Introducing Brackets: Quality in the Consumer Expenditure Interview Survey

GEOFFREY PAULIN

Nonresponse is a problem in surveys. Some potential participants may refuse to participate at all in a survey, while others may provide answers to some, but not all, questions asked. For those who participate at least partially, reasons for not responding to certain questions may include the sensitivity of the respondent to the question asked or simply a lack of knowledge on the part of the respondent. One situation in which either of these two reasons may be cited is when respondents are asked about income levels and sources. Some respondents may refuse to answer questions about income because they consider the matter too personal to divulge. Others may be willing to answer, but may not be able to do so completely, because they lack specific or detailed knowledge. This is often the case in "proxy reporting," wherein the respondent reports income information for another member of the consumer unit.[1] For example, a parent may not know precisely the amount of income earned by a teenaged daughter who is employed after school at a neighborhood fast-food restaurant.

In the case of complete refusal to participate in the survey, little can be done to obtain information. By contrast, as regards sensitive questions or lack of knowledge, information may be gained by allowing the respondent to give an answer that is not precise. For example, a person earning a salary of $300,000 may refuse to divulge that information precisely, but may be comfortable saying that the salary is "greater than $120,000." Similarly, the aforementioned parent may not know the precise salary of his teenaged daughter, but may know with confidence that it is "less than $5,000" per year. Prior to the second quarter of 2001, such information was lost in the Consumer Expenditure (CE) Interview survey, because the respondent could only report a value, assert "don't know," or refuse to answer. However, starting in April 2001, respondents were given the opportunity to provide an income range, or "bracket," when they were unable or unwilling to give a specific value. This article describes the collection of income data and the development of income brackets in the CE Interview survey.

Income data are collected in the second and fifth interviews for those who participate in those interviews. If the consumer unit does not complete its second interview (for example, if the family is unavailable during the survey period or if the family originally residing at the address during the second interview has moved away and the new

Geoffrey Paulin is a senior economist in the Branch of Information and Analysis, Division of Consumer Expenditure Surveys, Bureau of Labor Statistics.

[1] See "Glossary" in Appendix A at the end of this anthology for the definition of a *consumer unit*.

residents are now participating instead), the information is collected at the earliest possible interview (the third, fourth, or fifth). In either case, incomes are collected for the past year, as determined by the date of the interview. For example, a consumer unit interviewed in July 2002 would have been asked to recall income received from July 2001 to June 2002.

Data are collected on several sources of income. Some of these, such as data on wages and salaries, are collected for members of the consumer unit who are at least 14 years old. Others, such as information on interest income, are collected for the consumer unit as a whole. In addition to data on "labor" (wage and salary or self-employment) income and "nonlabor" (interest or dividend) income, information on other sources (such as alimony, child support, Food Stamps, and welfare income) also is collected. (For a complete listing of sources, see the appendix to this article.)

History of bracketing in the Interview survey

In May 1998, a 2-day seminar was held at Princeton University to discuss the utility of the CE Survey for measuring poverty and related issues. During the course of the seminar, many ideas for improving the quality of the data were proposed. One of these was to investigate the use of brackets for collecting data on income, assets, and liabilities, because these data are important, but frequently missing. Katharine G. Abraham, Commissioner of the Bureau of Labor Statistics (BLS) at the time, asked her organization's Division of Consumer Expenditure Surveys to study the feasibility of collecting bracketed data, starting with the 2000 survey.

In September 1998, a team was chartered to investigate and recommend strategies for the implementation of bracketing if it was deemed feasible. The team had two major questions to answer: first, does bracketing reduce nonresponse in practice? Second, which type or types of brackets, if any, should be used? Starting with a review

of the literature on the subject, the team discovered that, as expected, bracketing was useful for collecting data, because respondents with imprecise knowledge could provide at least some information. However, one unintended consequence described in the literature is that bracketing can lead to a loss of precision, because some respondents who report bracketed data might have reported actual values if the interviewer had probed sufficiently.[2] In addition, the team reasoned that brackets would increase respondents' burden, because, without them, a respondent could simply refuse to answer or respond "I don't know," and the next question would be asked. With brackets, once either of these occurs, the interviewer attempts to collect a bracketed value. Still, the team concluded that brackets would be useful despite these concerns. For example, the loss of precision might be outweighed by an increase in overall response when brackets were used. Interestingly, the literature also supported the hypothesis that brackets do *not* seriously increase respondents' burden: although it is true that there is one more question in cases where the initial response is "I don't know" or a refusal, it also is true that a large number of those who initially respond in either of those ways is subsequently willing and able to provide a bracketed value.[3]

Constructing the brackets

Given that brackets are indeed useful in data collection, the second question becomes operative. The team discovered that there are at least two types of brackets used in practice: "conventional" brackets and "unfolding" brackets. With both types, the respondent is first asked for a specific value. If he

[2] Kennickell, Arthur B., "Using Range Techniques with CAPI in the 1995 Survey of Consumer Finances" on the Internet at **http://www.federalreserve.gov/Pubs/oss/oss2/papers/rangepap0197.pdf**, January 1997.

[3] Juster F. Thomas and James P. Smith, "Improving the Quality of Economic Data: Lessons from the HRS and AHEAD," *Journal of the American Statistical Association,* vol. 92, no. 440, December 1997, pp. 1268–1278.

or she is unable to provide one, then, in a conventional-bracketing framework, the respondent is asked to identify, from a predetermined list, the range in which the income or asset is likely to fall (for example, less than $5,000; $5,000 to $9,999; $10,000 to $19,999; and so forth). In an unfolding-bracketing framework, the respondent is asked a series of questions designed to elicit ranges of values. For instance, the interviewer might say, "Is it at least $5,000?" If the response is "No," then a range of less than $5,000 would be recorded. If the response is "Yes," then the respondent would be asked, "Is it at least $10,000?" If "No," then a range of $5,000 to $9,999 would be recorded. If "Yes," the respondent would be asked, "Is it at least $20,000?" If "No," then a range of $10,000 to $19,999 would be entered. If "Yes," then a response of "at least $20,000" would be recorded, and the next question in the survey would be asked. The team recommended that conventional bracketing be adopted, for a couple of reasons: first, more precise answers would be obtained. (For some sources of income, such as wages and salaries, it is likely that a large percentage of recipients could accurately respond that their income from those sources was "at least $20,000"; narrower ranges, such as $20,000 to $29,999 and so forth, allow a more precise estimate of the value of such income.) Second, conventional brackets were thought to be less burdensome, because the respondent could be handed a card with the appropriate ranges and quickly scan it to find which was appropriate for the source in question. With unfolding brackets, the respondent might be asked three additional questions, instead of one.

Once the type of bracketing was selected, the next question was what the ranges of the brackets should be. One idea was to use standard publication ranges as a guide. For example, data currently are published for families whose total income is less than $5,000; $5,000 to $9,999; $10,000 to $14,999; and so forth. However, the Interview survey collects information from a variety of sources, some for each

member aged 14 and older, some for the consumer unit as a whole. The publication ranges may be appropriate for some sources of income (for instance, wage and salary income), but may not be appropriate for other sources. For example, almost all respondents who reported interest income reported a value less than $5,000, so, for this source, the publication range is too broad to be meaningful. To determine the most useful ranges, the distribution of each source was analyzed. Then, through a combination of empirical examination and normative analysis, a few sets of brackets were developed to fit the different kinds of data. The empirical examination involved looking at the percentiles for each source of income and seeing where breaks occurred. Normative analysis involved finding "reasonable" cutoff values for the data.

Refining the brackets, using the BLS cognitive laboratory

The next step in the implementation process required testing the results in the BLS cognitive laboratory. At this stage, a new team was formed that included a member of the Survey Research branch of the Division of Consumer Expenditure Surveys and a cognitive psychologist from the BLS Office of Survey Methods Research. Cognitive psychologists are trained in how respondents perceive certain questions. That is, when the interviewer asks about interest income, does the respondent correctly perceive what the interviewer is asking for (such as interest earned on checking and savings accounts), or might the respondent be confused and include other sources of income (such as dividends from stocks), or might the respondent even report no income received, when, in fact, he or she did receive such income, but thought it was something else? In the cognitive laboratory, tests are performed in which respondents are asked

for their answers and then are debriefed by the psychologist. During the testing, the psychologist might ask the respondent to define certain terms, to make sure that the respondent's definition matches the interviewer's; or the respondent might be asked questions about the survey in general—were the questions posed easy or difficult to understand and answer, for example.

After the brackets were refined on the basis of findings from the cognitive tests, the brackets were ready to be implemented. Various steps were involved in their implementation, including revising the survey instrument designed to collect the data, field-testing the instrument, and obtaining appropriate approvals from offices that regulate Government surveys. Bracketing finally appeared in the CE Interview Survey in the second quarter of 2001. That is, the first respondents to the survey who were asked to provide bracketed information began their participation in April 2001.[4] Currently, only income brackets have been implemented. The original team investigated the possibility of using brackets for assets and liabilities as well, but decided to start with income only and then apply any lessons learned therefrom to the implementation of assets and liabilities.

Conclusions

At present, the first year (2001) of data gathered with the use of brackets has been published, and a new team has been chartered to study how brackets have changed the collection of income data. Among the questions being investigated are the following: are many "don't knows" and refusals to answer

being converted to bracketed values? Have brackets improved the percentage reporting various sources of income? Has average income reported risen as a result of using brackets? and Are there any demographic differences in the propensity to provide bracketed information? As these issues are analyzed, further research results will be published documenting the findings.

APPENDIX: Income Sources and Bracket Ranges

Data on the following sources of income are collected for each individual member of the consumer unit who is at least 14 years old: Wages or salary; income (or loss) from nonfarm business, partnership, or professional practice; income (or loss) from own farm; Social Security or Railroad Retirement Income; and Supplemental Security Income.

The following sources of income are collected for the consumer unit as a whole: Unemployment compensation; workers' compensation and veterans' payments, including education; public assistance or welfare, including money received from job training grants such as Job Corps; Food Stamps and electronic benefits transfers; interest on savings accounts or bonds; regular income from dividends, royalties, estates, or trusts; pensions or annuities from private companies, the military, or government; income (or loss) from roomers or boarders; income (or loss) from payments from other rental units; child support; regular contributions from alimony or other sources, such as persons outside the consumer unit; and other money income, including money received from care of foster children, cash scholarships, fellowships, or stipends not based on working.

Table 1 shows the brackets applied by the interviewer to each source of income. ∎

---

[4] Although the initial goal was for implementation in 2000, it became apparent that to implement bracketing properly would require cognitive testing and other processes. Therefore, the implementation was delayed until 2001.

Table 1. Data collected by source of income and income brackets

Income source	Range for bracket—												
	0	1	2	3	4	5	6	7	8	9	10	11	12
Wages and salaries; nonfarm business income; own-farm income	Loss[1]	0– 4,999	5,000– 9,999	10,000– 14,999	15,000– 19,999	20,000– 29,999	30,000– 39,999	40,000– 49,999	50,000– 69,999	70,000– 89,999	90,000– 119,999	120,000 or more	—
Social Security and Railroad Retirement Income	—	Less than 300	300– 399	400– 499	500– 599	600– 699	700– 799	800– 899	900– 999	1,000– 1,499	1,500 or more	—	—
Supplemental Security Income	—	0– 999	1,000– 1,999	2,000– 2,999	3,000– 3,999	4,000– 4,999	5,000– 9,999	10,000– 14,999	15,000– 19,999	20,000– 29,999	30,000– 39,999	40,000– 49,999	50,000 or more
Data collected for the consumer unit as a whole	Loss[2]	0– 999	1,000– 1,999	2,000– 2,999	3,000– 3,999	4,000– 4,999	5,000– 9,999	10,000– 14,999	15,000– 19,999	20,000– 29,999	30,000– 39,999	40,000– 49,999	50,000 or more

[1] Self-employment income only
[2] Rental income only

NOTE: The foregoing brackets shown for sources collected for the consumer unit as a whole are also collected for the following sources of money, which are not considered income in the CE Interview survey: lump-sum payments received from estates, trusts, royalties, alimony, prizes, or games of chance or from persons outside the consumer unit; and sales of household furnishings, equipment, clothing, jewelry, and pets or other belongings, excluding the sale of vehicles or property.