

Response Rate Achieved in Government Surveys

Ruey-Pyng Lu

Energy Information Administration

KEY WORDS: Response rate; Nonresponse bias; Information collection request

Purpose

This paper presents the recommendations made to OMB for processing Information Collection Requests (ICRs) from Federal Agencies. We recommend that OMB adopt a formula based on historical experience for calculating expected response rates. When a proposed collection has a predicted response rate of 80% or less, the sponsor must provide detailed information showing why it expects a higher response rate and must identify specific steps it will take to maximize the response rate, and a plan of assessing the nonresponse bias.

Introduction

This study is to gain an understanding of the achieved response rate in Federal Government surveys. A response rate form is used to calculate the actual response rate of surveys approved by the Office of Management and Budget (OMB) in 1998. After analyzing the response rates characterized and tabulated by various factors, a standard for OMB use in reviewing Information Collection Requests (ICRs) is recommended. These analyses also provide recommendations to agencies on predicting response rates and on how to assess nonresponse bias.

Predicting the Response Rate

This section estimates a rule for predicting a proposed survey's response rate based on the response rate of similar surveys conducted in the past. The rule is restricted to general-purpose statistics for individuals and organizations within the United States. The basic idea is that the predicted response rate depends upon the kind of organizations collecting the data, whether it is a survey or Census, whether it is voluntary or mandatory, the target population (individuals/households, establishments or government), and how the data are collected (personal interviews, mail, telephone, or multimode).

The Sample

This project studies the "unit response rates" of statistical ICRs approved in 1998. When an agency submits an ICR to OMB for review, there are several items in the OMB Form 83-I (see <http://www.whitehouse.gov/omb/inforeg/83i-fill.pdf>) to be

completed by the agency to fulfill the requirements of Paperwork Reduction Act. The ICRs, except customer surveys, with the following three characteristics were selected to make up the universe of this study:

1. "**General purpose statistics**" was marked as "primary" in item 15 - Purpose of information collection.
2. "**Individuals or households**", "**Establishment** (Business or other for-profit, Not-for-profit institutions, Farms)" or "State, Local or Tribal **Government**" was marked as "primary" in item 11 - Affected public.
3. "**Voluntary**" or "**Mandatory**" was marked as "primary" in item 12 - Obligation to respond.

There were 130 ICRs that met these criteria. See table 1.

Table 1. Item 12: Obligation to respond			
Item 11: Affected public	Voluntary	Mandatory	Total
Individuals or households	26	5	31
Establishments (Business or other for-profit, Not-for-profit institutions, Farms)	38	40	78
State, Local or Tribal Government	19	2	21
Total	83	47	130

The agency contacts for all these ICRs were asked to provide the information on the OMB Response Evaluation Form (see Appendix A) and Data Collection Mode Form (see Appendix B). Several agencies (Census Bureau, NCHS, NCES, and EIA) bundled similar collections together in one ICR. As a result we have 216 potential information collections and they are used to evaluate response rates. See table 2.

Table 2: Information collections	
Eligible and returned	199
Ineligible	17
Total	216

Among these 216 collections, 10 were classified as ineligible by the agency, because they did not have the characteristics specified in this study, and the analyst excluded another 7 collections. Finally, a total of 199 information collections were analyzed in this study.

Methods

Factors Impacting the Response Rate

The following six items were collected for each survey,

1. c = Number of respondents completing the survey.
2. e = Number of potential respondents that were eligible but did not respond.
3. i = Number of potential respondents that were identified as being ineligible.
4. u = Number of potential respondents of unknown eligibility.
5. x = Estimated proportion of potential respondents of unknown eligibility that are eligible.
6. n = Total number of potential respondents in the survey (or population, if for census)

where $n = c + e + i + u$. If x is not provided, then the best estimate of $x = (c + e) / (c + e + i)$ will be used to calculate response rate, and our target variable, the response rate, RR , is calculated as $c / (c + e + xu)$.

If a weighted response rate was used in the information collection, the weighted response rate formula was provided by the agency for further research. Response rates were computed separately for several different characteristics:

1. **Primary function of the agency submitting the ICR** (ICSP: Interagency Council of Statistical Policy vs Non-ICSP);
2. **ICR type** (Census vs Survey);
3. **Obligation to respond** (Voluntary vs Mandatory);
4. **Affected public** (Households vs Establishments vs Government).
5. **Collection mode:** (Self-administered questionnaires; Personal or group interviews [including CAPI]; Mail survey; Telephone interview [including CATI]; Multi-mode [combination of the above and other methods].)

The combination of these factors will also be used to investigate their impact on response rate.

Across these 199 surveys, the mean unweighted response rate is 82.2% and the median unweighted response rate is 84.7%.

The distribution of the calculated unweighted response rate is displayed as

Table 3: Calculated unweighted response rate	Percentage achieved
Above 90%	37.7%
Above 80%	65.8%
Above 75%	73.4%
Above 70%	80.4%
Above 60%	90.0%
Above 50%	95.5%

There are 68 Surveys that have a calculated response rate below 80 percent. The above tables show that 65.8% (131/199) of ICRs have a calculated response rate above 80 percent; 73.4% (146/199) of ICRs have a calculated response rate above 75 percent; and 80.4% (160/199) of ICRs have a calculated response rate above 70 percent.

Single factor analysis

Six factors were characterized to evaluate their impacts on response rates. They are

I. Primary function of the agency submitting the ICR: ICSP (Agency is a member of Interagency Council of Statistical Policy) or Non-ICSP. The members of the Interagency Council on Statistical Policy are: Economic Research Service (ERS); National Agricultural Statistical Service (NASS); Bureau of Economic Analysis (BEA); Bureau of Census (BOC); National Center of Education Statistics (NCES); Energy Information Administration (EIA); National Center of Health Statistics (NCHS); Bureau of Justice Statistics (BJS); Bureau of Labor Statistics (BLS); Bureau of Transportation Statistics (BTS); Statistics of Income (SOI); Environmental Protection Agency (EPA); National Science Foundation (NSF); and Social Security Administration (SSA).

The distribution of the RR by "Primary function of the agency submitting the ICR" is shown in Table 4.

Table 4: Primary function the agency			
RR (no decimal)	ICSP	Non-ICSP	Total
[30%, 40%)	2	0	2
[40%, 50%)	5	2	7
[50%, 60%)	6	5	11
[60%, 70%)	17	2	19
[70%, 75%)	11	3	14
[75%, 80%)	13	2	15
[80%, 90%)	49	7	56
[90%, 100%]	56	19	75
Total	159	40	199
Average response rate	82%	82.8%	82.2%

II. ICR type: Census vs Survey; the distribution of the RR by (Census or Survey) is shown in Table 5.

Table 5. ICR type			
RR (no decimal)	Census	Sample	Total
[30%, 40%)	0	2	2
[40%, 50%)	0	7	7
[50%, 60%)	0	11	11
[60%, 70%)	0	19	19
[70%, 75%)	2	12	14
[75%, 80%)	1	14	15
[80%, 90%)	2	54	56
[90%, 100%]	13	62	75
Total	18	181	199
Average response rate	92.5%	81.8%	82.2%

III. Obligation to respond: Voluntary vs Mandatory; the distribution of the RR is shown below.

Table 6. Obligation to respond			
RR (no decimal)	Mandatory	Voluntary	Total
[30%, 40%)	2	0	2
[40%, 50%)	5	2	7
[50%, 60%)	2	9	11
[60%, 70%)	10	9	19
[70%, 75%)	4	10	14
[75%, 80%)	7	8	15
[80%, 90%)	18	38	56
[90%, 100%]	39	36	75
Total	87	112	199
Average response rate	82.8%	81.7%	82.2%

IV. Affected public: Households vs Establishments vs Government; the distribution of the RR is found in this table.

Table 7. Affected public				
RR (no decimal)	Individuals or households	Establishments	Government	Total
[30%, 40%)	0	2	0	2
[40%, 50%)	1	5	1	7
[50%, 60%)	6	5	0	11
[60%, 70%)	4	15	0	19
[70%, 75%)	5	9	0	14
[75%, 80%)	4	11	0	15
[80%, 90%)	13	34	9	56
[90%, 100%]	16	51	8	75
Total	49	132	18	199
Average response rate	79.6%	82.2%	88.9%	82.2%

V. Data collection mode:

The categories of "Data collection mode" are:

Self-administered questionnaires;

Personal or group interviews (including CAPI);

Mail survey;

Telephone interview (including CATI);

Multi-mode (combination of the above and other methods)

The distribution of the RR by "Data collection mode" is shown in Table 8.

Table 8: Data collection mode					
RR (no decimal)	Personal Interview	Mail Survey	Telephone interview	Multi-mode	Total
[30%, 40%)	0	2	0	0	2
[40%, 50%)	0	7	0	0	7
[50%, 60%)	1	4	2	4	11
[60%, 70%)	0	14	0	5	19
[70%, 75%)	3	3	1	7	14
[75%, 80%)	3	6	3	3	15
[80%, 90%)	3	29	2	18	56
[90%, 100%]	10	26	3	36	75
Total	20	91	11	73	199
Average response rate	84.7%	78.2%	80.2%	86.6%	82.2%

Note: There are 4 self-administered questionnaires that have response rates within [80%, 90%) with an average response rate of 85.3%.

Following the suggestion of Bob Grove to explore the relationship between the estimated response rate and the calculated response rate, I tried to establish a simple linear regression model using the estimated response rate in the ICR to predict the actual calculated response rate. The result was quite promising.

There were 82 information collections that reported an estimated response rate and the information of six items to calculate the achieved response rate. Using the calculated response rate as the dependent variable, y , and the estimated response rate as the independent variable, x , we fit the data into the simple regression model with zero y -intercept $y = b_1x$ since the calculated response rate should be zero when the estimated response rate is zero. The fitted simple linear regression equation is $y = 0.9901x$ with R -square = 0.998. There is an interesting scenario. When the information collection request states that the estimated response rate is 81%, using the fitted regression equation, the predicted response rate is 80% with standard error of 0.092, which is the threshold of approved expected response rate.

Conclusion

Recommendation to OMB about Response Rates

From the above analyses, the following statement about response rate is recommended to OMB in reviewing information collection requests:

1. Agencies that submit ICRs with an expected response rate of 80% or more should provide a complete description of how they arrived at the expected response rate.
2. Agencies that submit ICRs with expected response rates between 60% and 79% should provide a complete description of how they arrived at the expected response rate, a detailed description on steps they will take to maximize the response rate; and a discussion of how they plan to evaluate nonresponse bias.
3. Agencies that submit ICRs with expected response rates of less than 60% should generally not expect the ICR to be approved. However, agencies can, on occasion, justify conducting the information collection depending on the purpose of the study, the population being studied, past experience with response rates when studying this population, plans to evaluate nonresponse bias, and plans for an aggressive survey methodology to achieve at least 60%.

Maximizing Response Rates

Sponsors of surveys with predicted response rates of less than 80% should describe what they are doing to increase response by specific reference to the listed attributes of high response surveys.

A sample of efforts is summarized from those information collectors that achieved RR above 90%. Some tips for accomplishing this include:

For any survey:

- (1) Send a letter in advance to inform respondents of what, why, who, and how.
- (2) Provide a toll-free phone number for respondents to verify the legitimacy of survey.
- (3) Provide information about the survey on the agency's website.
- (4) Address confidentiality and anonymity.
- (5) Use more than 1 collection mode: e.g. face-to-face interviews in non-telephone households.
- (6) Mention a threat of fines for noncompliance in the mailout letter (for mandatory surveys only).
- (7) Keep the survey brief.
- (8) Make sure the survey is easily understood by respondents.
- (9) Try to make the content of the survey relevant to respondents.
- (10) Offer small incentives up front or a prize for responding (see Incentives Frequently Asked Survey Questions).
- (11) Identify strategies for contacting hard-to-reach populations.
- (12) Obtain endorsement of the survey by relevant organizations.
- (13) Conduct outreach sessions with presentations in several cities.
- (14) Provide news releases to trade journals, state associations, and other interested parties.
- (15) Maintain contact with respondents between waves in a longitudinal survey through holiday cards and postcards to be used in case of address changes.

For personal visit and telephone surveys:

- (16) Increase the number and timing of contact attempts.
- (17) Increase the length of the field period.
- (18) Enhance interviewer training.
- (19) Use senior, experienced interviewers to do "refusal conversion."

For mail surveys:

- (20) Follow-up the first mailout with a phone contact (or a Fax if a business).
- (21) Use Priority mail.
- (22) Address letters to specific individuals.

- (23) Send reminder/Thank you cards.
- (24) Send replacement questionnaires as part of nonresponse follow-up.
- (25) Follow-up phone calls to second mailing (or a Fax if a business).
- (26) Allow respondents to complete the survey on the web or via phone.

For internet surveys:

- (27) Use e-mail for advance notification, reminders, and follow-ups.
- (28) Allow respondents to complete the survey on a hardcopy (to mail in) or via phone.
- (29) Follow-up nonresponses with phone contact (or Fax if a business).

The evaluations conducted to assess the impact of possible non-response bias were requested in the OMB response rate evaluation form. A number of statements were reported and summarized. These are recommendations to agencies on how to assess nonresponse bias.

Evaluations Conducted to Assess the Impact of Possible Non-response Bias

Examples of suggested activities for information collectors that fail to meet the agency's target response rate:

In one case a large-scale evaluation was performed to examine nonresponse bias in the survey. It used information from the master files to compare respondents with nonrespondents and it used information from a nonresponse mail back follow-up to compare respondents and nonrespondents. The study found several interesting results. (1) Break off was most likely to occur at the stage of the telephone screener (43 percent) and that often the refusal is from the office staff rather than the respondent. This is consistent with information from the nonresponse follow-up that shows that a majority of nonrespondents do not remember being contacted about the survey. (2) A comparison of cooperation rates for many variables including respondent's specialty, gender, age, geographical region. (3) Effects of the observed differential nonresponse were not found to have much effect on the visit statistics produced by the survey as the nonresponse adjustment factor which takes respondent's specialty, and region into account, reduced the nonresponse bias for most of the visit statistics examined.

In one particular case, considerable resources were committed to studying potential nonresponse bias. A summary has been provided. 1. A comparison of estimates to extant data sources was conducted. 2. A Chi-square Automatic Interaction Detection (CHAID) analysis was

conducted on an extensive set of variables to determine which, if any, were significantly related to nonresponse.

(1) Use of the maximizing response rate methods discussed above (i.e., prioritizing nonresponse to follow-up on larger companies) results in a weighted response rate much higher than the unweighted response rate. While the unweighted response rate is approximately 80% of the eligible companies, the respondents are estimated to account for over 95% of the data of interest; (2) For the nonrespondents, agency imputations using previously reported data adjusted for changes and general trends in the industry.

Usage of Weighted Response Rate:

The distribution of the RR by "Usage of weighted response rate" is shown in Table 9.

RR (no decimal)	Weighted RR used	Weighted RR not used	Un-known	Total
[30%, 40%)	0	0	2	2
[40%, 50%)	0	1	6	7
[50%, 60%)	2	6	3	11
[60%, 70%)	0	5	14	19
[70%, 75%)	4	6	4	14
[75%, 80%)	1	8	6	15
[80%, 90%)	14	20	22	56
[90%, 100%]	3	47	25	75
Total	24	93	82	199
Average response rate	80.9%	85.7%	78.5%	82.2%

"Used" means this type of response rate was used in the ICR or in publication.

Recommendation to OMB on Multi-stage and Longitudinal Surveys

When agencies submit an ICR with multi-stage or longitudinal surveys, the following is recommended to OMB in reviewing those information collection requests:

1. Ask the agency to provide the expected response rate at each stage of samplings.
2. Ask the agency the expected total response rate taking all stages into account.
3. The agency should always describe the steps to improve response rate.
4. The agency should include a discussion of how they plan to evaluate nonresponse bias.

The final recommendation to Form 83-I revision is to add one item for expected response rate, and ask the agency to provide supporting material when submitting Information Collection Requests.

References:

American Association for Public Opinion Research (May 1998). *Standard Definitions*.

Couper, M. & Groves, R. (1996). Household-Level Determinants of Survey Nonresponse. . *New Directions for Evaluation: Advances in Survey Research*, 70, 63-80.

Appendix A

OMB Response Evaluation Form

OMB control number _____ - _____

Survey ID

This survey was (mark one)

- A. Completed
- B. Not conducted
- C. Other – please describe

If you marked “A” above, please provide the following information for the completed survey.

1. Number of respondents completing the survey.
2. Number of potential respondents that were eligible but did not respond.
3. Number of potential respondents that were identified as being ineligible.
4. Number of potential respondents of unknown eligibility.
5. Estimated proportion of potential respondents of unknown eligibility that are eligible.
6. Total number of potential respondents in the survey (or population, if for census)
7. Describe all efforts to maximize the response rate.

8. Describe evaluations conducted to assess the impact of possible nonresponse bias.

Appendix B

OMB Response Evaluation Form - data collection mode

OMB control number _____ - _____

Survey ID

1. The data collection mode in this ICR:

The data collection mode was (mark one)

- _____ A. Personal or group interview (including CAPI)
- _____ B. Mail Survey
- _____ C. Telephone interview (including CATI)
- _____ D. Some other methods (e.g. Fax, ...) , please specify

- _____ E. Multi-mode (combination of the above)

2. Weighted response rate:

_____ Yes, a weighted response rate was used in this ICR and the weighted response rate formula was

_____ .

and the calculated weighted response rate in 1998 was

_____ .

_____ No, we did not use weighted response rate in this ICR.