intensifying interviewer training on avoiding refusals, increasing the number of call attempts made to numbers that are never answered or in which an answering machine is consistently encountered, or developing different messages to leave when an answering is encountered. Other changes in the design of the survey include sending advance letters to those numbers in an RDD sample that can be identified as households (Link, et al. 2003) or using a dual-frame design that includes both RDD and listed numbers (Traugott, Groves and Lepkowski, 1987; Traugott and Goldstein, 1989).

Given the substantial changes in the telephone system, such as the large increases in the number of area codes and the lower proportion of telephone numbers being assigned to residential units and the decrease in the proportion of numbers that appear in directories, together
with the increased use of technologies to screen calls and the public’s increasing reluctance to participate in surveys (Tucker, Lepkowski, and Pierarski, 2002), is the use of a list sample a viable alternative to selecting an RDD sample? In this paper we investigate this question and report the results of an experiment involving two telephone surveys that were identical in all features except that the sample for one was generated by RDD while the second was selected from a list. These samples are compared not only in terms of the demographic characteristics of the respondents and the results for attitude and opinion questions, but also in terms of their costs and the efficiency of data collection.

**Previous Research**

One of the fundamental reasons for conducting survey research is that it permits inferences about a population to be made on the basis of a sample drawn from that population. Striving to produce sample results that are representative of a population is a key element of sample surveys and some of the more memorable, yet infamous, events in the history of survey research are instances in which a representative sample was not achieved. One of the first things that novice survey researchers learn – and are warned about – is the consequence of a non-representative sample as was the case with the *Literary Digest* in 1936 (Squire, 1988). The classic photo of President Truman holding the *Chicago Tribune* headline “Dewey Defeats Truman” is another stark reminder of the need to ensure representativeness. Quota sampling, which was used in early pre-election polling also has its limitations in terms of the sample representativeness. As the Social Science Research Council’s report on the performance of the pre-election polls in 1948 noted, “[U]sers of quota samples should recognize their shortcomings and the implicit difficulty of defending their use when the accuracy of forecasts is important (Mosteller, et al., 1949: 116). Similarly, evaluations of and reactions to what were seen by some
as the “failure” of the polls during more recent elections have stressed the need to collect data from a sample of respondents that reflects those who will turn out on election day, and have identified a number of operational considerations such as sample selection, number of call-backs, refusal conversions, making calls through the night before the election, and adjusting the data to mirror the electorate as measures that have improved the accuracy of telephone surveys for election trial heats (Crespi, 1988; Ladd, 1992; Voss, Gelman and King, 1995; Mitofsky, 1998).

The quest for a representative sample is also evident in the development of RDD for telephone surveys. As expressed by Cooper, (1964: 45) conventional telephone sampling does not provide telephone numbers of subscribers who: 1) have asked specifically that their numbers not be listed; 2) have recently moved to this community, but whose names will not appear until the next directory is issued; 3) have recently moved within the community, but to a different exchange area; 4) are not listed because of oversight or clerical error. While RDD samples have greater coverage than list samples, the inclusion of a relatively large number of not-in-service and business numbers make them relatively inefficient and a number of procedures have been developed to increase the efficiency of such samples (Waksberg, 1978) or to improve the resultant response rates through some type of dual-frame approach (Traugott, Groves and Lepkowski, 1987; Traugott and Goldstein, 1989).

As response rates for surveys, particularly telephone surveys, have declined significantly over the past decade, survey researchers have taken various steps to address this issue (Brehm, 1993; Groves and Couper, 1998; Dillman, Eltinge, Groves, and Little, 2002; Pew Research Center, 2004). One approach that has been taken is to devote increasing resources to maintaining higher response rates. This includes additional interviewer training in how to elicit respondent
cooperation, making more calls to non-answered numbers, and devoting more interviewer hours
to refusal conversions (Curtin, Presser, and Singer, 2000).

Other research has been undertaken that examines the value of devoting more resources
to increasing the response rate on survey results. In a study by Keeter and his colleagues (2000:
125) they compared the results of “a ‘Standard’ survey conducted over a 5-day period that used
a sample of adults that were home when the interviewer called, and a ‘Rigorous’ survey
conducted over an 8-week period that used random selection from among all household
members.” While these two surveys had very different response rates and produced some
significant differences in terms of the demographic characteristics of the respondents, “[V]ery
few significant differences were found on attention to media and engagement in politics, social
trust and connectedness, and most social and political attitudes, including even those toward
surveys” (Keeter, et al., 2000: 125). In discussing the issue of increasing response rates, they
note the opposing views on the effect of increasing response rates. On one hand is the argument
that reducing nonresponse brings into the response pool a more diverse group of respondents,
thus providing a more balanced representation of the population. On the other, the contention is
that increasing the response rate adds more of those persons for whom the survey conditions are
minimally acceptable, and that the survey statistics stay relatively fixed. As they note,
“methodologists have not yet developed theories that are capable of predicting when
nonresponse rates imply nonresponse error and when they do not” (Keeter, 2000: 126).

Given the tremendous increase in the use of telephone screening devices, such as
telephone answering machines, caller ID, privacy managers, and call blocking, over the past 15
years (Tuckel and O’Neill, 1996; Link and Oldendick, 1998; Tuckel and Schulman, 2000),
telephone survey researchers have had to develop procedures for penetrating these screens. At a
minimum this has meant additional calls to such households, varying the day of the week and the
time of the call, in hopes of making contact at a time when the screen was not in use. Procedures
have also been developed for the interviewers to leave messages on answering machines, and
while these efforts have not achieved much success, they do not harm the response rate (Tuckel
and Schulman, 2000: 904).

As noted above, another operational strategy that has been used to increase response rates
in telephone surveys is to send advance letters to those households that can be identified through
cross-matching of the telephone number (Traugott, Groves and Lepkowski, 1987; Traugott and
Goldstein, 1989; Link, 2003). This approach has been useful in increasing response rates and
decreasing initial refusal rates.

Providing respondents with incentives is another means that has been employed to in an
effort to address the increasing problem of noncooperation (Singer, et al., 1999, Singer, Groves
and Corning, 1999; Singer, Van Hoewyk, and Maher, 2000; Trussell and Lavrakas, 2003). Based
on a meta-analysis of the effects of incentives, Singer and her colleagues concluded that “paying
an incentive is effective in increasing response rates in telephone and face-to-face surveys …
This is true for all types of surveys, and not merely in those involving high burden for the
respondent” (Singer, et al., 1999: 225). They note, however, that the effects of incentives are
modest once other variables are controlled, and providing incentives can add substantially to
survey costs.1

1 Another response to declining response rates for some researchers has been increasing use of
Web-based surveys (Taylor, et al., 2001; Greenberg and Rivers, 2001). As with other modes of
data collection, one of the biggest challenges facing Web-based surveys is ensuring that the
sample is representative of the population being studied. As Mitofsky (1999: 24) has noted,
“people recruited into panels are self-selected, with characteristics that differ from the larger
population. People on the Internet do not represent the adult population of the country, and the
Another approach that has been tested – and the one of interest to us in this study – is to consider the use of list samples. For if the typical RDD sample is going to yield a response rate approaching 25%, is the threat of bias and lack of representativeness that much greater if a list is used as the basis for sample selection? Are there operational advantages to using a list sample that compensate for any potential increase in bias? Are there procedures that can be employed to mitigate the increased coverage error?

Several recent studies have addressed this issue. One such investigation was an Australian study designed to assess the differences between an RDD and Electronic White Pages (listed) sample on measures of health conditions and behaviors (Wilson, Starr, Taylor, and Dal Grande, 1999). In this study they reported that the use of directory listings produces only minimal bias over RDD samples and concluded that “the efficiency and methodological ease of conducting Electronic White Pages sampling makes it our preferred method of obtaining health behaviour data” (Wilson, Starr, Taylor, and Dal Grande, 1999: 633). Psaty and his colleagues conducted a similar comparison of RDD and list samples in their study of a “rare and elusive” population – those 65 and older. Their results produced a conclusion similar to that of the Australian study: a commercial telemarketing list can efficiently identify a sample of elderly respondents whose health status and health behavior appear to be comparable with those of elderly identified by random digit dialing techniques (Psaty, et al. 1991: 103). In a 2002 study, Guterbock, Hartman, and Hubbard conducted an experiment comparing results of a list-assisted RDD sample to the results from a directory-listed sample. They reached a somewhat different conclusion: “Despite the substantial cost savings to be gained, we would not recommend switching the … survey to an Electronic White Pages sampling method, because critical, Internet panelists do not even represent people on the Internet. At best, we end up with a large sample representing nothing but itself.”
substantive results of the survey would be artificially altered by the change in method” (Guterbock, Hartman, and Hubbard, 2003: 18). This conclusion is tempered somewhat, however, by the fact that this study involved cross-time comparisons. As they note, “[I]t is not clear what we might recommend to another state that is just beginning to measure customer satisfaction … if Electronic White Pages sampling were employed from the beginning of the survey series, the amount of survey error we have detected here might well be acceptable (Guterbock, Hartman, and Hubbard, 2003: 18).

The general conclusion from previous research on the comparison of RDD and listed samples is that listed samples are more efficient and less costly to administer, and that they generally produce results similar to those of RDD samples. There are, however, sufficient differences in some of the demographic characteristics of the samples and in some of the substantive results that questions about the utility of listed samples in producing representative results remain. In the remainder of this paper we provide some additional data on this issue from an experiment comparing the results of an RDD and a list sample in a statewide survey.

**Study Design**

To examine the differences in an RDD and list sample, we conducted two surveys in which the designs were identical in all aspects except for the method of sample selection. Data were collected by the interviewing staff of a university survey research unit. The same questionnaire was used in the two surveys, all interviewers were trained during a single training session, and each interviewer worked on both studies, though they were not aware whether they were working with the list or RDD sample. Data for both surveys were collected using the survey unit’s computer assisted telephone interviewing system. The field period for both studies was from October 23, 2003 to December 20, 2003. Samples for both studies were provided by
Survey Sampling, Inc. The RDD sample was a statewide list-assisted sample of 3,000 numbers in which numbers were drawn only from number blocks (100 groups) known to have three or more working numbers. The listed sample of 1,700 numbers was selected from telephone-directory listed households in the state. We compare the results from these two sampling methods on a number of factors, including response rate, completions per hour, costs, the demographic characteristics of the respondents, and the results of substantive questions. (Wording of the substantive questions used in this analysis is provided in the Appendix.)

Results

Response, cooperation, refusal and contact rates. Following AAPOR’s (2004) standard definitions, we report the response rate as the number of completed interviews with reporting units divided by the number of eligible reporting units; the cooperation rate as the proportion of all cases interviewed of all eligible units ever contacted; the refusal rate as the proportion of all cases in which a housing unit refuses to do an interview; and the contact rate as the proportion of all cases in which some responsible member of the housing unit was reached by the survey.

If one of the arguments for using a list sample over RDD is that it would improve the response rate, then the results presented in Table 1 indicate that this argument has little merit. The figures presented in this table show the percentages of all numbers that fell into the different categories and the percentage of eligible numbers appearing in these categories, that is after such dispositions as not in service, business, no eligible respondent in household, or changed numbers have been excluded from the denominator. These latter figures are typically used in reporting response rates, cooperation rates, refusal rates, and contact rates (AAPOR, 2004).

These results are largely what might have been expected given the differences in the way in which the samples were selected. For example, the proportion of not in service numbers is
much higher with RDD than in the list sample as is the percentage whose final disposition was “never answered.” Overall, the results for the other call dispositions – consistent answering machines, unable to complete from scheduled callback, refusals, etc. – are virtually the same with the two methods.

While the percentage of completed interviews in the listed sample is slightly higher (54.3% vs. 50.4%), this is due in part to the much higher proportion of never answered numbers in the RDD sample. If these numbers at which no contact has ever been made are excluded from the denominator (i.e., using the “minimum allocation” use for estimating the proportion of such numbers that are eligible used in AAPOR Standard Definition RR3), the resultant rates are 56.3% for the listed sample and 55.5% for RDD; if the percentage of these never answered numbers that would be eligible are estimated proportionally to the eligible numbers that were contacted in the two samples, the response rates, would be 54.7% and 52.7% respectively. Using any of the methods of calculation, the response rate for the two types of sample selection differ only slightly, and the cooperation rates (58.6% listed; 58.0% RDD) and refusal rates (27.8% listed; 26.5% RDD) are also quite similar. The one aspect of the disposition of calls in which the listed approach had some edge over RDD is in the contact rate, 92.7% vs. 87.0%. Overall, however, to the extent that there is some advantage in response rate from using a list sample, it is not very large.

Operational considerations and costs. While in this study the listed sample did not have a great advantage over RDD in terms of response rate, it should have been more efficient and less

---

2 This higher proportion of “never answered” numbers is a function of more “ring, no answer” numbers in the RDD sample. These are numbers in which calls are “merely switched into a ringing machine instead of a nonworking number recording or an operator intercept mode” (Groves and Kahn, 1979: 49). Determining the quantity of such numbers in a particular sample and how they are treated in the calculation of response rates makes a difference in the resulting rate (AAPOR, 2004: 23-24).
expensive to conduct. Previous research on this topic has consistently demonstrated the advantage of list samples over RDD in terms of efficiency and costs (Psaty, et al., 1991: 103; Wilson, Starr, Taylor, and Dal Grande, 1999: 630; Guterbock, Hartman, and Hubbard, 2003: 8).

Table 2 presents comparisons between the two samples in terms of the average number of calls made per completed interview; the average number of calls to those households in which an interview was completed; the average number of completed interviews per interviewer hour; and the average field cost per completed interview. These figures demonstrate that while there are some operational advantages to employing a list sample, the differences in this case are not as large as what would have been expected on the basis of previous research. While the average number of calls made per completed interview is less in the listed sample (16.52) than with RDD (17.96), the gain in efficiency is not nearly as great as that reported by Guterbock, Hartman, and Hubbard (2003), who reported an average of 18.0 attempts per completed interview with RDD compared with 13.6 attempts in their Electronic White Pages sample.

In the present test, most of the differences were in the number of not in service and business disposition calls, most of which were identified on the first attempt. Most of the calls and, more importantly, the interviewers’ time, were spent on households in attempting to locate

---

3 Field costs include interviewer and supervisor training; interviewer, supervisor, and Field Director data collection; the cost of the sample, and long distance telephone charges. Since the training for the two studies was done as apart of a single training session, the training costs were divided equally between the two studies. During data collection, the interviewers were randomly assigned to the two studies and each interviewer worked on both studies. Across the 21 interviewers who worked on these surveys, the biggest difference in time worked on the two studies was an interviewer who worked 103 hours on the RDD study and 78 hours on the listed. Given that the RDD survey required more interviewing hours to complete, this interviewer should have worked 91 hours on the RDD survey, for the relative time spent on the two studies to have been equivalent. The differences in the time that other interviewers spent on the two studies were less than this, with most interviewers spending slightly more hours on the RDD study than on the listed, which would be expected given the greater number of overall hours required to complete it.
an eligible respondent, gain their cooperation, and complete the interview. Overall, the average number of interviews completed per hour was 1.15 for the listed sample and 1.10 for RDD. The average field costs per completed interview were $15.31, listed and $16.02, RDD. In this study, then, the listed sample was more efficient than RDD and resulted in a savings of 4-5% in field costs.

*Sample comparisons: demographic characteristics.* While a listed sample is more efficient that RDD, the primary limitation of using a list sample is that the results are more likely not to be representative of the population of interest. To the extent the list is incomplete or out-of-date, there is the potential that the characteristics of the sample may not represent those of the population. Table 3 presents a comparison of the two samples for seven background characteristics.

As these data indicate, for two of these characteristics – race and age – the differences are statistically significant at the .05 level. The biggest difference is in the racial composition of the two samples. The listed sample yielded a higher percentage of white respondents (78.0% to 69.8%), a smaller percentage of African-Americans (19.6% to 27.1%) and about the same percentage of those of other races (2.4% to 3.2%). On the age variable, the listed sample produced a slightly lower percentage of respondents in the younger age categories (18 to 29 and 30 to 45) and a slightly higher percentage in the older age groups.

In both cases in which there are significant differences, the results from the RDD sample are closer to those from the Census (U.S. Census Bureau, 2002) than are those from the listed sample. Census estimates are that 69.8% the state’s adult residents are white, 27.2% are African-American, and 3.0% are of other races, percentages that are virtually identical to those from the RDD sample. In this study, African-Americans were underrepresented in the list sample.
terms of age, the Census estimates are as follows: 18 – 29, 22.9%; 30 – 45, 30.2%; 46 – 64, 30.7%; and 65 and over, 16.2%. The RDD results were again closer to the Census estimates than those for the listed sample, although both methods tended to underrepresent those in the younger age groups, particularly those ages 18 to 29.\(^4\)

While differences between the RDD results and Census estimates for these characteristics are similar to those generally found in telephone surveys (Lavrakas, 1993), the larger departure of the listed sample for the race and age variables is attributable, at least in part, to lower rates of listed phone numbers among these groups. It would be expected that those in the listed sample would have a higher percentage of listed phone numbers, since that was the basis upon which they were included in the list, and this is in fact the case: 91.6% of the respondents in the listed sample reported that there telephone number was listed as opposed to 78.3% in the RDD sample. In the RDD sample, only 63.6% of African-Americans reported having a listed phone, compared to 84.3% of whites. Similarly, the listed phone number rate varied across age groups, with those in the younger age groups (who were underrepresented in the listed sample) having a lower percentage of listed phones (based on the RDD results). The percentages with a listed phone number across age groups in the RDD sample were as follows: 18 – 29, 71.9%; 30 – 44, 64.9%; 45 – 65, 84.4%; and 65 and over, 89.8%.

Overall, these results demonstrate that a listed sample can vary significantly on some important background characteristics from those produced in an RDD sample. In this study, both methods tended to underrepresent those groups that are typically missed in telephone surveys, but the listed sample had a greater variation from Census estimates than did the RDD sample for

---
\(^4\) When compared with Census estimates, men are also significantly underrepresented in both surveys, although the differences between the RDD and listed samples are not statistically significant. More generally, the groups “missed” in these surveys, particularly younger people and men, are those who tend to be missed in using telephone surveys (Lavrakas, 1993).
African-Americans and for younger people. While this study represents only a single test, the evidence shows that the greater efficiency of using a listed sample may be offset by a less representative sample.

Substantive comparisons: unweighted data. Do these differences in the demographic composition of the two samples lead to differences in substantive responses? To what extent would we reach different conclusions based on the two samples? To examine this, responses to 24 closed-ended questions included in this survey were compared for the two samples. Table 4 presents the modal category, the percent in the modal category, and the chi-square value for the overall distribution for each of these questions.

In only one of these 24 cases was the modal category different, only one comparison was statistically significant at the .05 level, and the absolute value of the mean difference in the modal category across items was 2.0%. On the question of affirmative action in which the modal categories differed, the differences in the overall distribution were not statistically significant (.05 level), and the distinction in the modal category was a function of the virtual 50-50 split in the overall distribution. In the listed sample, 49.0% of those responding said that affirmative action programs were necessary and 51.0% thought they were not needed; in the RDD sample, 53.7% said that such programs were needed and 46.3% thought they were not. For the item on rating the community as a place to live, for which the overall differences were statistically significant, the distributions were as follows:

<table>
<thead>
<tr>
<th></th>
<th>RDD</th>
<th>Listed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excellent</td>
<td>36.2</td>
<td>38.6</td>
</tr>
<tr>
<td>Good</td>
<td>45.2</td>
<td>44.3</td>
</tr>
<tr>
<td>Only Fair</td>
<td>15.7</td>
<td>12.2</td>
</tr>
<tr>
<td>Poor</td>
<td>2.9</td>
<td>4.8</td>
</tr>
</tbody>
</table>

(N) (869) (769)
Even though the distribution is significantly different, one would reach the same substantive conclusion about the public’s impression of their community as a place to live whether one uses the results of the RDD or the listed sample.

Overall, the results of the substantive questions were essentially the same in the RDD and listed samples. The substantive conclusions that would be reached on the basis of the data from the two samples would be the same.

*Substantive comparisons: weighted data.* In many surveys, some type of weight is often applied to the data in order to compensate for nonresponse. Data are commonly adjusted to match some external source, such as Census data (Kalton, 1983). Such weighting is done under the assumption that if the data match the population (based on some external source) on the basis of some known characteristics, then the results for other questions (for which no external source is available) will more closely reflect the true population figures. In this study we weighted the cases in both the RDD and listed samples to match the state’s population on the basis of sex, race, and age, based on Census population data. We then repeated the analysis presented in Table 4, using these weighted data. The results are presented in Table 5.

Given that very few differences were found with the unweighted data and that the impact of weighting should be to reduce differences between the two samples, we would expect only minor differences when weighted data from the two samples are compared; as these data demonstrate, this is in fact the case. As with the unweighted data, there is only one item for which the modal category is different, two of the differences in the overall distributions are statistically significant, and the average mean difference of the percentages in the modal category is 2.3%. The one item for which the modal categories are different – describing race relations in the state – had the following distribution in the two samples:
<table>
<thead>
<tr>
<th></th>
<th>RDD</th>
<th>Listed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excellent</td>
<td>4.9</td>
<td>4.3</td>
</tr>
<tr>
<td>Good</td>
<td>41.3</td>
<td>42.0</td>
</tr>
<tr>
<td>Fair</td>
<td>40.5</td>
<td>42.4</td>
</tr>
<tr>
<td>Poor</td>
<td>13.3</td>
<td>11.3</td>
</tr>
<tr>
<td>(N)</td>
<td>(869)</td>
<td>(769)</td>
</tr>
</tbody>
</table>

These differences are not statistically significant, and the difference in the modal categories occurs because the percentages in the two middle categories are so similar in the two surveys.

The item which showed the largest difference between the two studies was one on whether the person served as an officer or on a committee of any local club or organization in their community. In the RDD sample 67.3% said they had not served, compared to 61.3% in the listed sample.

Comparing the results presented in Tables 4 and 5, there was only a minimal effect of applying the weighting adjustments to these data. The mean difference in the percentage in the modal category was 0.9% between the weighted and unweighted RDD samples and 1.4% for the weighted and unweighted listed samples. The largest difference in the RDD sample was 2.7% (for the trust in government item) and 5.6% (for the item on belonging to any civic, church, business, or other type of club or organization in the community) in the listed sample. In sum, the conclusion from the analysis using weighted data is the same as that for the unweighted data. In this study, the results of the substantive questions were essentially the same in the RDD and listed samples; the conclusions that would be reached with either sample would be the same.\(^5\)

\(^5\) Another way in which the data from these two samples might differ is in the percentage of “don’t know” or other nonresponse to these items. Using this measure, the results of these two samples are again quite comparable. The mean difference in the percentage of “don’t know” responses to the items shown in Table 4 between the two samples (unweighted) is 0.7% and the largest difference (2.9%) is for the item on whether of not the “State does too much,” for which 4.6% of the RDD sample and 7.5% of the listed sample gave a “don’t know” response. In the


Discussion

In Keeter et al.’s (2000) examination of the effect of reducing nonresponse on survey results they posed the question of what differences arise with different response rates, and we pose a similar question: what difference does it make if survey respondents are selected by RDD are from a list? Based on this investigation, is the use of a list sample a viable alternative to selecting an RDD sample? These results provide evidence both for and against the use of a list sample, and the decision about whether to employ a list or an RDD sample in a particular study should be made in the context of total survey error and survey costs (Groves, 1989).

The evidence in support of using a list sample is that, at least in the current study, it achieved a slightly higher response rate, was less costly than the RDD sample, and the results based on the two samples were largely similar. Overall, the conclusions reached on the basis of these two samples would be the same.

In addition to the theoretical concerns raised by the potentially large number of households that do not appear on the listed frame (Landenberger, Groves, and Lepkowski, 1984; Traugott, Groves, and Lepkowski, 1987), we would recommend against the use of list samples in surveys of a general population sample because of one of the differences that was found between the two samples – race. As described earlier, African-Americans are significantly underrepresented in the list sample, and this is likely to be the case whether this study was conducted at a state level or nationally, given the differences in listed household numbers between blacks and whites, a difference which has persisted over an extended period of time (Leuthold and Scheele, 1971; Lavrakas, 1993; Guterbock, Hartman, and Hubbard, 2003). Given weighted data, the mean difference in the percentage of “don’t know” responses was 0.6%, with the largest difference, 2.8%, also found for the “State does too much” item.
the rather large and systematic differences between blacks and whites on a range of political and social issues (Schuman, Steeh, Bobo, and Krysan, 1997; Bardes and Oldendick, 2002), a method which systematically underrepresents blacks has the potential for producing biased conclusions that is not justified in terms of the rather modest gains in efficiency that a listed sample provides.

This is not to argue that list samples should not be used, for there are certainly situations in which the relative cost benefits in using a list sample are much greater than those in the current example. The general similarity of the results in this investigation, as well as those from other research on this topic (Psaty, et al., 1991; Wilson, Starr, Taylor, and Dal Grande, 1999; Guterbock, Hartman, and Hubbard, 2003) suggest that the results from RDD and list samples can be quite comparable, particularly when appropriate weighting adjustments for noncoverage and nonresponse are applied. In situations where a listed sample may be much more cost effective than an RDD sample, as in a study of older respondents, such as that reported by Psaty and his colleagues (1991), or in a study in which households with respondents in a relatively narrow age range (e.g., youth ages 16 to 20) are the population of interest (Wolfson, Altman, and DuRant, 2004), the results of this research suggest that the use of a list sample is justified in terms of survey costs relative to the resultant potential error.
Table 1

Final Disposition of Calls – RDD and Listed Samples

<table>
<thead>
<tr>
<th></th>
<th>RDD</th>
<th>Listed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>% of total</td>
</tr>
<tr>
<td>Completed interview</td>
<td>854</td>
<td>28.5</td>
</tr>
<tr>
<td>Partially completed interview</td>
<td>18</td>
<td>0.6</td>
</tr>
<tr>
<td>Never answered numbers</td>
<td>154</td>
<td>5.1</td>
</tr>
<tr>
<td>Consistent answering machines</td>
<td>66</td>
<td>2.2</td>
</tr>
<tr>
<td>Not available during fielding period</td>
<td>2</td>
<td>0.1</td>
</tr>
<tr>
<td>Callbacks – respondent not available at follow-up</td>
<td>101</td>
<td>3.4</td>
</tr>
<tr>
<td>Language barrier</td>
<td>18</td>
<td>0.6</td>
</tr>
<tr>
<td>Respondent ill; senile; unable to complete</td>
<td>31</td>
<td>1.0</td>
</tr>
<tr>
<td>Refusals</td>
<td>449</td>
<td>15.0</td>
</tr>
<tr>
<td>Changed number</td>
<td>65</td>
<td>2.2</td>
</tr>
<tr>
<td>No eligible respondent in household</td>
<td>15</td>
<td>0.5</td>
</tr>
<tr>
<td>Business</td>
<td>174</td>
<td>5.8</td>
</tr>
<tr>
<td>Fax/Data Line</td>
<td>166</td>
<td>5.5</td>
</tr>
<tr>
<td>Not in service</td>
<td>887</td>
<td>29.6</td>
</tr>
</tbody>
</table>

3000 100.0 1700 100.0

Response rate (RR1) 50.4 54.3
Cooperation rate (COOP1) 58.0 58.6
Refusal rate (REF1) 26.5 27.8
Contact rate (CON1) 87.0 92.7
Table 2

Call Efficiency and Costs – RDD and Listed Samples

<table>
<thead>
<tr>
<th></th>
<th>RDD</th>
<th>Listed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average number of calls per completed interview</td>
<td>17.96</td>
<td>16.52</td>
</tr>
<tr>
<td>Average number of calls to households who completed an interview</td>
<td>5.10</td>
<td>4.96</td>
</tr>
<tr>
<td>Average number of completed interviews per hour</td>
<td>1.10</td>
<td>1.15</td>
</tr>
<tr>
<td>Average field costs per completed interview</td>
<td>$16.02</td>
<td>$15.31</td>
</tr>
<tr>
<td></td>
<td>RDD</td>
<td>Listed</td>
</tr>
<tr>
<td>----------------</td>
<td>--------------</td>
<td>---------------</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>SEX</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>38.4</td>
<td>335</td>
</tr>
<tr>
<td>Female</td>
<td>61.6</td>
<td>537</td>
</tr>
<tr>
<td>RACE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>27.1</td>
<td>230</td>
</tr>
<tr>
<td>White</td>
<td>69.8</td>
<td>593</td>
</tr>
<tr>
<td>Other</td>
<td>3.2</td>
<td>27</td>
</tr>
<tr>
<td>AGE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18 – 29</td>
<td>13.5</td>
<td>114</td>
</tr>
<tr>
<td>30 – 45</td>
<td>28.5</td>
<td>240</td>
</tr>
<tr>
<td>46 – 64</td>
<td>36.9</td>
<td>311</td>
</tr>
<tr>
<td>65 and Over</td>
<td>21.1</td>
<td>178</td>
</tr>
<tr>
<td>EDUCATION</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than High School</td>
<td>12.0</td>
<td>102</td>
</tr>
<tr>
<td>High School Diploma</td>
<td>30.7</td>
<td>261</td>
</tr>
<tr>
<td>Some College</td>
<td>28.4</td>
<td>241</td>
</tr>
<tr>
<td>College Degree</td>
<td>28.9</td>
<td>246</td>
</tr>
<tr>
<td>INCOME</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Under $15,000</td>
<td>14.8</td>
<td>102</td>
</tr>
<tr>
<td>$15,000 - $29,999</td>
<td>22.3</td>
<td>153</td>
</tr>
<tr>
<td>$30,000 - $49,999</td>
<td>25.3</td>
<td>174</td>
</tr>
<tr>
<td>$50,000 and Over</td>
<td>37.6</td>
<td>258</td>
</tr>
<tr>
<td>TYPE OF AREA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>26.9</td>
<td>229</td>
</tr>
<tr>
<td>Suburban</td>
<td>35.9</td>
<td>305</td>
</tr>
<tr>
<td>Rural</td>
<td>37.2</td>
<td>316</td>
</tr>
<tr>
<td>REGION</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upstate</td>
<td>39.5</td>
<td>342</td>
</tr>
<tr>
<td>Midlands</td>
<td>32.5</td>
<td>281</td>
</tr>
<tr>
<td>Lowcountry</td>
<td>28.0</td>
<td>242</td>
</tr>
</tbody>
</table>
Table 4
Unweighted Responses to Substantive Questions

<table>
<thead>
<tr>
<th>Item Description</th>
<th>Modal Category</th>
<th>% in Modal Category</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RDD</td>
<td>LISTED</td>
<td>RDD</td>
</tr>
<tr>
<td>Heard of DHEC</td>
<td>Yes</td>
<td>Yes</td>
<td>90.8%</td>
</tr>
<tr>
<td>Impression of DHEC</td>
<td>Somewhat Positive</td>
<td>Somewhat Positive</td>
<td>46.3%</td>
</tr>
<tr>
<td>Used DHEC Services</td>
<td>No</td>
<td>No</td>
<td>56.3%</td>
</tr>
<tr>
<td>Quality of Service</td>
<td>Very Satisfied</td>
<td>Very Satisfied</td>
<td>55.5%</td>
</tr>
<tr>
<td>Satisfied with Wait Time</td>
<td>Very Satisfied</td>
<td>Very Satisfied</td>
<td>43.8%</td>
</tr>
<tr>
<td>Courtesy and Attitude</td>
<td>Very Satisfied</td>
<td>Very Satisfied</td>
<td>66.5%</td>
</tr>
<tr>
<td>Ability to Answer Questions</td>
<td>Very Satisfied</td>
<td>Very Satisfied</td>
<td>63.9%</td>
</tr>
<tr>
<td>Problems Locating/Getting Services</td>
<td>No</td>
<td>No</td>
<td>96.2%</td>
</tr>
<tr>
<td>Rate Community as Place to Live</td>
<td>Good</td>
<td>Good</td>
<td>45.2%</td>
</tr>
<tr>
<td>Describe Race Relations</td>
<td>Good</td>
<td>Good</td>
<td>42.0%</td>
</tr>
<tr>
<td>Change in Race Relations</td>
<td>Stayed the same</td>
<td>Stayed the same</td>
<td>47.0%</td>
</tr>
<tr>
<td>Trust State Government</td>
<td>Some of the Time</td>
<td>Some of the Time</td>
<td>55.6%</td>
</tr>
<tr>
<td>State Does Too Much</td>
<td>Should do More</td>
<td>Should do More</td>
<td>58.9%</td>
</tr>
<tr>
<td>Education by Public Schools</td>
<td>Good</td>
<td>Good</td>
<td>35.6%</td>
</tr>
<tr>
<td>Education by Local Schools</td>
<td>Good</td>
<td>Good</td>
<td>42.4%</td>
</tr>
<tr>
<td>Health Care to Poor</td>
<td>Fair</td>
<td>Fair</td>
<td>35.1%</td>
</tr>
<tr>
<td>Death Penalty for Murder</td>
<td>Favor</td>
<td>Favor</td>
<td>66.5%</td>
</tr>
<tr>
<td>Affirmative Action Programs</td>
<td>Needed</td>
<td>Not Needed</td>
<td>53.7%</td>
</tr>
<tr>
<td>Services of Health Department</td>
<td>Do Not Know</td>
<td>Do Not Know</td>
<td>61.3%</td>
</tr>
<tr>
<td>Belong to Club or Organization</td>
<td>Yes</td>
<td>Yes</td>
<td>68.6%</td>
</tr>
<tr>
<td>Served as Officer/On Committee</td>
<td>No</td>
<td>No</td>
<td>64.9%</td>
</tr>
<tr>
<td>Race of Club Members</td>
<td>Most</td>
<td>Most</td>
<td>40.7%</td>
</tr>
<tr>
<td>Sex of Club Members</td>
<td>Some</td>
<td>Some</td>
<td>44.5%</td>
</tr>
<tr>
<td>Currently Registered to Vote</td>
<td>Yes</td>
<td>Yes</td>
<td>84.2%</td>
</tr>
</tbody>
</table>
Table 5  
Weighted Responses to Substantive Questions

<table>
<thead>
<tr>
<th>Item Description</th>
<th>Modal Category</th>
<th>% in Modal Category</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heard of DHEC</td>
<td>Yes</td>
<td>Yes</td>
<td>90.0%</td>
</tr>
<tr>
<td>Impression of DHEC</td>
<td>Somewhat Positive</td>
<td>Somewhat Positive</td>
<td>45.3%</td>
</tr>
<tr>
<td>Used DHEC Services</td>
<td>No</td>
<td>No</td>
<td>55.0%</td>
</tr>
<tr>
<td>Quality of Service</td>
<td>Very Satisfied</td>
<td>Very Satisfied</td>
<td>54.9%</td>
</tr>
<tr>
<td>Satisfied with Wait Time</td>
<td>Very Satisfied</td>
<td>Very Satisfied</td>
<td>44.3%</td>
</tr>
<tr>
<td>Courtesy and Attitude</td>
<td>Very Satisfied</td>
<td>Very Satisfied</td>
<td>66.9%</td>
</tr>
<tr>
<td>Ability to Answer Questions</td>
<td>Very Satisfied</td>
<td>Very Satisfied</td>
<td>63.8%</td>
</tr>
<tr>
<td>Problems Locating/Getting Services</td>
<td>No</td>
<td>No</td>
<td>96.8%</td>
</tr>
<tr>
<td>Rate Community as Place to Live</td>
<td>Good</td>
<td>Good</td>
<td>45.8%</td>
</tr>
<tr>
<td>Describe Race Relations</td>
<td>Good</td>
<td>Fair</td>
<td>41.3%</td>
</tr>
<tr>
<td>Change in Race Relations</td>
<td>Stayed the same</td>
<td>Stayed the same</td>
<td>45.8%</td>
</tr>
<tr>
<td>Trust State Government</td>
<td>Some of the Time</td>
<td>Some of the Time</td>
<td>52.9%</td>
</tr>
<tr>
<td>State Does Too Much</td>
<td>Should do More</td>
<td>Should do More</td>
<td>57.8%</td>
</tr>
<tr>
<td>Education by Public Schools</td>
<td>Good</td>
<td>Good</td>
<td>35.7%</td>
</tr>
<tr>
<td>Education by Local Schools</td>
<td>Good</td>
<td>Good</td>
<td>41.2%</td>
</tr>
<tr>
<td>Health Care to Poor</td>
<td>Fair</td>
<td>Fair</td>
<td>35.1%</td>
</tr>
<tr>
<td>Death Penalty for Murder</td>
<td>Favor</td>
<td>Favor</td>
<td>68.6%</td>
</tr>
<tr>
<td>Affirmative Action Programs</td>
<td>Needed</td>
<td>Needed</td>
<td>54.1%</td>
</tr>
<tr>
<td>Services of Health Department</td>
<td>Do Not Know</td>
<td>Do Not Know</td>
<td>61.4%</td>
</tr>
<tr>
<td>Belong to Club or Organization</td>
<td>Yes</td>
<td>Yes</td>
<td>67.1%</td>
</tr>
<tr>
<td>Served as Officer/On Committee</td>
<td>No</td>
<td>No</td>
<td>67.3%</td>
</tr>
<tr>
<td>Race of Club Members</td>
<td>Most</td>
<td>Most</td>
<td>41.2%</td>
</tr>
<tr>
<td>Sex of Club Members</td>
<td>Some</td>
<td>Some</td>
<td>46.1%</td>
</tr>
<tr>
<td>Currently Registered to Vote</td>
<td>Yes</td>
<td>Yes</td>
<td>83.5%</td>
</tr>
</tbody>
</table>
REFERENCES


*Public Perspective* 12 (March/April): 40-41.


Appendix

Heard of DHEC. Have you ever heard of the State Department of Health and Environmental Control or DHEC?

1. YES
2. NO
3. DON'T KNOW

Impression of DHEC. In general, what is your impression of DHEC ... would you say it is very positive, somewhat positive, neither positive nor negative, somewhat negative, or very negative?

1. VERY POSITIVE
2. SOMewhat POSITIVE
3. NEITHER POSITIVE NOR NEGATIVE
4. SOMEWHAT NEGATIVE
5. VERY NEGATIVE
6. DON'T KNOW

Used DHEC Services. The Department of Health and Environmental Control provides a number of environmental and health-related services such as private well testing, septic tank permitting, immunizations, women's and children's health services at local health departments, birth and death certificates, and disease investigation. Have you EVER used any of the services provided by DHEC?

1. YES
2. NO
3. DON'T KNOW

Quality of Service. In general, how satisfied were you with the quality of service you received during your last contact with or visit to DHEC ... would you say you were very satisfied, somewhat satisfied, somewhat dissatisfied, or very dissatisfied?

1. VERY SATISFIED
2. SOMEWHAT SATISFIED
3. SOMEWHAT DISSATISFIED
4. VERY DISSATISFIED
5. DOES NOT APPLY TO LAST CONTACT (VOL.)
6. DON'T KNOW
Satisfied with Wait Time. How satisfied were you with the time you had to wait for service (during your last contact with DHEC) … very satisfied, somewhat satisfied, somewhat dissatisfied, or very dissatisfied?

1. VERY SATISFIED
2. SOMEWHAT SATISFIED
3. SOMEWHAT DISSATISFIED
4. VERY DISSATISFIED
5. DOES NOT APPLY TO LAST CONTACT (VOL.)
6. DON’T KNOW

Courtesy and Attitude. How satisfied were you with the courtesy and attitude of the staff (during your last contact with DHEC) … very satisfied, somewhat satisfied, somewhat dissatisfied, or very dissatisfied?

1. VERY SATISFIED
2. SOMEWHAT SATISFIED
3. SOMEWHAT DISSATISFIED
4. VERY DISSATISFIED
5. DOES NOT APPLY TO LAST CONTACT (VOL.)
6. DON’T KNOW

Ability to Answer Questions. Overall, how satisfied were you with the staff’s ability to answer questions and provide information (during your last contact with DHEC) … very satisfied, somewhat satisfied, somewhat dissatisfied, or very dissatisfied?

1. VERY SATISFIED
2. SOMEWHAT SATISFIED
3. SOMEWHAT DISSATISFIED
4. VERY DISSATISFIED
5. DOES NOT APPLY TO LAST CONTACT (VOL.)
6. DON’T KNOW

Problems Locating/Getting Services. Did you have any problems locating or getting DHEC’s services?

1. YES
2. NO
3. DON’T KNOW

Rate Community as Place to Live. How would you rate your community as a place to live … would you rate it as excellent, good, only fair, or poor?

1. EXCELLENT
2. GOOD
3. ONLY FAIR
4. POOR
5. DON’T KNOW
*Describe Race Relations.* How would you describe race relations in the state today... would you say they are excellent, good, fair or poor?

1. EXCELLENT
2. GOOD
3. FAIR
4. POOR
5. DON’T KNOW

*Change in Race Relations.* In the last two years do you think that race relations in the state have gotten much worse, somewhat worse, have stayed about the same, have gotten somewhat better, or have gotten much better?

1. GOTTEN MUCH WORSE
2. SOMEWHAT WORSE
3. STAYED ABOUT THE SAME
4. GOTTEN SOMEWHAT BETTER
5. GOTTEN MUCH BETTER
6. DON’T KNOW

*Trust in Government.* And what about state government … How much of the time do you think you can trust the state government … just about always, most of the time, some of the time, or never?

1. JUST ABOUT ALWAYS
2. MOST OF THE TIME
3. SOME OF THE TIME
4. NEVER
5. DON’T KNOW

*State Does Too Much.* Some people think that state government does too many things that should be left to individuals and private businesses. Others think that state government should do more to help solve the state’s problems. Which of these do you agree with more (1) State government does too many things that should be left to individuals or private businesses; – OR – (2) State government should do more to help solve the state’s problems?

1. STATE GOVERNMENT DOES TOO MANY THINGS
2. STATE GOVERNMENT SHOULD DO MORE
3. DON’T KNOW
Education by State Public Schools. How would you rate the quality of education being given by public schools in the state ... would you say it is excellent, good, fair, poor, or very poor?

1. EXCELLENT
2. GOOD
3. FAIR
4. POOR
5. VERY POOR
6. DON’T KNOW

Education by State Local Schools. How would you rate the quality of education being given by your local public schools ... would you say it is excellent, good, fair, poor, or very poor?

1. EXCELLENT
2. GOOD
3. FAIR
4. POOR
5. VERY POOR
6. DON’T KNOW

Health Care to Poor. How would you rate state government for the way it goes about providing health care to the poor ... would you say this is excellent, good, fair, poor, or very poor.

1. EXCELLENT
2. GOOD
3. FAIR
4. POOR
5. VERY POOR
6. DON’T KNOW

Death Penalty for Murder. On a different topic ... Do you favor or oppose the death penalty for persons convicted of murder?

1. FAVOR
2. OPPOSE
3. DEPENDS
4. DON’T KNOW

Affirmative Action. Do you think affirmative action programs are needed today to help minorities overcome discrimination or do you think such programs are not needed?

1. AFFIRMATIVE ACTION PROGRAMS NEEDED
2. PROGRAMS NOT NEEDED
3. DON’T KNOW
Services of Health Department. Are the services of the health department in your area provided by the county or by the Department of Health and Environmental Control (DHEC), or are you not sure who provides health department services in your area?

1. COUNTY
2. DHEC
3. OTHER
4. DON’T KNOW

Belong to Club or Organization. Do you belong to any civic, church, business, or other type of club or organization in your community?

1. YES
2. NO
3. DON’T KNOW

Served as Officer/On Committee. In the past twelve months, have you served as an officer or served on a committee of any local club or organization in your community?

1. YES
2. NO
3. DON’T KNOW

Race of Club Members. Of all the groups that you are involved with in your community, please think of the one that is most important to you and about the members of this group you are involved with. About how many would you say are the same race as you – all, most, some, only a few, or none?

1. ALL
2. MOST
3. SOME
4. ONLY A FEW
5. NONE
6. DON’T KNOW

Sex of Club Members. About how many of this group are (men/women) … all, most, some, only a few, or none?

1. ALL
2. MOST
3. SOME
4. ONLY A FEW
5. NONE
6. DON’T KNOW
Currently Registered to Vote. Some people are registered to vote and others are not. Are you currently registered to vote in this state?

1. YES
2. NO
3. DK