METHODOLOGICAL ISSUES IN THE COLLECTION
OF CONSUMER EXPENDITURES IN THE UNITED STATES
ADRIANA SILBERSTEIN
BUREAU OF LABOR STATISTICS (BLS)
Paper prepared for the Methodology Workshop at Statistics Sweden
Stockholm - February 18-20, 1991

The Consumer Expenditure Survey (CE) has been a continuing survey since 1980, including two components, Diary and Interview, with separate samples of households. The Diary is used to collect all expenses for two consecutive one-week periods, with special emphasis given to detailed frequent expenses, such as food expenses. The Interview has five panel waves to the same respondents and collects all expenses, especially sizable expenses. Data collected in the first wave, with one-month recall, are primarily used for bounding purposes, and are excluded from expenditure estimates. A three-month recall is used for most expenses in the other waves. Data from both components contribute to the estimation of annual expenditure means and the development of expenditure weights for the Consumer Price Index.

Aside from cost considerations, cooperation levels, respondent burden, and data quality are critical issues in surveys. The length of the reference period, the questionnaire length, and the number of reporting intervals are factors affecting all three aspects. Diary and interview methods present different challenges for interviewers and respondents, but underreporting is a common effect. Many types of consumer expenditure surveys have resulted in an attempt to balance this problem vis a vis the ever-increasing demands for data. While nonresponse has been kept within reasonable levels in the United States (15% for both components), comparisons with independent sources of expenditure estimates suggest underreporting levels of 40 percent or greater for certain commodities of the CE. For this reason, the development of improved methodologies has become very important.

Recent research findings in the areas of recall bias and conditioning effects in the Interview, first-day bias and incomplete diarykeeping in the Diary, and comparisons with independent sources are included in section 1. Selected issues with regard to a redesign are discussed in section 2. Alternative collection methods have been suggested by BLS and other researchers to address these issues, and are briefly presented in section 3. They are meant to serve only as examples, with no analysis of the potential benefits or drawbacks. The workshop should serve as a source of discussion on the following topics:
1) experience in countries where diary and interview methods are used sequentially or concurrently for the same household,
2) other collection methods that may be feasible for national programs,
3) criteria that should be used to match the collection method(s) with the type of expense, and
4) ideas for evaluating alternative data collection methods.

1. RESEARCH FINDINGS

1.1 Repeated-Interview, Recall, and Telescoping Effects

Panel surveys collecting retrospective data exhibit conditioning and recall effects. These effects are usually in the form of reporting declines from one wave to the next and for events further back in the past. Differences in expenditure means by wave have been tested systematically, based on replies by respondents to the second through fifth waves in the years 1982/83 (n = 6,600 per wave, 74% of all respondents). Judging from these aggregate comparisons, it appears that

time-in-sample effects are moderate to small, whereas recall effects are large and widespread across commodities. Table 1 displays ratios to overall means for three commodities. Compared to the second wave, the fifth wave is 18% lower for home furnishings and 6 percent lower for apparel, when all recall months are combined (1 - col.5 / col.2). These declines are significant, but the decline for home maintenance is not significant. The declines from the first recall month (the most recent) to the third recall month (the most distant) are significant for each commodity. For waves 2 to 5 combined, the third recall month is 24 percent lower than the first recall month for home furnishings and 39 percent lower for apparel.

Further data analysis has compared the first wave to subsequent waves, in order to assess the extent of telescoping effects. These tend to inflate the reports due to the inclusion of expenses outside the reference period. The analysis was based on replies by respondents in all five waves for the year 1984 (n = 3,200 per wave, 70% of all respondents). Selected findings are displayed in Table 2. The overstatement due to unbounded data collection is estimated to be substantial. At the commodity level of aggregation, estimates are 18 percent for apparel and 37 percent for home furnishings, under certain assumptions (1 - col.2 / col..1). These results underscore the need to use bounding methods. The first wave, with a short recall period of one month, appears to provide higher reporting after the exclusion of telescoping effects. Mean expenditure estimates of one-month recall, net of telescoping, are 16 percent greater than comparable estimates for the three-month recall, for apparel (col.2 / col.3 - 1). Noteworthy, potential gains from a short recall period tend to increase for smaller expenses, but may become marginal for big-ticket items.

1.2 Diary Performance

Declines in reporting by diary day and week are well known, and are found again in the current CE. This analysis was based on diaries completed solely by respondents for the two weeks in the year 1987 (n = 3,935 per week, 66% of the respondents). On average, the first day of the first week exhibits an expenditure mean 35 percent greater than the overall mean. Mean expenditure estimates for the second week are 11 percent lower than first-week estimates. These effects are in remarkable agreement with findings from the 1972/73 CE (Pearl, 1979).

The procedure followed in the United States allows diaries to be completed (partially or totally) by recall at diary pickup. In most years, 10 percent of the diaries have partial recall, and 15 percent have total recall. These diaries exhibit lower expenditure levels than diaries completed by respondents, indicating that these respondents either have fewer expenses or report less expenses. These diaries also lack specificity in the reported data more often than diaries completed by respondents, and this seems to be a logical outcome of the procedure.

The Diary sample design allows continuous data collection throughout the year, within a placement schedule. Seasonal patterns from the two components show some dissimilarities for selected commodities. The month of December has relatively lower reporting in the Diary compared to the Interview. Ratios of monthly means to the overall monthly average for the year 1987 are shown in Table 3. While there are many differences in the two methods of data collection, a Diary-specific problem may be the potential disclosure of gift items for family members.

1.3 Expenditure Estimates

Certain commodities are believed to be underreported in both CE components, judging from comparisons with independent sources, such as the personal consumption sector of the National Accounts (NA). Ratios of aggregate CE to PCE estimates for the years 1984 to 1987 are displayed for selected commodities in Table 4. The methodology used in the estimation system

selects data from both sources at the detailed level according to the estimate with the lower mean square error (MSE). (In addition to variances from the two components, NA estimates are used to develop the MSE.) In some instances, 50 percent of the categories within a commodity are derived from each component. The current source selection for apparel categories is shown in Table 5.

The source selection is not always consistent with a priori expectations, although specialized comparisons of means and distributions give, in general, more expected results, confirming that the Diary captures smaller expenses to a greater degree, whereas the Interview yields stronger data for less frequent or more salient expenses (Silberstein and Scott, 1990). Despite possible improvements in the source selection methodology, a number of problems would remain. The estimates so derived are still quite low, and the selection of one or the other source does not translate into any reduction in respondent burden.

2. ISSUES IN A REDESIGN

Reduction in respondent burden

If the questionnaire is excessively long, it must be conducted at a very fast pace in order to complete it within a reasonable time. (On average, it takes two hours to complete the interview.) Respondents should be given fewer questions to reduce fatigue associated with the interview and have more time to think about the answers or consult records. Examples may be:
- Reduce the number of times certain sections are asked in the panel, thus reducing the number of sections in a given interview
- Reduce the amount of detail for certain sections
- Streamline the interviewing technique

Improvements in expenditure estimates

Recall problems are severe for certain types of expenses. A three-month recall may be unrealistic for frequent purchases, especially for households with numerous members and expenses. Other factors influencing underreporting are proxy reporting and an interview style that tends to induce negative answers. Alternatives may be:
- Shorten recall for certain expenditures
- Have respondents keep a diary during the intervening months of the panel
- Transfer certain sections to a Diary component for nonfood expenses
- Provide personal diaries

Constraints due to data needs

Besides providing sufficient sample size for expenditure estimates and CPI weights, the design should allow extended research. Two issues are briefly noted. Complete coverage of expenses within the Interview is desired, stemming from the idea that meaningful economic research can only be derived from a complete picture of yearly spending at the microlevel. This data need is a major obstacle to a redesign that would limit the data to be collected from each respondent. Collecting some of the data as summary (global or usual) expenses can be introduced as an alternative, although there are problems associated with this method.

Statistical linkage between survey components is also desired for research. This tends to increase the size of the questionnaire. The list of variables that may be useful can be quite extensive and,

at a minimum, should include demographic, housing, and income characteristics, and data on the estimated market (and rental) value of owned homes.

3. ALTERNATIVE METHODS

Many specific data collection techniques have been tested in several countries, and there is a considerable body of literature on the subject, dating back at least 40 years. Each country has developed methods believed to be acceptable, if not optimal. More information is needed on the success of current methods. The criteria for defining what is "successful" should be developed.

There seems to be agreement at BLS that the current Interview questionnaire is too long, although it is still expanding due to increased data demands and more precise wording of questions. One approach to develop alternative collection methods would be to restructure the CE into a modular-type survey, where specific techniques would be addressing different collection problems, with the overall goal to reduce respondent burden. Another approach would limit the redesign effort towards improving areas of expenses with the greatest problems in the current collection methods.

Selected changes to the three-month recall may be needed for improving the quality of responses. Shortening the recall period for some expenditures would reduce respondent burden and improve response quality, especially for frequently-made expenses. This technique can also be applied in the form of limiting the collection of repetitive expenses to the most current one (e.g., the last utility bill). As for apparel and other expenses that are difficult to recall, diary and interview methods may have to be used before substantial improvements can be achieved. Switching to one-month recall may not be sufficient to overcome underreporting. A special diary could be designed for these expenses, since the diary method with emphasis on food items discourages reporting of other expenses. The use of personal diaries should be considered to curb underreporting which results from proxy reporting and disclosure problems in the case of gifts.

Lengthening the recall period for certain expenditures can reduce respondent burden, especially for large infrequent expenses. The rationale for this technique is that respondent burden may be less when asking about large purchases, for instance, one time every six months rather than two times every three months. If these questions are asked only once during the panel, telescoping effects may be sizable. For large expenses, however, fewer telescoping errors are made and more efficient estimates can be derived with a longer recall.

Examples of possible techniques to restructure the CE are outlined within three types of methods: 1) modular interview, 2) core and supplement, and 3) designing a special component for selected commodities to be excluded from the current components.

1) Modular panel interviews

This method reduces the panel of five waves with all sections to five waves with a varying number of sections. Some sections may be transferred to a second Diary. Examples may be as follows:

- Rotated sections by wave: all sections would be collected from each reporter with three-month recall, but only twice, first time for bounding
- Split questionnaire: after a bounding first wave, collect Part A in the 2nd and 4th waves, Part B in the 3rd and 5th waves (4th and 5th waves with six-month recall).

2) Panel interviews with core and supplement(s)

One version of this method would collect all expenses in a core questionnaire, although not all at the same level of detail. The supplements would then be used to obtain detailed information on specific sets of expenses. Another version may select a number of sections to be collected in each wave (core), and a number of sections to be collected only as supplements. In this case, only a small part of the questionnaire would be selected as a core, and most of the questionnaire would be selected as supplement, although infrequent expenses may have to be included in the core to insure sufficient sample. A set of selection criteria would have to be decided.

Possible features:
- Variable number of waves by section
- One-time supplements for certain sections, two-time supplements for others
- Variable recall length by section and/or wave.

3) Design a third component

The three components would include one Panel Interview and two Diaries, with limited overlap. One option would be to design a Food Diary and a Nonfood Diary. Alternatively, the second Diary would be devoted primarily to sections known to be underreported, rather than to all nonfood items. These sections would be eliminated from the two existing components. An added feature of the second Diary could be to include a follow-up interview one month after the end of the diary in order to capture additional (large) expenses.

NOTE:
Three research methods are used at BLS: 1) statistical analysis of data to obtain estimates of response error and other effects, 2) field experiments to test the effectiveness of new techniques, and 3) cognitive laboratory studies to improve collection procedures. Of these, only selected findings from statistical analysis are discussed here. Research on respondent characteristics and their potential association with measurement error and nonresponse is reported in several papers. (See, for instance, Garner and Blanciforti, 1987; Silberstein, 1989; Tucker, 1990.) Results from field tests and cognitive laboratory studies are included in papers by E. Jacobs et al. and by C. Dippo, both presented at the 1989 ISI, and are discussed in detail in numerous papers (Miller and Downes-Le Guin, 1989; Tucker, 1986; Tucker and Bennett, 1988; Tucker et al., 1989).

REFERENCES

Dippo, C.S., "The Use of Cognitive Laboratory Techniques for Investigating Memory Retrieval Errors in Retrospective Surveys," Proceedings of the 47th Session of the International Statistical Institute, Paris, 1989, VIII (2), pp. 363-381.

Garner, T.I. and Blanciforti, L.A., "Reporting of Household Income: Complete versus Incomplete Response," Proceedings of the Third Annual Research Conference, Bureau of the Census, Washington, D.C., 1987, pp. 644-661.

Jacobs, E., Jacobs, C., and Dippo, C., "The U.S. Consumer Expenditure Survey," Proceedings of the 47th Session of the International Statistical Institute, Paris, 1989, VIII (2), pp. 123-142.

Miller, L., and Downes-Le Guin T., "Improving Comprehension and Recall in the Consumer Expenditure Interview Survey," Proceedings of the Section on Survey Research Methods, American Statistical Association, 1989, pp. 502-507.

Pearl, R.B., "Reevaluation of the 1972-73 U.S. Consumer Expenditure Survey," Census Bureau Technical Papers, No. 46, U.S. Government Printing Office, Washington, DC, July 1979.

Silberstein, A.R., and Jacobs, C.A., "Symptoms of Repeated Interview Effects in the Consumer Expenditure Interview Survey," in: Panel Surveys, (Eds. D.Kasprzyk, G.Duncan, G.Kalton, and M.P.Singh.), New York: Wiley, 1989, pp. 289-303.

Silberstein, A.R., "Recall Effects in the U.S. Consumer Expenditure Interview Survey," Journal of Official Statistics, Vol. 5, No. 2, 1989, pp. 125-142.

Silberstein, A.R., "First Wave Effects in the U.S. Consumer Expenditure Interview Survey," Survey Methodology, a Journal of Statistics Canada, 2, 1990, IN PRESS.

Silberstein, A.R., and Scott, S., "Expenditure Diary Surveys and Their Associated Errors," in: Measurement Errors in Surveys, New York: Wiley, 1991, IN PRESS.

Tucker, C., "An Analysis of the Dynamic in the CE Diary Survey," Proceedings of the Section on Social Statistics, American Statistical Association, 1986, pp. 18-27.

Tucker, C., and Bennett, C., "Procedural Effects in the Collection of Consumer Expenditure Information," Proceedings of the Section on Survey Research Methods, American Statistical Association, 1988, pp. 256-261.

Tucker, C., Vitrano, F., Miller, L., and Doddy, J., "Cognitive Issues and Research on the Consumer Expenditure Diary Survey," Paper presented at the annual conference of the American Association for Public Opinion Research, 1989.

Tucker, C., "The Estimation of Instrument Effects on Data Quality in the Consumer Expenditure Diary Survey," Paper presented at the International Conference on Measurement Error in Surveys, Tucson, AZ, November 11-14, 1990.

U.S. Bureau of Labor Statistics (BLS), "Consumer Expenditure Survey, 1987," BLS Bulletin, No. 2354, U.S. Government Printing Office, Washington, D.C., June 1990.

Table 1. Time-in-sample and Recall Effects: 1982/83 Interview
(Percent ratios of expenditure means by wave and recall month to overall expenditure mean)

	Recall Month	Wave				
		2 to 5	2	3	4	5
		(1)	(2)	(3)	(4)	(5)
Home maintenance	All	100	107	99	93	101
	1	112	119	112	109	108
	2	109	118	104	95	118
	3	79	85	80	77	76
Home furnishings	All	100	112	98	98	92
	1	116	122	115	111	115
	2	96	110	87	100	85
	3	88	103	92	81	77
Apparel a/	All	100	103	101	98	97
	1	124	126	127	123	120
	2	100	105	99	98	100
	3	76	80	78	73	73

a/ Includes apparel services.
Source: Silberstein & Jacobs (1989), pp.297, 298, 301.

Table 2. Comparisons of First Wave and Subsequent Waves: 1984 Interview
(Annual means in U.S. Dollars and, in parenthesis, standard errors)

	Wave 1		Waves 2 to 5	
	As reported	Net of telescoping a/	All months	1st recall month
	(1)	(2)	(3)	(4)
Home furnishings	1,972	1,235	1,179	1,327
	(85.0)	(n.a.)	(59.7)	(73.1)
Apparel b/	1,663	1,370	1,182	1,452
	(59.6)	(n.a.)	(61.7)	(71.0)

a/ Assuming conditioning.
b/ Includes apparel services.
Source: Silberstein (1990).

Table 3. Comparisons of Monthly Expenditures: 1987 Diary and Interview
(Percent ratios of expenditure means by month to overall monthly average mean expenditure)

			Month			
		All Months		OCT	NOV	DEC
		(1)		(2)	(3)	(4)
Small home furnishings	Diary	100	. .	91	143	154
	Interview	100	. .	85	99	184
Apparel	Diary	100	. .	94	156	162
	Interview	100	. .	86	99	251
Entertainment	Diary	100	. .	97	126	150
	Interview	100	. .	104	103	177

Source: Silberstein & Scott (1990).

Table 4. Comparisons of Aggregate Expenditures between CE and PCE: 1984 to 1987
(Ratios of CE to PCE estimates derived from the personal consumption sector of the National Accounts)

	1984	1985	1986	1987
	(1)	(2)	(3)	(4)
Food	0.72	0.74	0.71	0.71
Home furnishings	0.82	0.79	0.82	0.78
Apparel a/	0.66	0.68	0.62	0.63

a/ Includes apparel services.
Source: Bureau of Labor Statistics (1990), p. 6.

Table 5. Source Selection for Apparel Integrated Estimates: 1984 to present
(I - Interview, D - Diary)

	Men	Women	Boys	Girls
	(1)	(2)	(3)	(4)
Coats and jackets	D	D	D	I
Sportcoats	I	D	D	-
Suits	I	I	D	I
Dresses	-	D	-	I
Pants and shorts	D	D	D	I
Skirts	-	D	-	I
Sweaters	I	D	I	D
Shirts and blouses	D	D	D	D
Sportswear	I	D	I	D
Accessories	D	D	D	D
Uniforms	I	D	I	I
Nightwear	D	D	D	I
Underwear	D	D	D	I
Hosiery	D	I	D	D
Other clothing	I	D	I	I

Source: Bureau of Labor Statistics (1990), p. 139.