# USING ALTERNATIVE QUESTION STRATEGIES TO REDUCE INCOME NONRESPONSE

Jeffrey C. Moore and Laura S. Loomis[1/], U.S. Census Bureau
Jeffrey C. Moore, US Census Bureau, Center for Survey Methods Research/SRD, Washington, DC 20233-9150

**Key Words:** questionnaire design, income nonresponse, response brackets, QDERS

## 1. Introduction and Background

### 1.1 Income Nonresponse

Item nonresponse is a continuing problem for surveys that collect income data, including U.S. government surveys, despite their generally high overall response rates. Moore, Stinson, and Welniak (1999), for example, citing data from the Current Population Survey (CPS), report income nonresponse rates ranging from 20% up to almost 50%, with questions about income from assets showing the highest rates of nonresponse. As with any survey topic, elevated levels of nonresponse necessitate costly and time-consuming repair measures both in the field and in data processing, and increase the uncertainty surrounding income estimates derived from survey data. The importance of income data for social policy analysis adds an extra impetus for developing effective techniques to bring income nonresponse under better control.

Various reasons have been advanced to explain the high levels of income item nonresponse and other reporting problems: the complexity of the reporting task, confusion and uncertainty about income source labels and other terms, sensitivity, (Moore, Stinson, and Welniak, 1999), and, perhaps especially for asset income, lack of knowledge about income amounts, due to lack of salience and recall difficulties (Cantor, Brandt, and Green, 1991).

### 1.2 Questionnaire Design Solutions

Respondents' lack of knowledge and the sensitivity of income questions have proven difficult to combat. Recently, however, researchers have begun to report some success in reducing income nonresponse using a technique called "unfolding brackets" – asking a series of closed-ended income range questions (e.g., "Is it $10,000 or more?"), after the respondent has refused or said "don't know" to an open-ended exact amount question. Using this technique, which slowly "unfolds" a range containing the amount in question, several studies have shown that relatively large proportions of respondents who initially refuse or don't know the exact answer to an income question *will* answer follow-up income range questions (Juster and Smith, 1997; Kennickell, 1997; Ross and Reynolds, 1996; Heeringa, Hill, and Howell (1995); Hippler and Hippler, 1986; Bell, 1984). Another benefit of the bracket strategy is the increased willingness of relatively wealthy households to answer these questions, thus improving measurement of income at the upper end of the income distribution. When researchers have used information from bracketed questions in imputation procedures, much higher estimates of wealth and asset amounts have resulted, which is generally assumed to mean much better estimates (Juster and Smith, 1998, 1997; Kennickell, 1997).

While the unfolding brackets technique appears to have clear benefits, those benefits are not cost-free. One disadvantage is the tedium of having to read long lists of categories, or the painfully slow "unfolding" of the correct income category. In-person interviews can make use of show-cards to display income brackets, but telephone interviews are constrained to reading the series of categories or unfolding bracket questions. This is particularly problematic for surveys like the Census Bureau's Survey of Income and Program Participation (SIPP), and the March Income Supplement to the CPS, which focus much attention on income details, and thus ask a great number of amount questions about a great number of income sources. Also, the unfolding bracket technique is implemented after the respondent has already indicated reluctance to answer, running the risk of badgering the respondent. Again, this problem is exacerbated in income-focused surveys – interviewers in Kennickell's (1997) study, for example, which had a large number of income questions, complained that the follow-up bracket questions were "too pushy."

Another potential cost is loss of precision – the risk that exposure to bracket questions might affect respondents' willingness to answer later exact amount questions. The evidence on this issue is mixed. Juster and Smith (1997) found that respondents exposed to bracket questions were actually *more* likely to respond to subsequent exact amount questions; Kennickell (1997), in contrast, found that once they were offered as an option, the use of ranges persisted throughout the interview.

Of course, questionnaire design decisions may also affect interviewers' behaviors. In this case, interviewers, too, may come to over-rely on the bracket questions and, in the face of some respondent reluctance to answer an exact

income question, invoke them too readily in place of traditional probing techniques. Kennickell (1997) offers this as an explanation for the sharp decline he observed in exact answers to Survey of Consumer Finances amount questions following the introduction of bracketed question techniques.

A final concern about the unfolding brackets technique is the potential for anchoring effects - the dollar amount used as the initial entry point into the sequence of bracketing questions can influence the distribution of responses. Hurd and Rodgers (1998) found that the use of higher dollar amounts as entry points skewed the distribution of responses toward higher amounts.

### 1.3 Goals of the Current Research

This paper describes research which builds on the unfolding brackets approach, testing a new form of income range reporting, which we label "implicit brackets." The goals of the new procedure are: (1) reducing sensitivity concerns, by not forcing exact amount reports; (2) reducing cognitive burden, by abandoning almost all efforts to obtain details about very small amounts, and by allowing reasonable estimations; (3) avoiding the tedium of long lists of categories for each amount question, or multi-step question sequences to zero-in on appropriate amount brackets; while at the same time (4) allowing precise, point-estimate reports, if respondents are willing and able to provide them.

## 2. Methods and Procedures

### 2.1 The Questionnaire Design Experimental Research Survey (QDERS)

This research was part of the initial launch of the Census Bureau's Questionnaire Design Experimental Research Survey (QDERS), a special vehicle developed by Bureau staff for conducting questionnaire design research "off-line" from the agency's ongoing production surveys. The first QDERS, fielded in April 1999, included data collection experiments on disabilities, health insurance coverage, non-wage income sources, asset ownership, asset income, and household relationships. (See U.S. Census Bureau (1999) for a general description of QDERS and the 1999 implementation specifically.) This paper focuses on the asset income amount experiment.

### 2.2 Sampling and Experimental Design

QDERS was a split-sample controlled experiment, using paper-and-pencil questionnaires in a telephone interview. It used a nationally representative (excluding Alaska and Hawaii) RDD sample, with independent sub-samples for each treatment (GENESYS, 1997). Interviews were conducted with one household respondent, who reported for him/herself and up to five other adult household members. The initial sample consisted of 5,870 telephone numbers, pre-screened to identify working residential banks of numbers.

### 2.3. Questionnaires

The income amounts reporting experiment used two questionnaires - a standard "control" treatment, and the experimental "implicit brackets" approach. As noted, these were paper-and-pencil questionnaires, administered by telephone. The content of the income questions in each treatment was identical; only the manner in which the questions were asked differed. Both treatments included asset ownership questions, asked of all adult household members, consisting of the following five commonly-owned asset types: interest-earning checking accounts, savings accounts, certificates of deposit (CDs), mutual funds, and stocks. QDERS procedures called for income amounts to be collected only for assets owned by the household "reference person," the primary adult owner/renter of the sample unit[2].

#### 2.3.1 Control treatment

The control treatment amounts questions were quite straightforward. For each asset type reported in the "asset ownership" section of the interview, a question of the following form was asked: "How much interest did [NAME] earn from [asset type] in 1998?"

#### 2.3.2 Experimental, "implicit brackets" treatment

The question format for the experimental treatment consisted of two parts. The first asked whether the annual income amount for 1998 was "more or less than $X?," where X was a minimum amount which varied according to asset type. The goal was to vastly reduce the burden associated with reporting trivial amounts, by substituting a very easy judgment task for the usual exact amount question[3].

The implicit bracket procedure was invoked if the response to the initial "more/less" question was "more." The procedure was simply to ask a question of the form: "How much was it to the nearest $Z?," in effect establishing response brackets, of width Z, without having to present those brackets overtly. As with the minimum amount (see above), the specific value of Z

varied with different asset types. (See Moore and Loomis (2000) for specific examples of the two questionnaire formats, including the X (minimum value) and Z (bracket width) values for each asset type.)

2.4 Data Collection

2.4.1 Interviewers and interviewer training

A staff of 22 experienced telephone interviewers received approximately five hours of initial QDERS training, covering both income amount reporting treatments, and a brief refresher training session midway through the field period[4]. Throughout the field period interviewers' assignments included a mix of both treatments.

2.4.2 Response rates

As is generally true of RDD surveys, the unknown eligibility of "ring, no-answer" cases makes a precise estimate of the QDERS response rate impossible. Using American Association for Public Opinion Research (AAPOR, 1998) guidelines, we calculate a response rate range for QDERS of 36-46%. QDERS procedures did not include any special refusal conversion attempts, and as a result refusals were plentiful, accounting for about half of the observed non-response (30% of all cases). For both the response rate and refusal rate statistics the differences between the two income amount reporting treatments were trivial and nonsignificant. Interviewers completed interviews in 1,304 households, of which 13 were subsequently excluded due to missing data, for a final total of 1,291 completed interviews.

## 3. Results

3.1 Item Nonresponse

Table 1 summarizes the nonresponse results of the amount reporting experiment. The conclusion is clear: for all five asset income sources included in QDERS, the observed nonresponse rate for the experimental treatment is lower than the observed rate for the control treatment. The difference is statistically significant for only one of the five individual comparisons[5]; however, a simple sign test (Snedecor and Cochran, 1967) suggests that the complete, five-out-of-five consistency of the direction of differences across the individual comparisons is itself statistically significant (p=.0625), indicating that rejection of the null hypothesis is appropriate.

Additional analyses (not shown) reveal that the primary nonresponse problem for QDERS asset income amounts is "don't know's" (DK), which generally outnumber refusals by a four- or five-to-one ratio or more. More importantly for present purposes, they also reveal that the improvement in nonresponse in the experimental treatment is due to a reduction in DK nonresponse, and not to any improvement with regard to refusals. In all five comparisons, the control treatment suffered a higher rate of DK nonresponse than the experimental treatment; no such effect is evident for refusals. Thus, it appears that the implicit brackets approach made important inroads on resolving cognitive barriers to asset income amount reporting, but did not have the desired impact with regard to sensitivity concerns.

3.2 Response Quality Comparisons

Did the extreme ease of the "minimum value" component of the implicit brackets approach bias respondents toward reporting very low amounts? To address this question we compared the similarity of amount report distributions after recoding the continuous amount reports into four categories. We assigned categories by determining the approximate quartile values – 25%, 50%, and 75% – for the full array of reports, including both treatments. The results of this categorization, with "quartile 1" the lowest quartile, and "quartile 4" the highest, are summarized, by treatment, in Table 2.

According to a chi-square test, only for stocks do the distributions of amount reports differ significantly by questionnaire treatment. There is, however, no indication of a greater tendency for the proportion of experimental treatment cases in the lowest category to significantly exceed the comparable proportion in the control treatment – in fact, for the lowest quartile of responses the significant difference is in the opposite direction. The other asset types show a mix of non-significant treatment differences in the lowest amount category. Thus, the experimental treatment does not seem to have elicited more very low amount reports than the control treatment, and in general the response distributions do not seem to differ by questionnaire treatment.

Another potential consequence of the unfolding bracket technique is a reduction in precise, point-estimate reports, and an increase in the use of ranges, even as a first-choice response (e.g., Kennickell, 1997). Although the implicit brackets procedure does not use ranges explicitly, and is not a fallback response option which attempts to salvage information from an initial nonresponse, it is still of concern that the implicit invoking of ranges ("... to the nearest $10?") might reduce the precision of respondents' answers. In fact, an impetus for this procedure was the

desire to signal to respondents with sensitivity concerns, or who were uncertain as to an exact amount, that extreme precision was not necessary, and that it was acceptable to report a rounded amount.

We find, however, in an examination of the tendency to elicit rounded amount reports, no evidence of less precision in experimental treatment reports – in fact, the experimental treatment seems to have actually increased report precision (data not shown). It appears that the experimental treatment may have had an effect opposite to what was intended, and sent stronger signals for precision rather than weaker ones, (which may explain its failure to reduce refusal nonresponse).

Our final data quality measure concerns the reliability of the amount reports. QDERS procedures included a response variance reinterview to permit a comparison of the reliability of the data produced by the two forms of the amount report questions (see Hess et al., 2000, for details). All 1,291 originally-interviewed households were eligible for reinterview; altogether, 1,088 re-interviews were completed, or about 84% of all original interviews. By either of two standard measures of reliability – the "index of inconsistency" and the "gross difference rate" – the differences between the control and the experimental procedure were very small, and none approached statistical significance. In other words, the two treatments were roughly equivalent with regard to the random "noisiness" of the responses they produced.

### 3.3 Interviewers' Evaluations

At the end of the QDERS field period we administered a brief questionnaire to the interviewers, three items of which dealt explicitly with the income amount question experiment. Although there was no evidence that interviewers perceived one or the other form to be more likely to "encourage Rs to be careless," the two other items displayed interesting effects (data not shown; see Moore and Loomis, 2000, for details).

First, interviewers rated the experimental treatment significantly more positively on an "easy for Rs to answer" scale – reporting an average score of 1.4 (where 1 = "strongly agree" and 5 = "strongly disagree") for the experimental treatment, versus 3.7 for the control treatment. Interviewers clearly perceived a difference between the two amount question formats in terms of the burden they imposed on respondents.

The third evaluation item asked interviewers to assess whether each treatment "collected accurate answers." On this dimension, too, interviewers reported that they could perceive clear differences between the treatments, again in favor of the implicit brackets design. Using the same strongly agree - strongly disagree scale as before, the average scale rating assigned to the control treatment was 3.2, significantly higher (less perceived accuracy) than the 1.9 average score for the experimental treatment

## 4. Summary and Conclusions

In this test, the experimental, implicit brackets approach showed some significant advantages over a more standard approach to obtaining survey reports of asset income amounts. Most prominently, using the implicit brackets approach resulted in a reduction in item nonresponse, primarily through a reduction in "don't know" nonresponse. Reports were, if anything, more rather than less precise in the experimental treatment. These two major findings in the survey data are supported by interviewers' subjective judgments that the implicit brackets approach was easier for respondents and produced more accurate reports. We also find no important differences in the reliability of the income reports elicited by the two treatments.

We view the QDERS results as a positive step in the development of improved procedures for capturing survey reports of income amounts. The necessary next steps are, first, the refinement of procedures for reducing sensitivity concerns, and second, the replication of a test of the implicit brackets approach in a survey setting without QDERS' limitations – i.e., a larger scale, non-RDD sample survey, perhaps administered in-person, and certainly with a more satisfactory response rate.

## References

American Association for Public Opinion Research (1998), "Standard Definitions – Final Dispositions of Case Codes and Outcome Codes for RDD Telephone Surveys and In-Person Household Surveys." *[http://www.aapor.org/ethics/stddef.html]*

Bell, R. (1984), "Item Nonresponse in Telephone Surveys: An Analysis of Who Fails to Report Income." Social Science Quarterly. 65: 207-215.

Cantor, D., S. Brandt, and J. Green (1991), "Results of first wave of SIPP interviews." Unpublished Westat report to the U.S. Bureau of the Census (memorandum to Chet Bowie), February 21, 1991.

GENESYS Sampling Systems (1997), "GENESYS Sampling System Methodology." GENESYS Sampling Systems, Fort Washington, PA.

Heeringa, S., D. Hill, and D. Howell (1995), "Unfolding Brackets for Reducing Item Nonresponse in Economic Surveys." Health and Retirement Study Working Paper Series, Paper No. 94-029. Survey Research Center, Institute for Social Research, University of Michigan.

Hess, J., C. Keeley, J. Moore, J. Pascale, and J. Rothgeb (2000), "The Effects of Person-level vs. Household-level Questionnaire Design on Survey Estimates and Data Quality." Paper presented at the annual meetings of the American Association for Public Opinion Research, Portland, OR, May 18-21, 2000.

Hippler, H., and G. Hippler (1986), "Reducing Refusal Rates in the Case of Threatening Questions: The 'Door-in-the-Face' Technique." Journal of Official Statistics. 2(1): 25-33.

Hurd, M., and W. Rodgers (1998), "The Effects of Bracketing and Anchoring on Measurement in the Health and Retirement Study." *[source unknown]*

Juster, F., and J. Smith (1997), "Improving the Quality of Economic Data: Lessons from the HRS and AHEAD." Journal of the American Statistical Association. 92(440): 1268-1278.

Juster, F., and J. Smith (1998), "Enhancing the Quality of Data on Income and Wealth: Recent Developments in Survey Methodology." Paper prepared for the 25th General Conference of the International Association for Research in Income and Wealth. Cambridge, England, August 23-29, 1998.

Kennickell, A. (1997), "Using Range Techniques with CAPI in the 1995 Survey of Consumer Finances." Survey of Consumer Finances Working Paper. Board of Governors of the Federal Reserve System, Washington, DC.

Moore, J. and L. Loomis (2000), "Using Alternative Question Strategies to Reduce Income Nonresponse." Unpublished U.S. Census Bureau report, May 30, 2000.

Moore, J., L. Stinson, and E. Welniak (1999), "Income Reporting in Surveys: Cognitive Issues and Measurement Error." Chapter 10 in Sirkin, M., D. et al. (eds.), Cognition and Survey Research. New York: John Wiley & Sons, 1999.

Ross, C., and J. Reynolds (1996), "The Effects of Power, Knowledge, and Trust on Income Disclosure in Surveys." Social Science Quarterly. 77(4): 899-911.

Snedecor, G. and W. Cochran (1967), Statistical Methods. Ames, IA: The Iowa State University Press.

U.S. Census Bureau (1999), "QDERS Training Guide." Unpublished U.S. Census Bureau document prepared for the 1999 QDERS project, (undated, spring 1999).

**Notes**

1. This paper reports the results of research and analysis undertaken by Census Bureau staff. It has undergone a more limited review than official Census Bureau publications. This report is released to inform interested parties of research and to encourage discussion.

2. The complexity of the QDERS questionnaires caused interviewers to collect income amount information mistakenly for many non-"reference persons" as well. Because these extra income reports derived equally from both income reporting treatments, we elected to include all income reporters in our analyses, both intended and non-intended, exploiting the procedural errors in the interests of increased statistical power.

3. In fact, there were two subcomponents to this initial task. The first was a lead-in question identical to the control treatment amount question (e.g., "How much interest did you earn from all CDs in 1998?"). Interviewers were instructed to pause briefly at the end of this lead-in, to allow respondents who were willing and able to supply an exact amount to do so. The second part was the "more or less than" question itself (e.g., "Would you say it was more or less than $50?").

4. The refresher training session coincided with a shift in interviewers' assignments with regard to a different QDERS experiment involving the use of household screener questions (see Hess, et al., 2000).

5. As noted, our analyses use the full set of income amount respondents, including both reference persons and non-reference-persons, the latter of whom were administered the questions in error. In this case - and in general throughout these analyses - the results of an analysis of the more restricted reference-person-only data are completely consistent with the larger data set.

**Table 1: Asset Income Amount Item Nonresponse, by Asset Type and Question Treatment**

| Asset Type: | Question Treatment | | | | |
|---|---|---|---|---|---|
| | Control | | | Experimental | |
| | % nr | (n) | | % nr | (n) |
| Checking Accts | 48.8 | (301) | > | 43.7 | (348) |
| Savings Accts | **50.6**[1/] | (425) | > | **44.5**[1/] | (425) |
| CDs | 62.2 | (119) | > | 57.0 | (128) |
| Mutual Funds | 57.4 | (148) | > | 52.2 | (136) |
| Stocks | 50.9 | (159) | > | 47.8 | (136) |

1/ chi-square=3.19, 1df, p<.10

**Table 2: Percent Distributions of Reported Dollar Amounts ("Quartile" Recode), by Asset Type and Question Treatment**

| Asset Type: | | Question Treatment | | | chi-square |
|---|---|---|---|---|---|
| | | Control | | Experimental | |
| Checking Accts | quartile 1 | 18.8 | < | 27.0 | n.s. |
| (n)C=154 | quartile 2 | 29.9 | > | 24.0 | |
| (n)E=196 | quartile 3 | 26.0 | > | 24.0 | |
| | quartile 4 | 25.3 | > | 25.0 | |
| Savings Accts | quartile 1 | 21.0 | < | 22.9 | n.s. |
| (n)C=210 | quartile 2 | 30.5 | > | 30.1 | |
| (n)E=236 | quartile 3 | 24.8 | > | 23.7 | |
| | quartile 4 | 23.8 | > | 23.3 | |
| CDs | quartile 1 | 20.0 | < | 27.3 | n.s. |
| (n)C=45 | quartile 2 | 24.4 | < | 29.1 | |
| (n)E=55 | quartile 3 | 24.4 | > | 23.6 | |
| | quartile 4 | 31.1 | > | 20.0 | |
| Mutual Funds | quartile 1 | 28.6 | > | 26.2 | n.s. |
| (n)C=63 | quartile 2 | 22.2 | < | 27.7 | |
| (n)E=65 | quartile 3 | 20.6 | < | 26.2 | |
| | quartile 4 | 28.6 | > | 20.0 | |
| Stocks | quartile 1 | **35.9**[1/] | > | **11.3**[1/] | 34.49, 3df, p<.001 |
| (n)C=78 | quartile 2 | **6.4**[2/] | < | **46.5**[2/] | |
| (n)E=71 | quartile 3 | 30.8 | > | 22.5 | |
| | quartile 4 | 26.9 | > | 19.7 | |

1/ t=3.71, 147df, p<.001 2/ t=6.09, 147df, p<.001