THE ROLE OF TOPIC INTEREST IN SURVEY PARTICIPATION DECISIONS

ROBERT M. GROVES
STANLEY PRESSER
SARAH DIPKO

Abstract While a low survey response rate may indicate that the risk of nonresponse error is high, we know little about when nonresponse causes such error and when nonresponse is ignorable. Leverage-salience theory of survey participation suggests that when the survey topic is a factor in the decision to participate, noncooperation will cause nonresponse error. We test three hypotheses derived from the theory: (1) those faced with a survey request on a topic of interest to them cooperate at higher rates than do those less interested in the topic; (2) this tendency for the "interested" to cooperate more readily is diminished when monetary incentives are offered; and (3) the impact of interest on cooperation has nonignorability implications for key statistics. The data come from a three-factor experiment examining the impact on cooperation with surveys on (a) five different topics, using (b) samples from five different populations that have known attributes related to the topics, with (c) two different incentive conditions.

Introduction

Response rates have traditionally been used as indicators of survey quality, based on the fact that nonresponse error is partially a function of the response rate. Recently, three studies have challenged this practice (Curtin, Presser, and Singer 2000; Keeter et al. 2000; Merkle and Edelman 2002). Each found little relation between variation in response rates and changes in nonresponse error.

ROBERT M. GROVES is at the University of Michigan and the Joint Program in Survey Methodology (JPSM). STANLEY PRESSER and SARAH DIPKO are at the University of Maryland and the JPSM. The authors thank the interagency advisory committee for the 2000 Joint Program in Survey Methodology practicum (B. K. Atrostic, Susan Schecter, and Clyde Tucker), the JPSM students in the practicum, Steve Heeringa, Roger Tourangeau, Sonja Ziniel, and two anonymous reviewers, all of whom were helpful in preparing the article. Address correspondence to Robert M. Groves (e-mail: bgroves@isr.umich.edu) or Stanley Presser (e-mail: spresser@survey.umd.edu).

Public Opinion Quarterly, Vol. 68 No. 1 Pp. 2–31, DOI:10.1093/poq/nfh002

Since nonresponse error in survey statistics such as the unadjusted sample mean is a function of the response rate and the nonrespondents' characteristics, these studies suggest that respondents and nonrespondents often do not differ on commonly estimated statistics. Thus, research is needed to identify the circumstances under which nonresponse produces nonresponse error. Such identification will be facilitated by theories that distinguish between causes of participation that are related to statistics produced by a survey and causes that are unrelated to those statistics.

One effort along these lines is leverage-salience theory (Groves, Singer, and Corning 2000), which posits that people vary in the importance they assign to different aspects of a survey request. For example, for some individuals, the topic may be important; for others, whether a reputable organization is conducting the survey may be significant; and for still others, a chance to receive a cash reward may be of consequence. According to the theory, the influence of each component of the request depends both on the weight accorded it by a sampled individual (leverage) and on its prominence in the request protocol (salience).

The idea that people differ in their reactions to the features of a survey request is consistent with a diverse set of findings: for instance, that monetary incentives have different effects for different groups (Kulka 1994; Singer 2002), that interviewers who tailor their introductions to the concerns of the respondent obtain higher response rates (Groves and Couper 1998; Morton-Williams 1993), that some individuals respond more readily to a survey conducted by a government agency than by a university (National Academy of Sciences 1979), and that people with more involvement with a survey topic respond at higher levels than those with less involvement (Goyder 1987).

Of the many features of a survey request, we believe that subject matter, or topic, is particularly likely to lead to "nonignorable" nonresponse (Rubin 1987), that which produces nonresponse error. This follows from the fact that people more interested in the topic tend to have attributes on the key survey variables different from those of people less interested in the topic. Hence, statistics computed on variables central to the topic are apt to be among those most susceptible to nonresponse error, especially when the topic is made salient in the recruitment protocol.

Leverage-salience theory does not simply predict that persons interested in the survey topic will be overrepresented among respondents (and underrepresented among nonrespondents), relative to those uninterested. It predicts that the degree of overrepresentation will be a function of the salience of (and attitude toward) the survey topic among those deciding whether to cooperate, relative to the salience of (and attitudes toward) the other factors that are part of the survey request. If there are no other positive features to participation, the effect of topic should dominate the decision. If there are other influences toward participation, then the topic effect should be diminished. Indeed, if the information considered by the sample persons does not include the survey

topic, we might expect little difference between respondents and nonrespondents on variables related to the topic.

This perspective on the survey participation decision is compatible with a major social psychological theory of persuasion (Petty and Cacioppo 1986). The theory specifies two principal routes to decision making in reaction to a persuasion attempt. If the topic is of interest, people tend to engage in extensive cognitive processing of the message itself. If the topic is not of interest, people tend to rely on peripheral cues (e.g., characteristics of the speaker as opposed to the message) to make a judgment. This suggests, as does leverage-salience theory, that people uninterested in the topic of a survey are more apt to be influenced by other aspects of the survey request (e.g., the appeal of the interviewer or the offer of a monetary incentive) in deciding whether to participate.

Such a pattern has in fact been reported by Baumgartner and Rathbun (1997), who found that a monetary incentive increased cooperation more among those less interested in the survey topic than it did among the more interested, and by Groves, Singer, and Corning (2000), who found that such an incentive affected cooperation with a community survey more among those who were uninvolved in community affairs than among others. Thus, we expect that a survey on any specific topic will recruit more persons uninterested in the topic when it offers a monetary incentive. In other words, the addition of a monetary incentive (a non-topic-related reason for participation) should dampen the link between interest in the topic and cooperation.

The ideal research design to test this hypothesis would pose a series of independent survey requests to persons whose interest levels on the survey topics were known. The survey requests would make salient only one attribute of the survey—its topic; no other features of the request (e.g., sponsor, length of interview, interviewer characteristics) would be apparent. One request would be for a survey on the topic of greatest interest; another request would be for a survey on the topic of second-most interest; and so on. All the survey requests would be repeated in independent trials with and without a monetary incentive. We would hypothesize that the propensity to cooperate would be a positive function of both topic interest and incentive level and that there would be a negative interaction effect between topic interest and incentive.

The challenges of this ideal design are essentially insurmountable. Truly independent repeated survey requests of the same subject are impossible, so a between-subjects design is needed. Direct measures of people's interest sets are difficult to obtain, so indirect measures, such as membership on a list that suggests an interest, are a logical compromise (in essence using a binary indicator of interest, although it is inherently a continuous variable). Making topic the only feature of a survey that is salient in the introduction is impractical (given common guidelines on informed consent), so other features, such as interview length, sponsor, and the interviewer will inevitably have effects. Each of these departures from the ideal design makes it more difficult to observe the hypothesized relationship between topic interest and response propensity.

Thus, the test of the hypothesis we will report in this article is likely to be a conservative one.

We carried out a randomized factorial experimental design that exposed different populations with known characteristics to telephone survey requests on different topics. To test the hypothesized moderating effects of incentives, random half-samples were sent monetary incentives via an advance letter. We chose survey topics that appeared related (i.e., of interest) to the different populations. We hypothesized that when the topic of the survey was related to known characteristics of sample persons, their response rates would be higher than when the topic of the survey was not directly related to those characteristics. Further, we tested whether offering a prepaid incentive would reduce the effect of topic interest. Finally, we measured the impact of these factors on estimates from the survey, thereby assessing the nature of the nonresponse error associated with response rate differences.

Research Design

SAMPLE DESIGN

The sample for the study was drawn from five separate frames: four list samples representing four different groups, and a random digit dial (RDD) sample intended as a control (to contrast response propensities of the list frames to those of the general telephone household population). We hoped to obtain a total of 2,500 completed interviews, with approximately 500 from each frame, and thus we designed the sample to yield about 4,000 eligible households, 800 from each of the frames (see table 1).

There were four sample replicates released during the course of data collection. Each of the four replicates consisted of two randomly assigned parts: letter and non-letter. For cases assigned to the letter condition, advance letters with a study description and a five-dollar bill were mailed one week prior to the sample's release. Letters were mailed to one-half of the households from the four list samples and to one-half of the households from the RDD frame for which mailing addresses could be obtained. The letters mentioned only the survey topic that matched the randomly assigned introduction condition for the household. In retrospect, we realize that mentioning the survey topic in the letter could have indirectly acted to increase the salience of the topic. Thus, our test of the hypothesis that monetary incentives reduce the impact of salience on cooperation is a conservative one. (In all tables and analyses in this paper, RDD cases without mailing addresses are omitted.)

Given the design of the experiment, the selection of the frames suggested the topics of the survey. The four sampling frames were elementary and secondary schoolteachers (for whom we assumed an interest in education and

Table 1. Sample Sizes by Sampling Frame and by Survey Topic

	Sampling Frame					
Survey Topic	Teachers	New Parents	65 and Older	Political Contributors	RDD (control)	Total
Education and Schools	173	184	172	175	115	819
Child Care and Parents	174	183	172	175	116	820
Medicare and Health	172	182	172	175	112	813
Voting and Elections	172	181	172	176	98	799
Issues Facing the Nation (control)	175	182	171	177	114	819
Total	866	912	859	878	555	4070

schools), parents of children 0–6 months old (for whom we assumed an interest in child care and problems of parents), contributors to several presidential candidates in the year 2000 (for whom we assumed an interest in voting and participation in elections), and people aged 65 and older (for whom we assumed an interest in Medicare and health). The fifth, RDD, sample was selected from a one-plus list-assisted frame, purchased from Genesys. A fifth survey topic, assumed to be relatively content-free, was used as a control topic (i.e., "important issues facing the nation"). All five sample frames were restricted to the contiguous United States. The five topics—education and schools, child care and problems of parents, voting and participation in elections, Medicare and health, and important issues facing the nation—were randomly assigned across the five samples in equal proportions. That is, subsamples from all five frames were assigned at random to the five different topics.

Although the goal for the four list samples was to speak to the person on the list, we wanted to make the within-household selection process appear random to both interviewers and respondents. This was done to avoid questions that might arise from asking for a named individual and to blind interviewers to the origin of the sample. To accomplish this for the four list samples, the gender of the target respondent was identified in advance; when no such determination was possible, a random assignment of gender was used. Random assignment of gender was also used for the entire RDD sample. In the actual selection procedure, the interviewers asked to interview either a male or female head of the household. In the event that there was no such person, interviewers were allowed to switch and ask for the other gender.

ELEMENTARY AND SECONDARY SCHOOLTEACHERS

A frame of schoolteachers was obtained from a commercial data base list broker, Direct Mail and Computer Systems. The list owner creates this frame every year in the early spring, starting with a list of people who purchase products for classroom materials. Information from other sources is merged into the list to compile a "teachers at home" database. Our sample was drawn from a list with 955,001 entries: 555,403 from elementary schools and 399,598 from high schools. Although it was not possible to specify public versus private school as part of the sample order, the list broker claimed that public school teachers comprised approximately 95 percent of the database. (Later analysis of respondent characteristics from the survey indicated that the frame included some school employees who were not teachers.)

The database owner was asked to sort the list by school type (primary versus secondary) and the first three zip code digits before selecting 10,000 records (the minimum size order). Following the selection, records without telephone numbers were sent to Donnelly Marketing to have that information appended. Then prior to drawing the final sample from the 5,033 teacher records with names, addresses, and telephone numbers, the file was sorted by school type (primary versus secondary), gender (if available), and first three zip code digits.

PARENTS OF CHILDREN 0–6 MONTHS OLD

A list of parents of recently born children was also accessed through Direct Mail and Computer Services. It consisted of 1.1 million households with children 0–6 months old and was compiled from public birth records as well as from requests for diaper service and other products for infants. Some records included only the mother's name, others included only the father's name, and some contained both.

We instructed the source to sort the data base by the newborn's date of birth (month and year) and first three zip code digits. Following the selection of 10,000 records (the minimum size order), those without telephone numbers were sent to Donnelly Marketing to have that information appended. Before selecting the final sample from the 4,861 records that had telephone numbers, the file was sorted by the child's date of birth (month and year), gender of parent's name (male name, female name, or both), and first three zip code digits.

PERSONS 65 AND OLDER

The source for the 65 and older frame was Survey Sampling, Inc. (SSI), which maintains a database of listed telephone households with information on age, drawn from sources such as driver's license, motor vehicle, and voter registration records. The SSI frame consisted of 369,868 households (in the contiguous United States), with one or more persons 65 and older.

We ordered 2,500 records to ensure adequate sample size due to an expected higher nonworking and nonresidential rate than for the other list samples. The sample was drawn using SSI's standard "probability proportional to size" by state selection procedures. Prior to drawing the subsample for the experiment, the 2,500 records were sorted by gender, date of birth, and first three zip code digits.

CAMPAIGN CONTRIBUTORS TO NON-MAINSTREAM U.S. PRESIDENTIAL CANDIDATES IN THE YEAR 2000

Presidential candidates who wish to obtain federal matching funds are required by the U.S. Federal Election Commission (FEC) to report information on contributors to their campaigns. This information is posted quarterly on the FEC Web site, and it includes each individual contributor's full name, address, occupation, and employer, as well as the date and dollar amount of the contribution and a cumulative total of contributions by that individual.

We wanted to sample contributors with high interest in politics and strong commitments to their political beliefs. We expected the greatest intensity of interest would be among contributors to candidates with little chance of winning. For this reason we sampled only contributors to non-mainstream presidential candidates. The candidates chosen were Pat Buchanan, Alan Keyes, and Lyndon Larouche, thus ensuring a mix of contributors from different parts of the political spectrum.

The frame consisted of 9,350 individual records with full names and mailing addresses. Nearly half (43.5 percent) were contributors to Buchanan, roughly one-third (31.0 percent) were Keyes contributors, and the remainder (25.4 percent) were Larouche contributors. The names and addresses of all records were submitted to Telematch to obtain telephone numbers. Over half of the records (60.5 percent) yielded a telephone number, leading to a final sampling frame of 5,656 individuals. The telephone match rate did not differ substantially by candidate, so the final sampling frame was 42 percent Buchanan contributors, 32.5 percent Keyes contributors, and 25.5 percent Larouche contributors. Prior to drawing the sample, the frame was sorted by presidential candidate, quarters in which contributions were made, and the contributor's first three zip code digits.

RANDOM DIGIT DIAL SAMPLE

A one-plus list-assisted RDD sample of 2,000 telephone numbers was obtained from Genesys, and all records were submitted to Telematch to obtain addresses.

QUESTIONNAIRE STRUCTURE

The samples from each frame were divided into five equal-sized subsamples. Each subsample was assigned to a different survey introduction. The voting introduction, for example, was:

> Hello, my name is _____ from the University of Maryland Survey Research Center. Here at the university we are doing a study on important issues such as *voting and participation in elections.*
>
> We are gathering information on people's experiences and opinions about *voting and participation in elections.*
>
> For your household we have randomly selected the [male/female] head of the household. May I please speak with [him/her]?

The four other introductions replaced "voting and participation in elections" with "education and schools," "child care and problems of parents," "Medicare and health," or "important issues facing the nation."

A key design decision was how salient to make the topic of the survey. While the force of the experimental manipulation of topic could have easily been increased by having the interviewer give more elaborate descriptions, we chose to create the kinds of brief introductions that are commonly used in household telephone surveys.

The order in which the questions about the four topics were asked was varied by condition. For the four specific topic survey introductions, the questionnaire's opening section matched the introduction topic. When the "issues facing the nation" version of the introduction was used, the questionnaire was randomly assigned to start with one of the four topical sections.

INTERVIEWERS

The interviewers included twelve first-year students from the Joint Program in Survey Methodology (JPSM) master's program as well as the regular interviewing staff of the University of Maryland Survey Research Center (SRC). Roughly 10 percent of all call attempts and 15 percent of completed interviews were conducted by the JPSM students during the first three weeks of March 2000. The regular SRC interviewers were primarily undergraduate students, and their work was conducted over the course of three months, from mid-February through early May 2000.

Final Sample Dispositions and Response Rates

Of the 4,070 total telephone numbers from all five sample frames, 369 were non-households, including businesses and nonworking numbers. Non-household numbers ranged from a low of 6.4 percent among the contributor sample to 12.1 percent among parents of newborns, compared to 17.1 percent among RDD cases.

Of the 3,701 remaining telephone numbers, 835 were refusals and break-offs, 181 posed difficulties such as a language other than English, illness, or hearing problems, 311 were noncontacts in which a respondent was never

selected (e.g., the telephone was always answered by a household answering machine) or the selected respondent was never reached for an interview, and 75 were never answered after at least 20 callbacks. A total of 2,330 interviews were completed (of which 31 were partial interviews), leading to an overall response rate of between 63.0 and 63.4 percent, depending on the proportion of never-answered numbers that are treated as eligible. The separate results for each sample frame are shown in table 2.

Compared to the special population list frames, the RDD frame suffers from a higher noncontact rate (due to the indeterminacy of never-answered numbers), and thus the cooperation rate (AAPOR COOP2) is a useful comparative measure. The RDD cooperation rate of 59.7 percent is similar to that of the 65 and older frame (61.6 percent), both of which are much lower than those of the other frames. The political contributors exhibit the highest cooperation rates (79.5 percent).

Results

THE DEPENDENT VARIABLE

Cases in which the introduction was never heard are not likely to show an effect of our main experimental manipulation. Consequently, our main analyses include only those cases where a contact was made with someone in the household, and the interviewer was able to deliver the introduction at least as far as mentioning the survey topic. The dependent variable is the outcome of the first contact with the sampled household during which the survey topic was mentioned. (We chose that outcome as one that is minimally affected by differential skills of interviewers and other potential influences not key to the research design.) The variable is coded "1" if the interview was begun on that call, and "0" if it was not. Thus, completed and partial interviews on the first contact are contrasted with other outcomes of that contact such as refusals, appointments, and unavailable respondents. (Cases in which all contacts with the household yielded a language barrier or other respondent problem, such as deafness, are excluded from our analyses.) The fact that the first contact in which the introduction was read may have occurred with a household member other than the targeted respondent is another reason our tests of the salience hypothesis are conservative.

CONTRASTING FRAMES, TOPICS, AND INCENTIVE TREATMENTS

A two-factor display of the dependent measure, the cooperation rate at first exposure to the stimulus, is presented in table 3 for the five frame populations and the five survey topics. The table contains both the final response rates for each experimental treatment cell and the percentage of sample persons who

Table 2. Percentage in Final Disposition Categories by Sampling Frame (Pooled over 5 Survey Topics)

	Sampling Frame					
					RDD (control)	
Final Result	Teachers	New Parents	65 and Older	Political Contributors	$e=1$[a]	$e=0$[a]
Complete interviews (%)	66.0	64.0	56.6	73.6	48.3	50.9
Refusals (%)	23.6	20.0	27.2	14.1	25.4	26.8
Noncontact[a] (%)	9.2	10.9	8.2	8.5	19.1	14.7
Other (%)	1.3	5.2	8.0	3.8	7.2	7.6
Total (%)	100	100	100	100	100	100
N	802	802	815	822	460	436
AAPOR COOP2 rate (%)	72.7	71.8	61.6	80.5	59.7	

[a] For the four list frames, the "noncontact" category includes telephone numbers that were never answered by a human or machine. For the RDD frame, two distributions are shown: $e=1$ using the AAPOR guidelines notation, in which all unanswered phone numbers are counted as eligible, and $e=0$ in which none of the unanswered phone numbers are counted as eligible.

Table 3. Response Rates (RR) and Rate of Cooperation on First Hearing of Study Introduction (CR) by Survey Topic and Sampling Frame

	Sampling Frame									
	Teachers		New Parents		65 and Older		Political Contributors		RDD (control)	
Survey Topic	RR	CR	RR	CR	RR	CR	RR	CR	RR[a]	CR
Education and schools	**73.8**	**55.9**	60.0	34.7	50.0	42.3	82.5	62.5	41.1 44.1	31.9
Children and parents	64.2	44.0	**69.7**	**45.1**	60.5	41.7	64.8	49.0	49.0 50.5	37.2
Medicare and health	64.2	39.5	60.5	39.9	**67.5**	**52.4**	73.6	54.4	57.3 59.8	31.7
Voting and elections	70.8	45.4	67.7	34.2	50.0	38.8	**73.3**	**52.3**	45.2 48.1	32.8
Issues facing the nation (control)	56.6	39.2	61.5	34.0	55.0	39.5	73.5	55.0	47.8 51.2	36.4
Average of all topics	66.0	44.8	64.0	37.6	56.6	42.9	73.6	54.8	48.3 50.9	34.1
Average of irrelevant topics	63.9	42.0	62.4	35.6	53.9	40.6	73.6	55.4		
Difference (due to topic interest)	9.9	13.9	7.3	9.5	13.6	11.9	–0.3	–3.1		
SE	(5.12)	(5.69)	(5.25)	(5.80)	(5.37)	(5.80)	(4.88)	(5.58)		

NOTE.—Results for relevant topic survey are in boldface.

[a] The first response rate for the RDD cases was calculated using $e = 1$, whereas the second is based on $e = 0$, using the AAPOR guidelines notation.

cooperated on the same contact on which the survey introduction was first given. The political contributors had the highest average cooperation rate among the different frame populations (54.8 percent) and the RDD frame, the lowest (34.1 percent). The various survey topics exhibit smaller differences in level of cooperation. "Education and schools" has the highest rate (47.5 percent); "voting and elections" and the control topic "issues facing the nation" have the lowest (about 41 percent). These main effects of frame and topic are not the focus of our research, but they suggest how response rates differ across common topics and populations.

Table 3 permits an initial look at our first hypothesis. The table row "Average of irrelevant topics" shows the average cooperation among the four survey topics judged *not* relevant to the given frame. For example, for the sample of teachers, the average cooperation rate for the four other survey topics was 42.0 percent, compared to 55.9 percent for the "education and schools" topic ($p < .05$). The 13.9 percentage point difference supports the hypothesis that persons will cooperate with surveys on topics that interest them more than on other topics. Similar support for the topic interest hypothesis appears for new parents and people 65 and older. The political contributors, possibly because all of the topics may be seen as having political content (admittedly a post hoc hypothesis), show no preference for the "voting and elections" topic over the others.

Leverage-salience theory predicts that the treatment effects should be smaller for the response rate (which reflects the impact of callbacks and refusal conversion designed in part to offset respondents' lack of interest in the topic) than for the cooperation rate at the first contact. For the teacher frame, as expected, the differences in *response* rates due to topic relevance are smaller than the differences in *cooperation* rates (9.9 percentage points versus. 13.9 percentage points; $t = 9.55$, $df = 748$, $p < .001$). For other frames the evidence is more mixed. This could reflect a lack of diminished salience of the topic stimulus, relative to the effects of callbacks and interviewer persuasion, which involve survey features in addition to the topic (e.g., sponsorship, importance of respondent cooperation, interviewer professionalism).

THE EFFECTS OF MONETARY INCENTIVES

We expected the topic interest effects we observed for three of the four frames to be diminished by the provision of the $5 incentive. Table 4 shows that this occurs in two of the three cases: teachers and individuals 65 and older. Surprisingly, however, the reverse occurs for new parents.

MULTIVARIATE MODELS

We then constructed multivariate models that simultaneously included the effects of topic and frame membership in an equation predicting the likelihood of cooperation at the first contact. The specification of the model is

$$\ln\left(\frac{p}{1-p}\right) = \beta_0 + \sum_{k=1}^{4} \beta_k Frame_k + \sum_{l=5}^{8} \beta_l Topic_l + \sum_{m=9}^{12} \beta_m Frame, TopicMatch_m + \sum_{n=13}^{21} \beta_n Frame, TopicMismatch_n + \varepsilon$$

where p is the probability of providing an interview on the first contact, $Frame_k$ denotes membership in the k-th frame (omitting the RDD frame as the contrast group), $Topic_l$ denotes implementation of the l-th survey topic (with "issues facing the nation" as the omitted group), $Frame, TopicMatch_m$ equals 1 when each of the four frame samples was assigned to the "matched" survey topic, and $Frame, TopicMismatch_n$ equals 1, otherwise. The key hypothesis is that β_m is greater than the corresponding β_n's for each of the four list frames. (In addition, we included as control variables three other predictors of cooperation for which we had information about both respondents and nonrespondents: urbanicity, region, and sex.)

We fit the model in two steps. First, we examine whether, in the presence of the multivariate controls, the effect of the topic interest indicators remains. These results are listed in the column of table 5 labeled "topic interest model." The test statistic for the multiple contrast hypothesis involving all the $\beta_m > \beta_n$ contrasts is significant at the .001 level ($\chi^2 = 10.96$, $df = 1$). The estimated increase in the odds of cooperating with a relevant topic (pooling all cases of relevant topic together) versus all others is 38 percent (a marginal odds ratio of 1.378). Thus, the effect of topic interest on this indicator of cooperation is substantial.

Step two of the model-fitting process addresses the second hypothesis: that when incentives are offered, the higher cooperation rates of those interested in

Table 4. Percentage Cooperating at First Contact by Frame by Topic Interest by Incentive

	No Prepaid Incentive			Prepaid Incentive		
Sampling frame	Interesting Topic	Other Topic	Difference	Interesting Topic	Other Topic	Difference
Teachers	49.3	33.4	15.9	62.3	50.7	11.7
New parents	32.9	29.3	3.6	57.1	42.0	15.1
65 and older	46.6	27.8	18.8	58.3	52.5	5.9
Political contributors	50.7	49.0	1.6	54.0	61.9	–7.9

Table 5. Odds Ratios for Parameters in Three Logit Models of Cooperation: Topic Interest Model, Interaction Model of Interest and Incentives (Full Sample), and Interaction Model of Interest and Incentives (Excluding Young Parents)

		Interaction Model of Interest and Incentive	
	Topic Interest Model	Full Sample	Excluding Parents Frame
Covariates			
Nonrural	0.75**	0.75**	0.81**
Midwest	1.19	1.20	1.27**
South	0.98	0.99	1.10
West	1.14	1.15	1.18
Female	1.10	1.10	0.80**
Topic			
Education	1.20	0.99	1.03
Child Care	1.04	0.98	1.10
Politics	0.93	0.89	0.92
Medicare	1.09	0.97	0.93
Frame			
Teachers	1.50**	1.95**	1.93**
New Parents	1.13	1.47	—
Contributors	2.36**	3.33**	3.14**
65 and Older	1.42**	1.59**	1.55**
Topic Interest	1.38**	1.54**	1.75**
Incentive	2.08**	2.73**	2.97**
Topic Interest × Incentive			
Inc/Education		1.42	1.46
Inc/Child Care		1.11	0.83
Inc/Politics		1.07	0.95
Inc/Medicare		1.22	1.20
Incentive × Frame			
Inc/Teachers		0.65	0.70
Inc/Parents		0.64	—
Inc/Contributors		0.54**	0.56**
Inc/65 and Older		0.85	0.89
Incentive × Topic Interest		0.81	0.59**

** $p < .05$.

the topic are diminished, reflecting the stronger effects of incentives on those not interested in the topic of the survey. This hypothesis is tested by fitting interaction terms that reflect the combined effects of incentive and topic interest.

$$\ln\left(\frac{p}{1-p}\right) = \beta_0 + \sum_{k=1}^{4} \beta_k Frame_k + \sum_{l=5}^{8} \beta_l Topic_l + \beta_9(Incentive = 1) +$$

$$\sum_{m=10}^{13} \beta_m Frame, TopicMismatch_m +$$

$$\sum_{n=14}^{22} \beta_n Frame, TopicMismatch_n +$$

$$\sum_{o=23}^{26} \beta_o(Incentive = 1) \times (Frame, TopicMatch) +$$

$$\sum_{p=27}^{35} \beta_p(Incentive = 1) \times (Frame, TopicMismatch) + \varepsilon$$

When the joint hypothesis test is performed (that for each list frame β_o is less than the corresponding β_p), the magnitudes of the incentive interaction effects do not exceed those expected from sampling variability alone ($\chi^2 = 12.07$, $df = 12$, $p = .44$; data not shown).

An examination of the different sampling frames bears out the bivariate result from table 4 that the new parents' frame is an exception with respect to the interaction. When coefficients associated with that frame are omitted, the interaction effect is more pronounced but still nonsignificant ($\chi^2 = 13.2$, $df = 9$, $p = 0.15$). We are not certain why the new parents exhibit effects that appear distinctive from those of the other frames. An after-the-fact inquiry to the vendor from whom we purchased our sample revealed that the new parents' list was a "best-seller," very popular for direct telemarketing to new parents. Thus, it is possible that telephone contacts with strangers about their new birth are disproportionately sales efforts; hence, a stranger requesting an interview relevant to the new birth is precisely what the new parents are *not* interested in having. In any case, this anomaly needs further investigation.

CONSEQUENCE FOR SURVEY ESTIMATES

To this point we have demonstrated that people who possess characteristics that predispose them to be interested in a particular topic are generally more likely to cooperate with a survey request when that topic is mentioned in the introduction. This has potential implications for nonresponse error of survey estimates. To the extent that nonresponse is caused by interest in the topic, survey results on matters related to interest in that topic should be biased by the overrepresentation of those with high interest (and corresponding underrepresentation of those with low interest). Thus, we hypothesized that the

distributions of answers to some questions asked in our survey would vary by introduction. The sample exposed to the voting introduction, for instance, ought to have expressed higher levels of political interest and involvement to the survey questions than the samples that heard the other introductions.

We tested this hypothesis with three kinds of items. First, toward the end of the interview we asked respondents which of our four topics they considered the most important problem facing the nation: improving education, health care for the elderly, child care for preschoolers, or increasing participation in elections. We hypothesized that choice of a topic should be greatest when it was mentioned in the survey introduction, on the assumption that ratings of a topic's importance are linked to interest in the topic.

Second, from each of the four topic sections of the questionnaire we chose one item that appeared to us most related to concern about the topic: attitude toward teacher salaries, opinion about the difficulty of finding child care, enrollment in Medicare, and attention to the 2000 election campaign. We expected that respondents would say more often that teacher salaries were too high or too low (as opposed to saying "about right" or having no opinion) in the survey about "education and schools"; express more beliefs that high-quality child care was difficult to find in the survey about "child care and problems of parents"; report highest Medicare enrollment in the survey on "Medicare and health"; and claim greatest attention to the campaign in the survey on "voting and election participation." Our assumption was that each of these characteristics is apt to be related to interest in the topic.

Finally, the interview contained questions that allowed us to identify whether respondents were members of the four populations from which we drew our list samples: Are you a teacher? Do you have a child 0 to 6 months old? When were you born? Did you contribute money to a political candidate? Our hypothesis was that membership in a population would be greatest with the corresponding topic introduction, again on the assumption that interest in the topic is promoted by membership on the frame.

Because we drew on answers from the interview, these analyses switched from the total sample (interviewed and never interviewed cases) that we used in preceding analyses to only those cases that were interviewed (either on the first contact or at a later point). Each test used the interviewed sample from four of the five frames: the sample from the frame relevant to a topic was excluded from analyses comparing that introduction to the others. For instance, in comparing the level of contributors (or political interest, or the view that political participation was the most important problem facing the nation) between the politics introduction and the other introductions, the sample from the contributors' frame was excluded. This was done because that frame's likely homogeneity on these variables meant that contributors would not be apt to show variation on such variables by introduction.

Table 6 shows the results of these tests. Eight of the twelve tests are in line with the hypothesis. For example, among the four noncontributor frames,

Table 6. Percentage of Respondents Reporting Salient Characteristics for Relevant Survey Topic, by Introduction Condition, Excluding the Relevant Frame for Each Topic

Salient Characteristic	Introduction Condition: Relevant Topic	*(n)*	Other Topics	*(n)*	Odds Ratio
Frame membership					
Contributed money to political candidate	15.6	*340*	13.2	*1,361*	1.18
Employed as a schoolteacher	2.6	*341*	2.6	*1,431*	1.00
Parent of newborn child (in past 12 months)	0.9	*349*	2.8	*1,445*	0.31[a]
Age is 65 or older	23.1	*359*	21.2	*1,456*	1.09
Most important problem facing the country today is					
Increasing participation in elections	5.6	*340*	4.4	*1,361*	1.27
Improving education	49.9	*341*	44.4	*1,431*	1.12**
Child care for preschoolers	6.3	*349*	7.5	*1,445*	0.84
Health care for the elderly	27.5	*367*	21.5	*1,483*	1.28**
Other survey reports					
Followed 2000 political campaigns very closely (vs. fairly, not too, and not at all)	19.7	*340*	18.1	*1,361*	1.09
Thinks teachers' salaries are too high or low (vs. about right or no opinion)	46.9	*341*	48.6	*1,431*	0.97
Thinks it's extremely or very difficult for most American families to find high-quality child care (vs. somewhat, not too difficult)	46.4	*349*	42.3	*1,445*	1.10*
Currently has Medicare coverage	25.9	*367*	24.3	*1,483*	1.07

[a] *t*-test shows a significant difference in the wrong direction, $p < .05$.
** One-tailed *t*-test shows a significant difference in the expected direction, $p < .05$.
* One-tailed *t*-test shows a difference of borderline significance in the expected direction, $.05 < p < .10$.

more respondents claim to have made a campaign contribution in the "voting and elections" condition than in the other conditions combined. On the other hand, only three of these eight effects are significant at the .05 level; and of the four results not in line with expectation, one is statistically significant. Thus, overall, there is support for the hypothesis, but it is weak.

These analyses, however, use data from all respondents, irrespective of whether they initially cooperated or required refusal conversion and irrespective of whether they received a monetary incentive. Yet there is less reason to expect the hypothesis to apply to respondents who were reluctant to participate after hearing the introduction or to those who received the incentive (who had a non-topic-rationale for participation). Thus, we repeated the analyses separately for groups differing in the ease with which the interview was obtained and by incentive condition. We expected that our hypothesis would receive stronger confirmation among those who were easy to interview and among those who did not receive the monetary incentive.

As can be seen in tables 7 and 8, there is not much support for the expectation that either interview ease or the monetary incentive affects the association between introduction and survey result. There is a little evidence of an effect of the introduction, but it is about the same for hard cases and easy ones, incentive cases and non-incentive ones.

The relative weakness of the results concerning nonresponse error may be partly due to departures from the ideal experimental design that we discussed in the introduction. If the motivation for cooperation is multifaceted, the hypothesized effects are likely to be attenuated. In the case of our design, heterogeneity in topic interest within each frame, the possible influences of sponsorship ("the University of Maryland Survey Research Center"), and the moderating effects of interviewer behavior may explain some of these results.

But there are also other reasons that the impact on survey estimates was not more pronounced. Two of our groups (teachers and recent parents) constitute only tiny fractions of the total population, and the other two (contributors and those 65 and older) are relatively small (on the order of 15 to 20 percent). Moreover, membership in these four populations is not consistently related to our survey variables. Of the eight variables in the lower two panels of tables 6–8 ("most important problem" choices and "other survey reports"), only four are significantly related to membership in the relevant population, and two of the four involve teachers, who comprise a tiny fraction of the total population. Thus, the effect of topic interest on cooperation does not translate into much of an effect on the survey's estimates.

AN ANALYTIC DERIVATION

To put the size of these effects into perspective, we examined the maximum possible impact on nonresponse error in the kind of experiment we mounted. The appendix derives analytically the bias of the difference between a statistic

Table 7. Percentage of Respondents Reporting Salient Characteristics for Relevant Survey Topic, by Introduction Condition, by Ease of Interview, with Relevant Frame Excluded for Each Comparison

	Easy to Interview[a]					Hard to Interview				
	Introduction Condition					Introduction Condition				
Salient Characteristic	Relevant Topic	*(n)*	Other Topics	*(n)*	Odds Ratio	Relevant Topic	*(n)*	Other Topics	*(n)*	Odds Ratio
Frame membership										
Contributed to political candidate	17.0	*159*	12.8	*694*	1.32	14.4	*181*	13.6	*661*	1.05
Employed as a schoolteacher	3.7	*191*	1.8	*718*	2.02	1.3	*150*	3.4	*713*	0.39[†]
Parent of newborn child	0.0	*186*	2.7	*779*	0.00[†]	1.8	*163*	2.9	*666*	0.65
Age is 65 or older	29.6	*169*	24.2	*755*	1.22*	17.4	*190*	17.8	*701*	0.97
Most important problem facing the country today is										
Participation in elections	5.7	*159*	4.9	*696*	1.16	5.5	*181*	3.9	*665*	1.41
Improving education	52.9	*191*	42.3	*718*	1.25**	46.0	*150*	46.6	*713*	0.99
Child care for preschoolers	5.4	*186*	8.1	*779*	0.66	7.4	*163*	6.8	*666*	1.09

Health care for the elderly	28.5	*172*	22.4	*764*	1.27**	26.7	*195*	20.6	*719*	1.30**
Other survey reports										
Followed 2000 political campaigns very closely (vs. fairly, not too, and not at all)	20.8	*159*	19.4	*696*	1.07	18.8	*181*	16.7	*665*	1.13
Thinks teachers' salaries are too high or too low (vs. about right or no opinion)	46.6	*191*	46.9	*718*	0.99	47.3	*150*	50.4	*713*	0.94
Thinks extremely or very difficult for most American families to find high-quality child care (vs. somewhat, not too difficult)	48.4	*186*	45.2	*779*	1.07	44.2	*163*	38.9	*666*	1.14
Currently has Medicare coverage	31.4	*172*	28.1	*764*	1.12	21.0	*195*	20.3	*719*	1.04

[a] Easy cases are those for which the interview was obtained on the first household contact, without questions from the respondent, and with no refusal in the case history.

[†] *t*-test shows a significant difference in the wrong direction, $p < .05$.

* One-tailed *t*-test shows a difference of borderline significance in the expected direction, $.05 < p < .10$.

** One-tailed *t*-test shows a significant difference in the expected direction, $p < .05$.

Table 8. Percentage of Respondents Reporting Salient Characteristics for Relevant Survey Topic, by Introduction Condition, by Incentive Condition, with Relevant Frame Excluded for Each Comparison

	Non-Incentive Cases					Incentive Cases				
	Introduction Condition					Introduction Condition				
Salient Characteristic	Relevant Topic	*(n)*	Other Topics	*(n)*	Odds Ratio	Relevant Topic	*(n)*	Other Topics	*(n)*	Odds Ratio
Frame membership										
Contributed to political candidate	11.8	*153*	13.9	*553*	0.84	18.7	*187*	12.7	*802*	1.47**
Employed as a schoolteacher	3.6	*140*	2.1	*616*	1.69	2.0	*201*	2.9	*815*	0.68
Parent of newborn child	2.0	*148*	3.4	*617*	0.60	0.0	*201*	2.3	*828*	0.00†
Age is 65 or older	22.7	*150*	22.3	*641*	1.02	23.4	*209*	20.3	*815*	1.16
Most important problem facing the country today is										
Participation in elections	7.8	*153*	5.2	*555*	1.50	3.7	*187*	3.8	*806*	0.97
Improving education	45.7	*140*	44.3	*616*	1.03	52.7	*201*	44.5	*815*	1.18**

Child care for preschoolers	6.1	*148*	7.1	*617*	0.85	6.5	*201*	7.7	*828*	0.84
Health care for the elderly	21.3	*155*	20.3	*656*	1.05	32.1	*212*	22.5	*827*	1.43**
Other survey reports										
Followed 2000 political campaigns very closely (vs. fairly, not too, and not at all)	19.0	*153*	19.5	*555*	0.97	20.3	*187*	17.1	*806*	1.19
Thinks teachers' salaries are too high or too low (vs. about right or no opinion)	50.0	*140*	48.4	*616*	1.03	44.8	*201*	48.8	*815*	0.92
Thinks extremely or very difficult for most American families to find high-quality child care (vs. somewhat or not too difficult)	44.6	*148*	41.7	*617*	1.07	47.8	*201*	42.8	*828*	1.12**
Currently has Medicare coverage	27.1	*155*	25.6	*656*	1.06	25.0	*212*	23.3	*827*	1.07

† *t*-test shows a significant difference in the wrong direction, $p < .05$.
* One-tailed *t*-test shows a difference of borderline significance in the expected direction, $.05 < p < .10$.
** One-tailed *t*-test shows a significant difference in the expected direction, $p < .05$.

from a survey about topic Y, which is of interest to a subset of the population, versus that from a survey about an unrelated topic. It asks the question, for estimates of the percentage of persons in the population with an interest in Y, what would be the bias difference between the result from the relevant topic survey and the irrelevant topic survey? The answer is that the bias is a function of the true percentage of persons with the interest *and* the difference in response rates to the two surveys for the interested population relative to the response rate difference for the uninterested. On our surveys, the corresponding response rate differences are approximately 10, 7, 14, and 0 percentage points (last row of table 3). With response rate differences generated by topic interest differences in this range, the impact on nonresponse error ***must*** be rather small. For the case where the true population percentage is 50 percent, for example, the 7–10 percentage point response rate differences due to topic interest produce biases of less than 5 percentage points around the 50 percent true value. If the population statistic is smaller, say 20 percent, then the bias induced by topic interest is in the range of 1–2 percentage points around the 20 percent. Thus, given our response rate differences, the impact on survey estimates of interest have to be small and will be even smaller for estimates of variables that are not perfectly correlated with interest.

Discussion

The results of the experiment were not uniform over frame and topic. There are at least three alternative explanations for this:

a. Leverage-salience theory is incorrect;
b. Our frames inadequately operationalized interest; and
c. Our survey introductions inadequately made topic a salient decision criterion.

The finding that political contributors cooperate with all the topics (relative to the RDD control group and other frames) raises the possibility that all our topics are of interest to them. This suggests that the operationalization of the topic set yielded a poor test of the hypotheses for political contributors. If we had used a topic of no political relevance, say, sports, the results might have been different, and this would be a useful step for further research. Alternatively, the high cooperation rates for all topics may imply that political contributors have high response rates on all survey topics. In short, they are consistent responders to surveys. The possibility that some subgroup places high positive leverage on the very act of participating in surveys could be incorporated into the theory. If large parts of the population were indifferent to all features of the survey request, however, it would contradict the basic foundation of the theory. Yet this is incompatible with the findings cited at the beginning of the article (that different persons respond under different conditions).

We believe that explanations (*b*) and (*c*) are more likely: the frames incompletely identify interest groups, and the survey introduction made the topic insufficiently salient. There are many influences on whether one responds, in addition to topic. Interviewer voices have effects (Oksenberg, Coleman, and Cannell 1986), sponsorship of the survey may act independently of topic (Groves and Couper 1998), and other situational factors act to attenuate the influence of topic on cooperation. This implies that larger effects might be obtained if the introduction had been altered to heighten the salience of topic relative to other factors.

The upshot for practicing survey researchers, however, is good news. We have provided a theoretical and analytic (in the appendix) rationale, with further empirical support, for the conclusions of Curtin, Presser, and Singer (2000), Keeter et al. (2000), and Merkle and Edelman (2002). Leverage-salience theory suggests that there are many diverse influences on survey participation. Only those influences linked to the survey statistics of interest need cause concern to the analyst. Survey designs that induce multiple influences (mode differences, incentives, alternative interviewers) tend to activate alternative causes of participation, thereby increasing the diversity of the respondent pool and reducing the likely nonresponse error. On the other hand, a request that makes survey topic most salient should produce a linkage between nonresponse rates and nonresponse error (e.g., self-administered questionnaires prominently displaying questions on a single topic prior to the sample person's decision about participation).

Summary and Conclusions

We have tested whether topic interest effects might be one common source of nonignorable nonresponse in household surveys. That is, do people cooperate at higher rates to surveys on topics that interest them, and, if so, to what extent does that behavior affect survey estimates? We crafted a test that attempted to identify sets of people who shared an interest in a specific topic and then asked them to participate in a survey on that topic. The first challenge of such a design is choosing a survey topic such that all members of the specific population *do* indeed share that interest. A second challenge is crafting a survey introduction that is realistic (i.e., has high external validity) *and* makes salient the topic of the survey. All of our conclusions are conditioned on these features of our experiment.

Judging from behavior on the first contact containing the survey introduction, we found that persons cooperated at higher rates to surveys on topics of likely interest to them. The odds of cooperating are roughly 40 percent higher for topics of likely interest than for other topics, based on the four frames utilized in the experiment. Given these results and the deductions from leverage-salience theory, we suspect we could make the 40 percent much higher by

making the topic a much more salient aspect of the survey introduction. It is important to note that the overall effects on total response rates of these effects are dampened by noncontact nonresponse, as well as by physical-, mental-, and language-caused nonresponse.

The second hypothesis deduced from leverage-salience theory is that by making a monetary incentive a salient feature of the request, those persons attaching high leverage to personal benefits would cooperate at higher rates (in contrast to their behavior without the financial incentive). This implies that surveys with monetary incentives should show lower tendencies for the "interested" to respond at higher rates than others. The hypothesized direction of effect was found, but it did not attain statistical significance.

The final step in our tests of leverage-salience theory addressed the question, "What is the practical import on survey estimates of the effects of topic interest?" The analytic findings in the appendix show that such effects will be a function of (a) the relative size of the subpopulation interested in a given topic (i.e., prevalence of high leverage), and (b) the degree to which the survey request highlights the topic (i.e., saliency of the topic), producing relatively higher response rates for those interested. In our case, these magnitudes were not large enough to generate much bias due to the main effect of topic interest on cooperation.

Appendix

In this appendix we calculate the maximum nonresponse bias that can arise under various conditions given the experiment described in the article. Nonresponse bias only occurs for variables related to nonresponse propensities. Under our theoretical perspective the largest nonresponse bias is apt to occur for those variables that directly cause nonresponse, such as interest in the survey topic. Here we focus on how the estimate of the proportion of persons interested in topic Y differs between two surveys: Survey 1, identified as a study about topic Y and Survey 2, identified as a study about another topic, unrelated to Y.

Under leverage-salience theory, those interested in the topic ($Y_i = 1$, for the i-th person) will be more likely to cooperate with Survey 1 relative both to their own cooperation with Survey 2 and to the cooperation with Survey 1 of those not interested in the topic ($Y_i = 0$). We consider bias in estimating $\bar{Y} = \sum_{i=1}^{N} Y_i / N$ in the target population of size N. The estimate of $\bar{Y}$ from the first survey is the mean among the r_1 respondents of Survey 1, or $\bar{y}_{r1} = \sum_{i=1}^{r_1} y_i / r_1$. Similarly, the estimate from the second survey is the mean

among the r_2 respondents of Survey 2, or $\bar{y}_{r2} = \sum_{i=1}^{r_2} y_i / r_2$. Of the N persons in the population, $N\bar{Y}$ are persons with an interest in the topic.

A traditional expression for the nonresponse bias of a respondent mean is

$$Bias\ (\bar{y}_r) = (1 - \bar{p})E(\bar{y}_r - \bar{y}_m)$$

where $\bar{p} = \sum_{i=1}^{N} p_i / N$ is the mean probability of responding, and $\bar{y}_r$ and $\bar{y}_m$ are the means of the survey variable for the respondents and nonrespondents, respectively. For our purposes it is desirable to express the bias as a function of the covariance between the response probability of a case and its value on Y. By using the presentation of Lessler and Kalsbeek (1987, section 3.2), the bias of the respondent mean can be expressed as

$$Bias(\bar{Y}_r) \cong \sigma_{pY} + (1 - \bar{p})(E(\bar{y}_r) - \bar{Y})$$

where σ_{pY} is the covariance between the response probability and Y, and $\bar{Y}$ is the population mean of the survey variable. We compute the covariance on the entire population of size N, for which the i-th person manifests p_{1i} likelihood of responding to Survey 1 and p_{2i} likelihood of responding to Survey 2. We first express the difference of nonresponse biases of the two respondent means as

$$Bias(\bar{y}_{r1}) - Bias(\bar{y}_{r2}) \cong [\sigma_{p_1 Y} - \sigma_{p_2 Y}] + \{ (1 - \bar{p}_1)[E(\bar{y}_{r1}) - \bar{Y}] - (1 - \bar{p}_2)[E(\bar{y}_{r2}) - \bar{Y}] \} .$$

The expression is an approximation, because it ignores a term for the difference of technical estimation biases of the two respondent means (which are ratio means). This term can be ignored because (a) the difference in the technical biases is likely to be close to zero, and, further, (b) the technical bias of each mean becomes negligible relative to its standard error as the sample size increases. The expression can be rewritten as the sum of two terms, A and B. The A term, the difference in the covariance terms for the two surveys, is equivalent to

$$A = (\sigma_{p_1 Y} - \sigma_{p_2 Y}) = \frac{1}{N} \sum_{i=1}^{N} (\bar{p}_{1i} - \bar{p}_1)(Y_i - \bar{Y}) - \frac{1}{N} \sum_{i=1}^{N} (\bar{p}_{2i} - \bar{p}_2)(Y_i - \bar{Y})$$

$$= \frac{1}{N} \sum_{i=1}^{N} [(\bar{p}_{1i} - \bar{p}_{2i}) - (\bar{p}_1 - \bar{p}_2)][Y_i - \bar{Y}]$$

$$= \frac{1}{N}\left[\sum_{i=1}^{N}(\bar{p}_{1i}-\bar{p}_{2i})Y_i - N\bar{Y}(\bar{p}_1-\bar{p}_2)\right].$$

For the first term, the summation contains $N(1-\bar{Y})$ persons for whom $Y_i=0$, and the product $(p_{1i}-p_{2i})Y_i$ always equals zero, and $N\bar{Y}$ persons with $Y_i=1$, for whom the product is always equal to the difference in the response rates between Survey 1 and Survey 2. This is equivalent to

$$\begin{aligned} A &= \frac{1}{N}\left[\sum_{i=1}^{N\bar{Y}}(\bar{p}_{1i|Y_i=1}-\bar{p}_{2i|Y_i=1}) - N\bar{Y}(\bar{p}_1-\bar{p}_2)\right] \\ &= \frac{1}{N}(N\bar{Y}\bar{p}_{1|Y_i=1} - N\bar{Y}\bar{p}_{2|Y_i=1}) - \bar{Y}(\bar{p}_1-\bar{p}_2) \\ &= \bar{Y}(\bar{p}_{1|Y_i=1}-\bar{p}_{2|Y_i=1}) - \bar{Y}(\bar{p}_1-\bar{p}_2) \end{aligned}$$

where $p_{1i|Y_i=1}$ is the value of p_{1i} for those cases where $Y_i=1$, and similarly for $p_{2i|Y_i=1}$.

The B term is equivalent to

$$\begin{aligned} B &= \{(1-\bar{p}_1)[E(\bar{y}_{r1})-\bar{Y}] - (1-\bar{p}_2)[E(\bar{y}_{r2})-\bar{Y}]\} \\ &= (1-\bar{p}_1)E(\bar{y}_{r1}) - (1-\bar{p}_2)E(\bar{y}_{r2}) + (\bar{p}_1-\bar{p}_2)\bar{Y}. \end{aligned}$$

We observe that $E(\bar{y}_{r1}) = \dfrac{\bar{Y}\bar{p}_{1|Y_i=1}\cdot 1(1-\bar{Y})\bar{p}_{1|Y_i=1}\cdot 0}{\bar{Y}\bar{p}_{1|Y_i=1}+(1-\bar{Y})\bar{p}_{1|Y_i=0}} = \dfrac{\bar{Y}\bar{p}_{1|Y_i=1}}{\bar{p}_1}$, ignoring the technical estimation bias of the ratio mean, and similarly for $E(\bar{y}_{r2})$. After applying algebraic transformation,

$$\begin{aligned} A+B &= \bar{Y}(\bar{p}_{1|Y_i=1}-\bar{p}_{2|Y_i=1}) - \bar{Y}(\bar{p}_1-\bar{p}_2) + \\ &\quad (1-\bar{p}_1)E(\bar{y}_{r1}) - (1-\bar{p}_2)E(\bar{y}_{r2}) + (\bar{p}_1-\bar{p}_2)\bar{Y} \\ &= \bar{Y}(\bar{p}_{1|Y_i=1}-\bar{p}_{2|Y_i=1}) + (1-\bar{p}_1)E(\bar{y}_{r1}) - (1-\bar{p}_2)E(\bar{y}_{r2}) \\ &= \bar{Y}(\bar{p}_{1|Y_i=1}-\bar{p}_{2|Y_i=1}) - (1-\bar{p}_1)\left[\frac{\bar{Y}\bar{p}_{1|Y_i=1}}{\bar{p}_1}\right] - (1-\bar{p}_2)\left[\frac{\bar{Y}\bar{p}_{2|Y_i=1}}{\bar{p}_2}\right] \\ &= \bar{Y}\left[\frac{\bar{p}_{1|Y_i=1}}{\bar{p}_1} - \frac{\bar{p}_{2|Y_i=1}}{\bar{p}_2}\right]. \end{aligned}$$

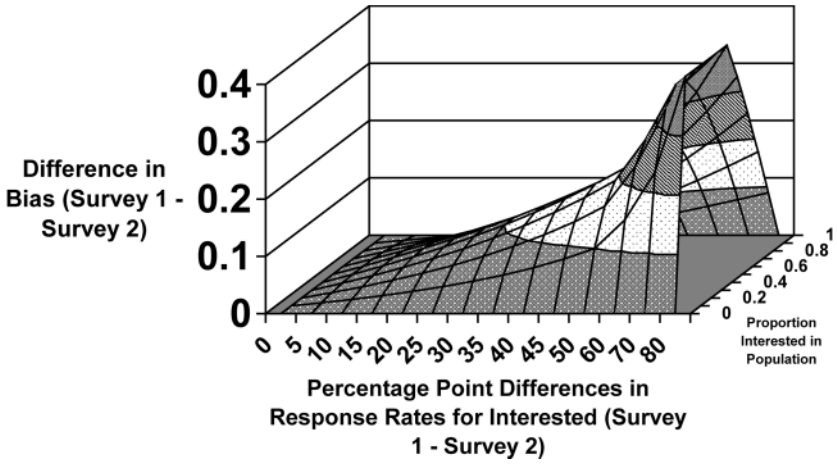

Figure A1. Nonresponse bias difference between Surveys 1 and 2 in estimated proportions interested in topic Y, by response rate differences between Surveys 1 and 2 for those interested in Y, by true proportion interested in Y in the target population (90% response rate for those interested in Y for Survey 1; equal and lower response rates for others for Survey 1 and for all persons for Survey 2).

Thus the magnitude of nonresponse error differences observed in the types of experimental designs we describe in this paper will be a function of

1. the prevalence in the population of interest in the topic, and
2. the ratio of the response rate difference between Survey 1 and Survey 2 among the interested group to the response rate difference among the not interested.

We expect that ratio to be higher in Survey 1 than Survey 2, because Survey 1 increases topic leverage for those interested in Y. Hence, we expect that $\bar{p}_{1|Y_i=1} > \bar{p}_{2|Y_i=1}$ and $\bar{y}_{r1} > \bar{y}_{r2}$. Further, we note that $\bar{p}_1$ and $\bar{p}_2$ are functions of $\bar{Y}$, so that the expression more generally is

$$Bias(\bar{y}_{r1}) - Bias(\bar{y}_{r2}) \cong \bar{Y}\left[\frac{\bar{p}_{1|Y_i=1}}{\bar{p}_1} - \frac{\bar{p}_{2|Y_i=1}}{\bar{p}_2}\right]$$

$$= \bar{Y}\left[\frac{\bar{p}_{1|Y_i=1}}{\bar{Y}\bar{p}_{1|Y_i=1} + (1-\bar{Y})\bar{p}_{1|Y_i=0}} - \frac{\bar{p}_{2|Y_i=1}}{\bar{Y}\bar{p}_{2|Y_i=1} + (1-\bar{Y})\bar{p}_{2|Y_i=0}}\right].$$

To get a sense of the magnitude of possible nonresponse bias differences, we let $\bar{Y}$, the proportion interested in Y in the target population, vary from 0.0 to 1.0; we set $\bar{p}_{1|Y_i=1}$, the response rate for Survey 1 among those interested

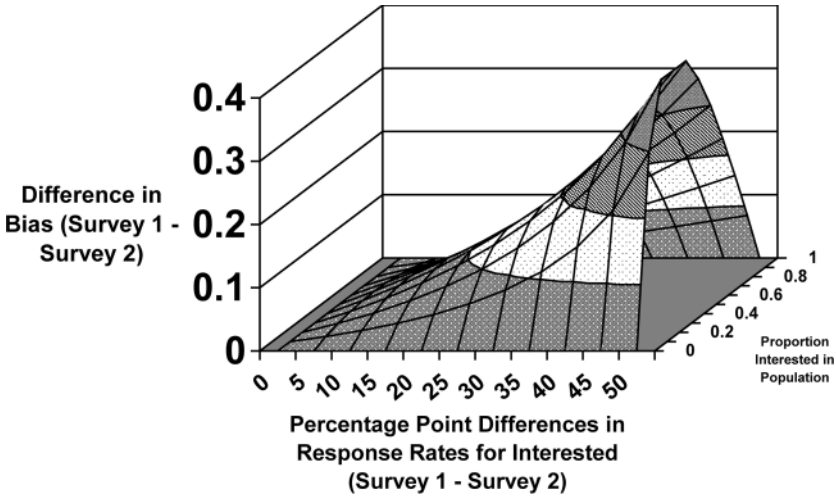

Figure A2. Nonresponse bias difference between Surveys 1 and 2 in estimated proportions interested in topic Y, by response rate differences between Surveys 1 and 2 for those interested in Y, by true proportion interested in Y in the target population (60% response rate for those interested in Y for Survey 1; equal and lower response rates for others for Survey 1 and for all persons for Survey 2).

in Y, at two different levels (90 percent in figure A1 and 60 percent in figure A2); we set the response rate for those uninterested in Y on Survey 1 equal to (a) their response rate on Survey 2 ($\bar{p}_{1|Y_i=0}=\bar{p}_{2|Y_i=0}$), and to (b) the response rate in Survey 2 for those interested in Y ($\bar{p}_{1|Y_i=0}=\bar{p}_{2|Y_i=0}=\bar{p}_{2|Y_i=1}$).

Figures A1 and A2 plot, on the vertical axis, the magnitude of the bias difference for the two estimates of the proportion of the target population interested in Y (i.e., $Bias(\bar{y}_{r1})-Bias(\bar{y}_{r2})$) as a function of the difference in response rates, assuming the average response propensities are all the same for people who are not interested in the topic of the survey (i.e., $(\bar{p}_{1|Y_i=1}-\bar{p}_{2|Y_i=1})=(\bar{p}_{1|Y_i=1}-\bar{p}_{1|Y_i=0})=(\bar{p}_{1|Y_i=1}-\bar{p}_{2|Y_i=0})$). The response rate difference between Survey 1 and Survey 2 is shown on the left-right axis; the proportion of the target population that is interested in Y is shown on the near-far axis.

When the proportion in the population interested in Y is either 0 or 100 percent, the difference in the bias between the surveys is zero. With increasing distance from 0 and 100 percent, the difference in bias increases, with the maximum bias difference when 50 percent of the population is interested in Y. Thus, the surfaces have curvilinear features.

The graphs show that the maximum nonresponse bias differences due to treatments like those described in this article (where the response rate differences between Survey 1 and 2 for those interested in the topic tended to be less than 15 percentage points) will be in the 2–7 percentage point range.

When those interested in Y respond at a very high rate in Survey 1 (e.g., 90 percent as in figure A1), nonresponse bias differences require larger nonresponse rate differences relative to cases where they respond at lower rates in Survey 1 (compare the slope of figure A1 to that of figure A2 on the left-right axis).

In the survey methodological literature there are many experiments that find response rate gains of 5–10 percentage points from some design feature (e.g., incentives, advance letters). Figures A1 and A2 imply that experiments with such outcomes will have relatively little impact on nonresponse error of estimated proportions. Following the analytic results above, this is especially true when (a) those sensitive to the manipulation are either a small minority of the population or the vast majority of the population, and (b) the survey variables of interest are less than perfectly correlated with sensitivity to the design feature found to increase participation.

References

American Association for Public Opinion Research. 2000. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys.* Ann Arbor, MI: AAPOR.

Baumgartner, Robert, and Pamela Rathbun. 1997. "Prepaid Monetary Incentives and Mail Survey Response Rates." Paper presented at the annual conference of the American Association for Public Opinion Research, Norfolk, VA.

Curtin, Richard, Stanley Presser, and Eleanor Singer. 2000. "The Effects of Response Rate Changes on the Index of Consumer Sentiment." *Public Opinion Quarterly* 64:413–28.

Goyder, John. 1987. *The Silent Minority: Nonrespondents on Sample Surveys*. Boulder, CO: Westview Press.

Groves, Robert, and Mick Couper. 1998. *Nonresponse in Household Interview Surveys.* New York: John Wiley and Sons.

Groves, Robert, Eleanor Singer, and Amy Corning. 2000. "Leverage-Salience Theory of Survey Participation: Description and an Illustration." *Public Opinion Quarterly* 64:299–308.

Keeter, Scott, Carolyn Miller, Andrew Kohut, Robert Groves, and Stanley Presser. 2000. "Consequences of Reducing Nonresponse in a Large National Telephone Survey." *Public Opinion Quarterly* 64:125–48.

Kulka, Richard A. 1994. "The Use of Incentives to Survey 'Hard-to-Reach' Respondents: A Brief Review of Empirical Research and Current Practice." Paper prepared for Seminar on New Directions in Statistical Methodology, Bethesda, MD.

Lessler, Judith, and William Kalsbeek. 1987. *Nonsampling Errors in Surveys*. New York: Wiley.

Merkle, Daniel, and Murray Edelman. 2002. "Nonresponse in Exit Polls: A Comprehensive Analysis." In *Survey Nonresponse*, ed. R. M. Groves, D. A. Dillman, J. L. Eltinge, and R. J. A. Little, pp. 243–58. New York: Wiley.

Morton-Williams, Jean. 1993. *Interviewer Approaches*. Brookfield, VT: Dartmouth.

National Academy of Sciences. 1979. *Panel on Privacy and Confidentiality as Factors in Survey Response*. Washington, DC: National Academy Press.

Oksenberg, Lois, Lerita Coleman, and Charles F. Cannell. 1986. "Interviewers' Voices and Refusal Rates in Telephone Surveys." *Public Opinion Quarterly* 50:97–111.

Petty, Richard E., and John T. Cacioppo. 1986. *Communication and Persuasion: Central and Peripheral Routes to Attitude Change*. New York: Springer-Verlag.

Rubin, Donald R. 1987. *Multiple Imputation for Nonresponse in Surveys*. New York: Wiley.

Singer, Eleanor. 2002. "The Use of Incentives to Reduce Nonresponse in Household Surveys." In *Survey Nonresponse*, ed. R. M. Groves, D. A. Dillman, J. L. Eltinge, and R. J. A. Little, pp. 163–78. New York: Wiley.