# THE EFFECTS OF RESPONSE RATE CHANGES ON THE INDEX OF CONSUMER SENTIMENT

RICHARD CURTIN
STANLEY PRESSER
ELEANOR SINGER

**Abstract** From 1979 to 1996, the Survey of Consumer Attitudes response rate remained roughly 70 percent. But number of calls to complete an interview and proportion of interviews requiring refusal conversion doubled. Using call-record histories, we explore what the consequences of lower response rates would have been if these additional efforts had not been undertaken. Both number of calls and initially cooperating (vs. initially refusing) are related to the Index of Consumer Sentiment (ICS), but only number of calls survives a control for demographic characteristics. We assess the impact of excluding respondents who required refusal conversion (which reduces the response rate 5–10 percentage points), respondents who required more than five calls to complete the interview (reducing the response rate about 25 percentage points), and those who required more than two calls (a reduction of about 50 percentage points). We found no effect of excluding any of these respondent groups on cross-sectional estimates of the ICS using monthly samples of hundreds of cases. For yearly estimates, based on thousands of cases, the exclusion of respondents who required more calls (though not of initial refusers) had an effect, but a very small one. One of the exclusions generally affected estimates of change over time in the ICS, irrespective of sample size.

A basic tenet of survey research is that high response rates are better than low ones. Indeed, a low response rate is one of the few outcomes or features that—taken by itself—is considered a major threat to the usefulness of a survey. For example, the "Best Practices" guide of the American Association for Public Opinion Research (AAPOR1997, p. 5) states that "a low cooperation

RICHARD CURTIN is director of the Survey of Consumer Attitudes at the University of Michigan Survey Research Center. STANLEY PRESSER is professor of sociology at the University of Maryland. ELEANOR SINGER is senior research scientist at the University of Michigan Survey Research Center.

Public Opinion Quarterly Volume 64:413–428. © 2000 by the American Association for Public Opinion Research
All rights reserved. 0033-362X/2000/6404-0001$02.50

or response rate does more damage in rendering a survey's results questionable than a small sample." Theoretical models of survey participation (Groves and Couper 1998) suggest that nonrespondents differ from respondents in systematic ways.

As a result, survey organizations devote extensive resources to reducing nonresponse: making multiple calls, attempting to convert initial refusals, and providing cash incentives to respondents or interviewers. Moreover, the fraction of survey budgets allocated for this purpose has grown over time as it has become more difficult to contact and gain cooperation from individuals. Thus, in Wonderland fashion, survey research must run harder to stay in place.

The emphasis on obtaining high response rates (or maintaining past rates) stems from the belief that increases in nonresponse lead to greater bias. Yet bias is not a simple function of nonresponse level. It is a multiplicative function of the nonresponse level and the nonrespondents' distinctiveness. Consequently, as nonresponse increases, bias increases only if the distinctiveness of the nonrespondents stays constant or becomes more pronounced.

Differences between respondents and nonrespondents are often not constant over changes in response rate. For instance, increases in response rate due to monetary incentives have been shown to come disproportionately from individuals who are otherwise less likely to cooperate, for example, those for whom the survey topic is less salient (Baumgartner and Rathbun 1997) and those with lower civic involvement (Groves, Singer, and Corning 2000). In these cases, lower response rates are accompanied by more distinctive nonrespondents, and thus increasing the response rate reduces nonresponse bias.

Studies have also shown that respondents interviewed on early calls differ from those interviewed on later calls. Distributions of respondent background characteristics such as household size, age, sex, education, and race alter with changes in response rate (e.g., Dunkelberg and Day 1973; Fitzgerald and Fuller 1982; Sharp and Feldt 1959). The relation of number of calls to sample composition in these studies, however, is mainly due to the distinctiveness of those interviewed on the first or second call; third and later calls usually show much less variation in respondent characteristics. This suggests that the effect of reducing nonresponse might depend on the level from which one starts: increasing the response rate from 20 percent to 40 percent (by moving, say, from a one-call design to a five-call design) may have a much greater impact on results than a parallel increase from 40 percent to 60 percent (moving, e.g., from a five-call design to a 15-call design).

A comparison like the latter one has recently been reported by Keeter et al. (2000). They compared estimates from the same omnibus questionnaire administered in two different designs, one conducted over 5 days yielding a 37 percent response rate and the other conducted over 2 months yielding a 61 percent response rate.[1] Remarkably, there were very few statistically sig-

1. The denominators of these response rates include all sampled phone numbers with the ex-

nificant differences in the results across a large set of demographic, behavioral, attitudinal, and knowledge items. But the extent to which one can generalize from Keeter et al.'s single omnibus survey is unclear.

The research reported below brings a large amount of additional data to bear on the issue by examining the impact of response rate on the Index of Consumer Sentiment (ICS). The ICS is derived from the University of Michigan's Survey of Consumer Attitudes (SCA), a major national study that is conducted every month to monitor changes in consumer sentiments about the economy. The call-record histories for the more than two hundred SCA surveys conducted in the last 20 years allow us to see how differing levels of effort, which yield different response rates, affect the results.

## METHOD

Although the SCA began in the late 1940s as a face-to-face study, for the past 2 decades it has been a random digit dial telephone survey of households in the coterminous United States. Other than the constraint imposed by the monthlong data collection period, there is no limit on the number of times unanswered numbers are called, and attempts to convert essentially all initial refusals are made by specially trained interviewers. Within households, a single adult, age 18 or older, is selected using the Kish procedure.

Each monthly survey consists of roughly 60 percent new cases and 40 percent reinterviews from an earlier month. In this article we use only the new cases. Most recently, there have been about three hundred new interviews a month, though the number was somewhat larger in earlier years. Our data come from the 211 surveys conducted between June 1979 and December 1996 consisting of more than seventy-two thousand independent interviews.

The response rate over these 17.5 years was fairly constant, averaging roughly 70 percent—slightly more (71 percent) in the first half of the period and slightly less (68 percent) in the second half (fig. 1).[2] But the effort to obtain that result has increased dramatically over time. As shown in figures 2 and 3, the mean number of calls to complete an interview more than doubled from 3.9 in 1979 to 7.9 in 1996, and interviews from refusal conversion likewise rose from 7.4 percent to 14.6 percent.[3] These are separate trends, as the mean number of calls to complete interviews that did not require refusal conversion also increased (from 3.7 to 7.4). Taken together, the additional

---

ception of both known ineligibles (e.g., businesses) and the fraction estimated to be ineligible of those whose eligibility was unknown. If all the unknown cases are included in the denominators, both response rates are 4 percentage points lower.

2. The denominators of these response rates include all sampled phone numbers with the exception of those known to be ineligible, and the numerators include a very small number of partial interviews. This corresponds to AAPOR's "Response Rate 2."

3. The curvilinearity in the figures is puzzling; we know of no change in survey procedures that would explain it.

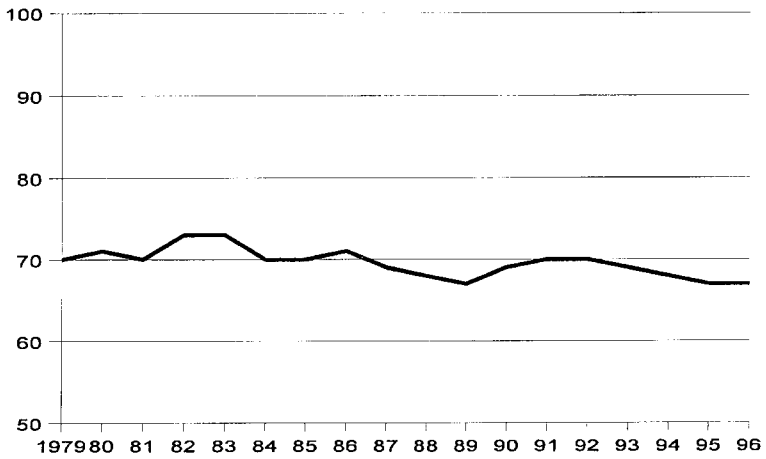


**Figure 1.** Survey of Consumer Attitudes response rates by year

calls and refusal conversions translated into a substantial increase in interviewers' time spent per completed interview (2.1 hours in 1981 vs. 2.7 in 1996).[4]

Would the survey results have been affected if this additional effort had not been made and lower response rates achieved? The key SCA result involves the Index of Consumer Sentiment (ICS), which is based on the answers to five questions:

> We are interested in how people are getting along financially these days. Would you say that you (and your family living there) are better off or worse off financially than you were a year ago?
>
> Now looking ahead—do you think that a year from now you (and your family living there) will be better off financially, or worse off, or just about the same as now?
>
> Now turning to business conditions in the country as a whole—do you think that during the next 12 months we'll have good times financially, or bad times, or what?
>
> Looking ahead, which would you say is more likely—that in the country as a whole we'll have continuous good times during the next 5 years or so, or that we will have periods of widespread unemployment, or depression, or what?
>
> About the big things people buy for their homes—such as furniture, a refrigerator, stove, television, and things like that. Generally speaking, do you think now is a good or a bad time for people to buy major household items?

4. We compare 1996 with 1981 because the average interview length in those 2 years was almost identical: 34 minutes in 1981 and 33 minutes in 1996. (Although the core consumer attitude questions are constant, the surveys contain varying numbers of additional questions.)


**Figure 2.** Percentage of interviews requiring refusal conversion by year

The appendix shows how the answers are combined to produce the ICS.

Our analysis proceeds as follows: in the section titled "Comparing Respondents Who Required Different Levels of Effort," we compare initial cooperators to refusal conversions and respondents interviewed on early calls to those interviewed on later calls. Then, in the section titled "Effects of Excluding Respondents Who Require More Effort," we assess the effect of

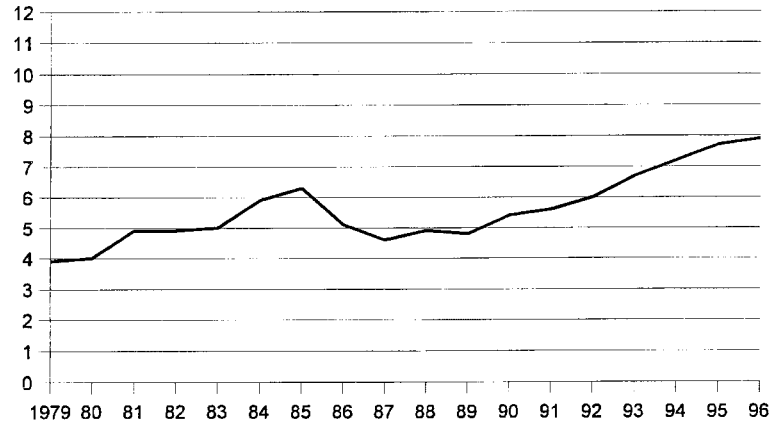


**Figure 3.** Mean number of calls to complete an interview by year

**Table 1.** Relationship of ICS to Number of Calls and Initial Cooperation, by Year (Ordinary Least Squares Estimates)

| | 1979–87 | | 1988–96 | |
|---|---|---|---|---|
| | $\beta$ | $p$ | $\beta$ | $p$ |
| Refusal (yes) | −4.26 | .0001 | −3.90 | .0001 |
| No. of calls | .35 | .0001 | .20 | .0001 |
| Survey year | −12.39 | .0001 | −97.00 | .0001 |
| Year squared | 4.61 | .0001 | 6.11 | .0001 |
| Year cubed | −.34 | .0001 | −.12 | .0001 |
| $R^2$ | .11 | | .03 | |
| $N$ | 40,464 | | 31,518 | |

reduced levels of survey effort in two ways. In one set of analyses, we estimate the impact on the ICS of excluding respondents who required refusal conversion. This reduces the response rate 5 to 10 percentage points depending on the year. In the other set of analyses, we gauge the effect on the ICS of (*a*) excluding respondents who required more than five calls to complete the interview and (*b*) excluding those who required more than two calls. This reduces the response rate on the order of 25 and 50 percentage points, respectively (less in the late 1970s to early 1980s and more in the 1990s). Where they differ, results are presented separately for the first and second halves of the period (1979–87 and 1988–96) in order to capture differences due to easier and more difficult survey environments.

## RESULTS

### COMPARING RESPONDENTS WHO REQUIRED DIFFERENT LEVELS OF EFFORT

Table 1 presents the relationship of the ICS to number of calls and to initially cooperating versus refusing, controlling on time of the survey (since the ICS as well as refusals and number of calls change over time). Respondents who did not require refusal conversion had higher ICS scores (were more optimistic about the economy) than those who initially refused. By contrast, respondents interviewed on early calls had lower ICS scores (were less optimistic) than those interviewed on later calls, with the effect more pronounced in the earlier period.[5]

Why should contact and cooperation be related to the level of consumer

5. There are, however, no differences between initial cooperators and refusal conversions or between interviews from early calls and those from later ones in the ICS variances.

**Table 2.** Effects of Demographic Characteristics on Number of Calls, Controlling for Whether or Not Refusal Conversion Was Required, 1979–96 (Ordinary Least Squares Estimates)

| | $\beta$ | $p$ |
|---|---|---|
| Refusal (yes) | 3.15 | .0001 |
| Log income (1996 $) | .18 | .0001 |
| Age in years | −.04 | .0001 |
| Years of schooling | .02 | .0133 |
| Sex (female) | −.40 | .0001 |
| Race (nonwhite) | 1.37 | .0001 |
| Survey year | .77 | .0001 |
| Year squared | −.09 | .0001 |
| Year cubed | .004 | .0001 |
| $R^2$ | .08 | |
| $N$ | 65,371 | |

sentiment? We believe it is at least partly because refusal conversions are disproportionately respondents of lower socioeconomic status (who are typically less optimistic about the economy), and those who are more difficult to reach are disproportionately of higher status (who are typically more optimistic). Indeed, tables 2 and 3 show that respondents requiring a larger

**Table 3.** Effect of Demographic Characteristics on Requiring Refusal Conversion, Controlling for Number of Calls, 1979–96 (Logistic Regression Estimates)

| | $\beta$ | $p$ |
|---|---|---|
| No. of calls | .07 | .0001 |
| Log income (1996 $) | −.001 | .9755 |
| Age in years | .017 | .0001 |
| Years of schooling | −.04 | .0001 |
| Sex (female) | .12 | .0001 |
| Race (nonwhite) | −.37 | .0001 |
| Survey year | .25 | .0001 |
| Year squared | −.02 | .0001 |
| Year cubed | .001 | .0001 |
| Pseudo $R^2$ | .03 | |
| $N$ | 65,372 | |

**Table 4.** Relationship of Index of Consumer Sentiment to Number of Calls and Refusal Conversion, Controlling for Demographics and Year, 1979–87 and 1988–96

| | 1979–87 | | 1988–96 | |
|---|---|---|---|---|
| | $\beta$ | $p$ | $\beta$ | $p$ |
| Refusal (yes) | −1.12 | .0719 | −1.09 | .0976 |
| No. of calls | .20 | .0001 | .06 | .0428 |
| Log income (1996 $) | 5.83 | .0001 | 3.79 | .0001 |
| Age in years | −.27 | .0001 | −.30 | .0001 |
| Years of schooling | 1.14 | .0001 | .94 | .0001 |
| Sex (female) | −8.60 | .0001 | −7.38 | .0001 |
| Race (nonwhite) | −6.42 | .0001 | −2.41 | .0001 |
| Survey year | −12.31 | .0001 | −100.73 | .0001 |
| Year squared | 4.66 | .0001 | 6.33 | .0001 |
| Year cubed | −.34 | .0001 | −.13 | .0001 |
| $R^2$ | .19 | | .09 | |
| $N$ | 40,464 | | 28,532 | |

number of calls are younger and more affluent, while those who required refusal conversion are older and less educated.[6]

Table 4 reveals that once these demographic variables are controlled, the association between the ICS and refusal conversion is no longer significant at the .05 level in either time period. Although number of calls remains significantly associated with the ICS after controlling for demographics, the size of the coefficient is considerably reduced.

To this point, the analyses have been based on sample sizes that greatly exceed those actually used to estimate monthly consumer confidence. In addition, the analyses have involved point estimates, whereas a key focus of the consumer confidence survey is change over time. Thus the next analyses examine estimates of change as well as of absolute level and do so using samples averaged over months, quarters (combining three consecutive surveys), and years (combining 12 surveys).[7]

Estimates of the ICS for these different time periods are shown in table 5. The absolute means are very similar across the periods, though the mean change increases as the time period lengthens. Because of sample size dif-

6. Because the number of calls to complete an interview is related to whether the interview required refusal conversion (over the entire period, initial cooperators averaged 5.2 calls vs. 8.2 calls for those who initially refused), we control for number of calls in our analysis of refusal conversion and for refusal conversion in our analysis of number of calls.

7. As noted earlier, published data on the ICS are now based on monthly samples of 500: 300 new cases (those utilized in our analysis) and 200 reinterviews of cases that were first interviewed 6 months earlier. As a result, the monthly estimates we report are based on smaller samples than those actually used.

**Table 5.** Means and Mean Change in the ICS Averaged over Months, Quarters, and Years, 1979–96

| | Level | | | | Change | | | |
|---|---|---|---|---|---|---|---|---|
| | *N* | Mean | SE | SD | *N* | Mean | SE | Mean (Absolute Values) |
| Months | 211 | 84.2 | 1.93 | 35.2 | 210 | .14 | 2.74 | 3.10 |
| Quarters | 70 | 84.2 | 1.12 | 35.3 | 69 | .47 | 1.59 | 3.73 |
| Years | 17 | 84.8 | .56 | 35.5 | 16 | 1.79 | .80 | 4.84 |
| Total sample | 72,424 | 83.2 | .14 | 37.1 | … | … | … | … |

NOTE.—The means for months, quarters, and years represent the averaged monthly, quarterly, or yearly levels of the Index of Consumer Sentiment (ICS). The total sample mean differs slightly because the monthly sample sizes in the early part of the period were somewhat larger, and the ICS was lower (reflecting the recession in the early 1980s). The standard errors (SE) of the estimated means show the impact of the different sample sizes when the data are pooled by different time periods. The "mean (absolute values)" is the mean of the differences without regard to sign.

ferences, the longer the time period, the smaller the standard errors for both the absolute means and the mean change.

How much do the means differ for cases obtained with varying levels of effort? The upper panel of table 6 shows the average differences in absolute level of the ICS between initial cooperators and refusal conversions and between early calls and later ones, and the lower panel shows the corresponding differences in the estimates of change in the ICS. A comparison of these results to the standard errors displayed in table 5 suggests a small impact on estimates of level but virtually none on estimates of change.

This is confirmed in table 7, the upper panel of which shows that differences in absolute level of the ICS between initial cooperators and refusal conversions and between early calls and later ones attain statistical significance in about one in 10 of the 211 monthly surveys. Consistent with expectations based on variations in sample size, a greater proportion of the differences attain significance for the quarterly estimates, and an even larger fraction reach significance for the annual estimates.

By contrast, for estimates of change in the ICS, table 7's lower panel reveals that, with one exception, the proportion of statistically significant differences ($p < .05$) between initial cooperators and refusal conversions and between early calls and later ones does not exceed, or is very close to, the 5 percent to be expected by chance. The notable exception is the difference between yearly samples based on more and less than six calls, where the estimates differ almost one-third of the time.

Overall, then, both number of calls and initial cooperation versus refusal are related to estimates of the level of the ICS, though only number of calls

**Table 6.** Mean Differences in the Index of Consumer Sentiment (ICS) between Estimates Based on Easier versus Harder Cases

| | Initial Cooperators vs. Refusal Conversions | 1–5 Calls vs. 6 or More Calls | 1–2 Calls vs. 3 or More Calls | Initial Cooperators | |
|---|---|---|---|---|---|
| | | | | 1–5 Calls vs. 6 or More Calls | 1–2 Calls vs. 3 or More Calls |
| Mean difference in ICS level (*N*): | | | | | |
| Months (211) | 3.45 | −2.33 | −2.88 | −2.88 | −3.47 |
| Quarters (70) | 3.42 | −2.32 | −2.87 | −2.83 | −3.46 |
| Years (17) | 3.35 | −2.35 | −2.86 | −2.81 | −3.41 |
| Mean difference in ICS change (*N*): | | | | | |
| Months (210) | −.04 | −.02 | −.01 | −.02 | −.002 |
| Quarters (69) | .03 | −.01 | −.01 | .01 | .003 |
| Years (16) | −.09 | −.19 | −.07 | −.20 | −.07 |

**Table 7.** Percentage of Significant Differences ($p < .05$) between Estimates of ICS Based on Easier and Harder Cases

| | Initial Cooperators vs. Refusal Conversions | 1–5 Calls vs. 6 or More Calls | 1–2 Calls vs. 3 or More Calls | Initial Cooperators | |
|---|---|---|---|---|---|
| | | | | 1–5 Calls vs. 6 or More Calls | 1–2 Calls vs. 3 or More Calls |
| Level of ICS (*N*): | | | | | |
| Monthly (211) | 10 | 11 | 11 | 13 | 11 |
| Quarterly (70) | 13 | 23 | 23 | 23 | 31 |
| Yearly (17) | 47 | 35 | 77 | 35 | 82 |
| Change in ICS (*N*): | | | | | |
| Monthly (210) | 4 | 6 | 4 | 6 | 4 |
| Quarterly (69) | 0 | 7 | 6 | 6 | 6 |
| Yearly (16) | 0 | 31 | 6 | 25 | 13 |

NOTE.—Table entries represent the proportion of differences that were significantly different at the 5% level. For example, 10% of the 211 monthly estimates of the Index of Consumer Sentiment (ICS) differed significantly between initial cooperators and refusal conversion cases, and 4% of the 210 estimates of month-to-month change (the change among initial cooperators compared with the change among refusal conversions) differed significantly.

survives a control for background variables at the .05 level. Only number of calls shows any relation to estimates of change in the ICS (and only with the largest—yearly—samples); refusal conversion is unrelated to change in the ICS.

### EFFECTS OF EXCLUDING RESPONDENTS WHO REQUIRE MORE EFFORT

So far we have investigated the differences between respondents who are easier and harder to interview (a part-to-part comparison). We now examine whether changes in response rate due to refusal conversions and additional calls would have an impact on the survey estimates (a part-to-whole comparison).

The ideal evidence about whether conducting refusal conversions or making a large number of calls affects the survey's results would come from an experimental comparison of the SCA with otherwise identical surveys done at the same time using lower levels of effort—for instance, no refusal conversions or no more than five calls. Although such surveys do not exist, we can simulate them by truncating the existing samples—for example, omitting the refusal conversion cases or the cases that were interviewed after more than two calls or more than five calls. We do not, however, compare the

**Table 8.** Correlations of ICS Estimates Based on All Call Design with Estimates from Restricted Call Designs, 1979–96

| | Quarterly Observations | | Half-Year Observations | |
|---|---|---|---|---|
| | Level | Change | Level | Change |
| Correlation of all calls to initial cooperators | .979 | .772 | .993 | .936 |
| Correlation of all calls to 1–5 calls | .980 | .781 | .995 | .961 |
| Correlation of all calls to 1–2 calls | .964 | .684 | .985 | .888 |
| Correlation of all calls to initial cooperators and 1–5 calls | .975 | .684 | .985 | .888 |
| Correlation of all calls to initial cooperators and 1–2 calls | .966 | .694 | .987 | .903 |
| *N* | 70 | 69 | 35 | 34 |

truncated samples to the total samples from which they came because the lack of independence between samples would artifactually increase the similarity of the estimates. Instead, we use a random subsample of each survey as our "total" estimate and truncate the remaining random subsample. To maintain equal size subsamples, the random allocation is done separately for the different truncations (e.g., refusal conversions, interviews after the fifth call, etc.). For example, a monthly survey with 12.5 percent refusal conversions was randomly divided into one subsample with 46.7 percent of the cases and the other with 53.3 percent. The refusal conversion cases were then omitted from the larger of the subsamples so that the two subsamples were of about equal size. This permits us to examine whether conclusions from independent total and truncated subsamples differ. (Using subsamples reduces sample size, so we restrict our comparisons to quarterly and half-year estimates.)

Table 8 shows, for both level and change, the correlations between ICS estimates based on subsamples of total calls and estimates based on independent subsamples restricted to five different groups: initial cooperators, respondents interviewed after 1–5 calls, respondents interviewed after 1–2 calls, initial cooperators interviewed on 1–5 calls, and initial cooperators interviewed on 1–2 calls. Because the correlations do not differ between the earlier and the later periods, they are presented only for the entire period 1979–96.

Columns 1 and 3 show that estimates of the ICS level from all of the restricted subsamples are almost perfectly correlated with those based on all calls. The correlations for estimates of change in the ICS are somewhat lower than those for the absolute level, especially for those based on quarters, but

**Table 9.** Percentage of Significant Differences ($p < .05$) between Estimates of Index of Consumer Sentiment Based on All Calls versus Restricted Calls

| | All Calls versus | | | | |
|---|---|---|---|---|---|
| | Initial Cooperators | 1–5 Calls | 1–2 Calls | Initial Cooperators and 1–5 Calls | Initial Cooperators and 1–2 Calls |
| Level of ICS: | | | | | |
| Quarterly (70) | 6 | 3 | 7 | 6 | 6 |
| Half-year (35) | 3 | 3 | 9 | 3 | 11 |
| Yearly (17) | 0 | 3 | 18 | 6 | 24 |
| Change in ICS: | | | | | |
| Quarterly (69) | 9 | 3 | 6 | 7 | 4 |
| Half-year (34) | 6 | 0 | 0 | 0 | 0 |
| Yearly (16) | 0 | 0 | 0 | 0 | 0 |

NOTE.—Table entries represent the proportion of comparisons between two independent estimates that were significantly different at the 5% level. For example, comparing the 70 quarterly estimates of the ICS from the random subsamples of all-call cases with those from the random subsamples of initial cooperators, 6% of the level comparisons were significantly different, and 9% of the quarter-to-quarter change estimates were significantly different.

we believe this is because, relative to the estimates, the sampling errors for change are larger than for absolute level. Increasing the sample size by computing half-year estimates bears this out: the correlation of the two estimates is close to 1.0 for the half-year to half-year change.[8]

Another way to assess whether survey estimates are affected by excluding either interviews from later calls or refusal conversions is to see how often the means for the truncated and total subsamples differ. The upper panel of table 9 shows that for refusal conversions and interviews after the fifth call, the number of statistically significant differences between the truncated and total subsamples falls in the range to be expected by chance. By contrast, exclusion of cases interviewed after the second call produces somewhat more differences than would be expected by chance, with the departure from chance levels increasing as the sample size increases. This is true whether or not the samples based on number of calls are restricted to initial cooperators or include refusal conversions.

8. All our analyses are based on raw data, whereas the published ICS is computed from data poststratified to census demographic totals as well as weighted to reflect differential selection probabilities (due to variation in household size and number of residential phone lines). The similarity of the ICS in the truncated and total subsamples might be even greater with the appropriate weights, since weighting would make the demographics of the truncated subsamples more similar to those of the total subsamples (i.e., weighting would correct for the differences shown in tables 2 and 3).

Turning to estimates of change (table 9, lower panel), the number of significant differences due to the exclusion of either refusal conversions or cases interviewed on later calls generally does not exceed sampling error, particularly as the sample size increases. Thus time series estimates seem especially immune to these changes in response rate.

## DISCUSSION

Truncating the SCA samples, thereby creating a lower response rate, is not the same as conducting a survey designed to yield a lower response rate. If the SCA were carried out using a limited number of calls, the calls would probably be scheduled differently. Under the SCA design, for instance, the fifth call has no special significance and thus is not restricted to times when respondents are most likely to be home. By contrast, under a five-call design, the fifth call (and in all likelihood earlier calls as well) might be made only during particularly productive times (e.g., weekday evenings). Nonetheless, we believe our simulations provide a useful approximation of response rate differences from different designs.

Our results both replicate and extend the finding of Keeter et al. (2000) that large differences in response rate had only minor effects on cross-sectional analyses in a single omnibus survey. We have replicated their work by assessing the impact of response rate in 211 additional surveys, and we have extended it by examining the effect of response rate on time series analysis.

We found systematic differences in the level of the ICS between easier and harder to interview respondents. However, detecting these differences requires sample sizes larger than those typically used in the consumer attitude surveys. Consequently, if one focuses on the impact of excluding the more difficult cases, not on the difference between the easier and more difficult cases, the effects are too small to be visible in monthly analyses of consumer confidence. The truncation effects are more evident when cases are aggregated to quarterly or yearly samples. Thus in applications using larger samples, for instance, the measurement of unemployment, the effects we were able to detect by cumulating across surveys could be consequential. Put another way, as sampling error declines with larger sample sizes, ceteris paribus, the impact of nonresponse bias becomes greater.

Time series analyses of the ICS appear even less vulnerable to nonresponse bias than cross-sectional analyses. Even when our harder to interview respondents differed from those easier to interview, the two changed at the same rate over time, that is, the nonresponse bias remained constant, and temporal analyses were thus unaffected. As a result, we believe the ICS's ability to predict future changes in economic conditions is unlikely to be affected by the nonresponse differences we examined.

Unlike Keeter et al. (2000), who analyzed a large number of measures, we

focused on a single indicator, albeit a very widely used one. The indicator is, however, based on five items, so we repeated our analyses for each of the individual items. Although the nonresponse effects are somewhat larger for the two items that refer to the respondents' own financial situation (as opposed to the economy at large), in general each of the items shows patterns similar to those we have reported with the index.

Much survey analysis focuses on associations between variables, whereas Keeter et al. (2000) examined only marginals, and the only relationship we examined involved time. Although research has generally found relationships to be more immune than marginals to methodological effects, further work on nonresponse's impact on associations between variables is needed.

Moreover, it is important to note that both our work and that of Keeter et al. (2000) address the issue of relative, not absolute, nonresponse bias. Observing little effect of nonresponse when comparing response rates of 60 to 70 percent with rates much lower does not mean that the surveys with 60 to 70 percent response rates do not themselves suffer from significant nonresponse bias, and research on this issue, too, is urgently needed.

We began by noting that changes in nonresponse bias depend on the relationship between the level of nonresponse and the distinctiveness of nonrespondents. The generally small differences we found between truncated and total subsamples suggest that in the SCA nonrespondents' distinctiveness decreases as nonresponse grows. Research focusing on this issue with variables that are more likely to be related to the causes of contactability and cooperation (e.g., time spent away from home, health impairments, and lack of interest in the survey topic) should be of high priority, as the results of such studies would not only greatly improve our understanding of the nature of nonresponse but also have important practical implications for deciding how to trade off between survey costs and response rate.[9]

But such work needs to go hand in hand with the development of better theories about the conditions under which nonresponse affects survey estimates. At present, sampling theory is our only justification for drawing inferences from samples to populations, and drawing on this theory assumes that response rates are high enough that achieved samples accurately reflect target samples. Better theories about when nonresponse matters will provide much needed guidance in deciding whether and how to generalize from findings such as ours.

## Appendix

The basic formula for the Index of Consumer Sentiment is

9. The variety of causes of noncooperation is apt to be more restricted in interviewer-administered surveys than in mail surveys, where literacy and interest in the survey topic are likely to be strongly related to nonresponse.

$$\mathrm{ICS}_t = \sum_{j=1}^{5} \left( P^{f}_{jt} - P^{u}_{jt} \right) 100 + 100,$$

where $P^{f}_{jt}$ = the sample proportion giving favorable replies to the $j$th question at time $t$, and $P^{u}_{jt}$ = the sample proportion giving unfavorable replies to the $j$th question at time $t$. Equivalently, the formula can be expressed in terms of the individual responses

$$\mathrm{ICS}_t = \sum_{j=1}^{5} \sum_{i=1}^{n} \frac{X_{ijt}}{n} (100) + 100,$$

where $X^{ijt} = 1$ if favorable response to $j$th question by $i$th respondent at time $t$, $X^{ijt} = -1$ if unfavorable response to $j$th question by $i$th respondent at time $t$, and $X^{ijt} = 0$ for all other responses to $j$th question by $i$th respondent at time $t$. The final figures are published as a proportion of the base year value (1966).

## References

American Association for Public Opinion Research. 1997. *Best Practices for Survey and Public Opinion Research.* Ann Arbor, MI: American Association for Public Opinion Research.

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.

Dunkleberg, William, and George Day. 1973. "Nonresponse Bias and Callbacks in Sample Surveys." *Journal of Marketing Research* 10:160–68.

Fitzgerald, Robert, and Linda Fuller. 1982. "I Hear You Knocking but You Can't Come In: The Effects of Reluctant Respondents and Refusers on Sample Survey Estimates." *Sociological Methods and Research* 11:3–32.

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

Groves, Robert, Eleanor Singer, and Amy Corning. 2000. "Decision Making in Survey Participation." *Public Opinion Quarterly* 64:000–000.

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.

Sharp, Harry, and Allan Feldt. 1959. "Some Factors in a Probability Sample Survey of a Metropolitan Community." *American Sociological Review* 24:650–61.

## ERRATA

Curtin, Richard, Stanley Presser, and Eleanor Singer, The Effects of Response Rate Changes on the Index of Consumer Sentiment, vol. 64, no. 4, pp. 413–428. On page 413, the last sentence of the **Abstract** should read: "None of the exclusions generally affected estimates of change over time in the ICS, irrespective of sample size."

Scott Keeter's review (vol. 64, no. 4, pp. 543–546) of *Reading Mixed Signals: Ambivalence in American Public Opinion about Government* by Albert H. Cantril and Susan Davis Cantril listed the book's publisher as Johns Hopkins University Press. The volume is published by the Woodrow Wilson Center Press (Washington, DC) and distributed by the Johns Hopkins University Press (Baltimore, MD)