Journal of Official Statistics
Vol. 5. No. 2, 1989, pp. 125–142

Recall Effects in the U.S. Consumer Expenditure Interview Survey[1]

Adriana R. Silberstein[2]

Abstract: The three month recall of apparel and house furnishing expenses from the 1984 U.S. Consumer Expenditure Interview Survey is analyzed by type of respondent. Main effects and interactions are tested for such variables as family type, age of respondent, and use of records, with a dependent variable derived by comparing the expenses reported for the most recent month to expenses reported for the three month period; the variable is interpreted as an indicator of potential underreporting. Results show that significant variation can be expected in the degree this phenomenon manifests itself with different respondents.

Key words: Response error; expenditure data; type of respondent.

1. Introduction

The U.S. Consumer Expenditure Interview Survey is an ongoing panel survey collecting expenditure data from a national sample of households interviewed five times at three month intervals (waves). The interview covers all household expenditures in great detail, for the most part. Only summary questions are asked for frequently purchased items, such as food items, which are collected in detail by a diary method with a separate sample of households. Panel effects due to repeated interviews have been found to be small, but recall effects are evident. These effects were analyzed at the macro-level in a study by Silberstein and Jacobs (1989).

This article discusses differential effects in quality of responses. The study analyzes data reported in 1984 for two expense categories, apparel and house furnishings, using log-linear models. The task of reporting this type of expenses, while easily understood by the respondents, is often difficult for a three month recall period. The general conclusions from the study are that recall effects tend to be consistent from one wave to the next and are likely to be heightened by certain interview and respondent characteristics.

Section 2 describes relevant aspects of the response process and major factors affecting quality. In Section 3, the model methodology and the variables used are discussed. Section 4 reports on the findings, and Section 5 concludes.

[1] An earlier version of this article was presented at the Annual Meeting of the American Statistical Association, San Francisco, 1987.

[2] Mathematical Statistician, Office of Prices and Living Conditions, Statistical Methods Division, U.S. Bureau of Labor Statistics, Washington, D.C. 20212, U.S.A.

Acknowledgements: Thanks are extended to Clyde Tucker, Stuart Scott, the referees, and David Mingay for their helpful comments.

2. Response Process

Recall bias results from the selective recollection of past events: less recent and less salient events tend to be reported less completely than more recent or more salient events. The error is generally greater for longer reporting periods, and is usually compounded by "forward telescoping," i.e., a tendency to postdate events when reporting them. Certain aspects of the interview, e.g., respondent fatigue and level of interest, are widely recognized as primary factors influencing the retrieval and reporting of retrospective information. (See Tanur (1981), Turner and Martin (1984), Moss and Goldstein (1979).)

Figure 1 illustrates in general terms various elements of the response process. The respondent's attitudes and cooperation are affected by the type of survey and the interview procedures; the survey content and complexity influence the respondent's motivation and ability to report accurately. The interaction between these factors varies by respondent, family, and interview characteristics, thus differences in data quality may result.

Major factors in an expenditure interview survey are the respondent's knowledge of the family's expenses, the length of the recall period, and the expense records available during the interview. The respondents to the Interview Survey answer a single questionnaire covering all pertinent members. There is no set procedure for the choice of respondents; more than one person may participate, depending on cooperation.

The unit of analysis is a consumer unit (CU), which includes those members of a household who are either related or pool their income to make joint expenditure decisions (U.S. Bureau of Labor Statistics (1986)). More than one CU may be identified within a sampled household, and report separately in the panel; new CUs may move into the sampled household after the beginning of the panel.

The first wave, with one month recall, is a

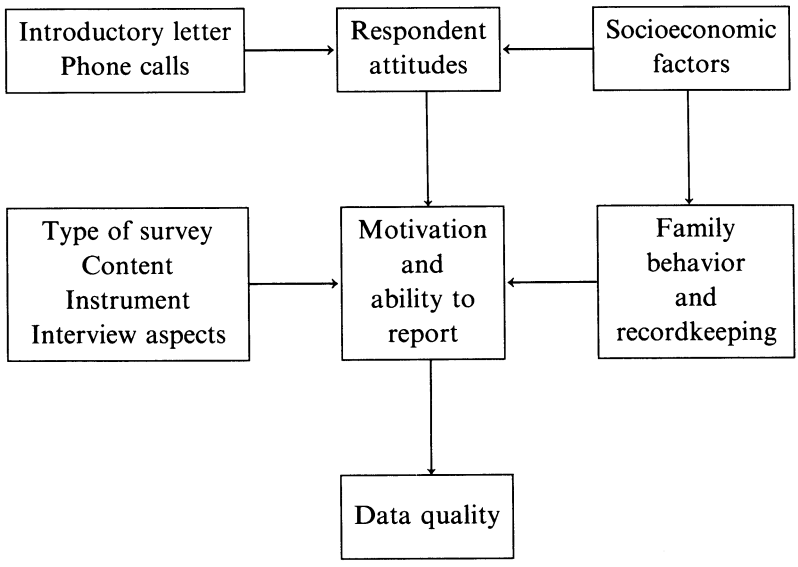

Fig. 1. Response process

Table 1. Percent CUs with no expenses in category by recall month

	Apparel		House furnishings	
Qtr:	842	843	842	843
CUs with expenses in three month period	3461	3408	2745	2762
Recall month:	%	%	%	%
FIRST	16	17	28	29
SECOND	33	31	45	40
THIRD	41	39	48	45

bounding interview. Expenses reported for the previous month of the first wave are not used in the estimates since they are thought to include expenses incurred in prior months (as a result of forward telescoping). The second through fifth waves have three month recall periods. The reporting period includes the portion of the current month up to the interview day; expenses reported for that month are made available to the interviewer in the following wave, to minimize duplicate reporting.

3. Methodology

3.1. The dependent variable

The study is based on respondents in two consecutive waves (90% of all respondents); CUs that do not participate in a previous wave experience an unbounded interview. All data pertain to the urban population.

Study cases are reporters in waves two to five with at least one expense in the categories analyzed: 3408 for apparel and 2762 for house furnishings in the third quarter of 1984 (Qtr 843). This quarter, covering expenses made between April and August, excludes major seasonal shopping peaks which would affect comparisons by month. Table 1 shows how reports with no expenses are less frequent in the first (most recent) month compared to the other two months.

The models' dependent variable is a categorical variable derived by grouping CUs according to the expenses (the dollar value) reported for the first recall month as percentage of the total expenses reported for the three recall months. These *recall effects* are subdivided into three levels: NONE – 0 to 35%, MODERATE – 35% to 75%, and GREAT – 75% to 100%. The distributions and the mean expenditures are given in Table 2. The first level indicates that no recall effects could be observed, since the first recall month had either no reported

Table 2. Dependent variable: Recall effects (Qtr 843)

Apparel	% CUs	Monthly mean expenditure	House furnishings	% CUs	Monthly mean expenditure
Recall effects:	n = 3408		Recall effects:	n = 2762	
Total	%	U.S.$102	Total	%	U.S.$141
NONE	44	106	NONE	53	146
MODERATE	29	125	MODERATE	18	163
GREAT	27	71	GREAT	29	119

Table 3. Number of items reported: Apparel (Qtr 843)

Number of records[1]	Number of CUs	Distribution by recall effects				1st month ratio[2]
		%	None	Moderate	Great	
ALL	3408	100	45	29	27	128
1	333	100	53	–	47	130
2 to 4	994	100	44	21	35	141
5 to 9	876	100	46	30	24	134
10 to 19	806	100	41	39	20	125
20+	399	100	41	49	10	124

[1] One record may include multiple purchases when the same type of item is reported for the same person and month.
[2] Percent ratio of first month expenses to three-month average.

expenses or comprised up to about a third of the reported expenses. A fourth level, EXTREME, was included in some models: it separates CUs that reported expenses only for the first recall month (half of the CUs in the third level).

The variable is referred to as recall effects, rather than recency effects, as otherwise called (Biderman and Linch (1981)), to emphasize the complexity of the recall task. While it is difficult to interpret the recall variable in terms of quality, it seems reasonable to assume that greater recall effects are symptomatic of greater underreporting. Greater effects were found for respondents (in bounded interviews) that did not report for the whole length of the panel, compared to reporters in all five waves.

Two reporting aspects appear to influence the dependent variable directly: the number of expenses and the interview week. The fewer items purchased the easier recalling should be. However, it is apparent that recall effects tend to decrease with increasing number of items reported (Table 3). A more even distribution by recall month results when more expenses are reported, partly due to the fact that more than one month is often involved. The data in this table refer to the number of "records." A record may include multiple purchases when the same type of item is reported for the same person and month. A categorical variable derived by using this variable, rather than the dollar value, produced consistent results in the analysis.

Table 4. Interview week of previous wave: Apparel (Qtr 843)

Interview week of previous wave	Number of CUs	Total		3rd recall month	
		Monthly mean expenditure	1st month ratio[1]	Monthly mean expenditure	Collected in previous wave
Total	3408	U.S.$102	128	U.S.$78	39%
FIRST	1418	90	132	60	17%
SECOND	1154	108	135	82	36%
OTHER	836	114	115	104	64%

[1] Percent ratio of first month expenses to three-month average.

Expenses reported for the "current month" in one wave refer to an overlap period with the third recall month of the next wave, which is usually not well reported. Interviews conducted late in the month capture many of these expenses and this tends to produce higher reports in the next wave; as Table 4 shows, these interviews (OTHER) have higher means and lower first month "bias" ratio. The lowest bias and the highest mean expenditure were observed when the previous interview was in the second half of the month and the current interview was in the second week of the month; this is also one of the shortest reporting periods, not surprisingly.

3.2. *Independent variables*

The topic of respondent effects in survey methodology involves many issues: differential understanding of the response task and varying degree of motivation in answering survey questions are just two examples. The study of these effects tends to produce results bound to a specific survey setting. Sudman and Bradburn (1977) point out that age of respondent might be expected to be the major characteristic related to memory error variation, but several interview conditions tend to reduce these effects. Cannel, Marquis, and Laurent (1977) reported that forgetting curves were similar for both male and female respondents in hospitalization surveys that were validated by record checks; self-reports appeared less affected by recall bias than reports given by another family member, while age and education were found to yield less conclusive results.

The direction of the relationship of a specific variable to reporting quality may reflect the survey subject matter and administration. Response from young people (college age) as well as from the elderly (65 and over) appear to be of lower quality in the U.S. Diary Survey (Tucker (1986)). Attempts have been made to provide generalized results from several surveys. Andrews and Herzog (1986) concluded that older respondents tend to give less precise answers, judging from a study of six opinion surveys that included methodological supplements.

Certain variables, especially the age distribution, are sensitive to exclusions due to the scope of the study. Weighted percent distributions of the major variables are given in Table 5. About 40% of respondents less than 25 years old did not participate in the previous wave, due mainly to mobility; this group decreased from 10% to 7% of the total, as a result. Older respondents (65 years old and over) presented a different situation: only 6% did not participate in a previous wave, but 26% did not report any apparel expenses (compared to 13%, on average).

The number and type of respondents, and record usage are variables pertaining to interview aspects of the recall process. Record usage refers to use of records during the interview rather than to a specific questionnaire section. It can be assumed that if no records are used, the recall process for all sections is done without records. The use of records eliminates telescoping effects and tends to decrease omissions (Sudman and Bradburn (1977)).

3.3. *Models*

Log-linear models were fitted to the cross-classified, categorical data using the computer program "Contingency Table Analysis for Complex Sample Designs (CPLX)" developed by Fay (1987); the program computes maximum likelihood estimates of the parameters and their standard errors. Many

Table 5. Main respondent and consumer unit characteristics (Qtr 843). Percent

Apparel n = 3408					
Age	100	CU size:	100	Respondent type:	100
Under 30	19	One person	22	(Husb–wife CUs only)	
30–39	25	Two persons	29	Both wife and husband	36
40–64	40	Three or four persons	35	Wife	48
65 and over	16	Five or more persons	14	Husband	16
Education:	100	Type of CU:	100	Number of respondents:	100
Less than high school	19	Husb–wife with others	41	(CUs with 2 or more persons)	
High school	35	Single or husb–wife alone	42	Only one	67
More than high school	46	Other	17	More than one	33
Sex:	100	CU income: U.S.$	100	Record usage:	100
Male	33	< 15 000	33	Some records used	77
Female	67	15 000–30 000	28	No records used	23
		30 000 and over	32		
		Reported incompletely	7		

sources are available describing log-linear models (Bishop, Fienberg, and Holland (1975), Goodman (1978) and Fienberg (1980)). "CPLX interprets these models as statements about the proportions in the underlying population from which the sample is drawn" (Fay (1987), 1.1), and is designed for analysis of data derived from complex sample surveys. Weighted frequencies were derived with the method of balanced repeated replication. This method takes into account design effects, and is one of the options in CPLX; the 1984 Interview Survey had 20 half-sample replicates.

Following an analysis of the individual relationships of the independent variables to recall effects (test of independence), models were developed using sets of variables found to be significantly related to recall effects, adding interaction terms up until a good fit was obtained. The overall fit of a model and the contribution made by additional parameters were tested through the jackknifed chi-square tests developed by Fay (1985). The three and four variable models shown below were used in the analysis

$$\log(F_{ijh}) = \lambda + \lambda_i^I + \lambda_j^J + \lambda_h^H + \lambda_{jh}^{JH} + \lambda_{ij}^{IJ} + \lambda_{ih}^{IH},$$

and

$$\log(F_{ijhk}) = \lambda + \lambda_i^I + \lambda_j^J + \lambda_h^H + \lambda_k^K + \lambda_{jh}^{JH} + \lambda_{jk}^{JK} + \lambda_{hk}^{HK} + \lambda_{jhk}^{JHK} + \lambda_{ij}^{IJ} + \lambda_{ih}^{IH} + \lambda_{ik}^{IK},$$

where i is the dependent variable, j, h, and k are independent variables.

The models measure the type of interactions of each independent variable with the dependent variable, after eliminating the effects of the interactions between the independent variables. No more than three independent variables were included in any model, due to sample size limitations; zero frequencies were not allowed in the complete cross-classification, although this is not a requirement in CPLX.

Most of the findings were derived from models fitted to apparel data since a larger sample was available for that category. Test results are displayed in Tables 6 to 13; with the exception of Table 6, which pertains to recall effects in two consecutive quarters, the tables refer to Qtr 843. Standardized values for the significant interaction terms involving the dependent variable are shown in the tables; these values are the ratios of parameter estimates to their estimated standard errors.

Some variables were defined in a dichotomous way, and the standardized values are only shown for one of the levels; the level not shown has the opposite sign. Standardized values $|z| = 1.96$ or larger can be considered significant at the .05 level. Critical values (.05 level) for the jackknifed chi-square test are: 1.5 for DF = 2 to 9, and 1.6 for DF = 10 to 20.

4. Test Results

4.1. Previous and current waves

Recall effects for the same respondents in two consecutive waves were compared; data for the second wave were excluded in this test, since the previous wave for those reports was the first wave. The results, shown in Table 6, indicate that consistent recall effects can be expected from one wave to the next, especially for respondents with moderate recall effects ($z = 4.0$). CUs that do not report apparel expenses in the previous wave exhibit greater recall effects in the current wave.

The relationship of week of previous and current interviews to recall effects is significant, as expected. The interview weeks were

Table 6. Apparel recall effects in two consecutive waves

Variables: Recall effects in Qtr 843 (dependent), recall effects in Qtr 842, week of previous and current interview.
Note: Second wave in Qtr 843 excluded.

	Apparel recall effects in Qtr 843		
	None	Moderate	Great
	Standardized values		
Interaction of apparel recall effects in two waves			
Qtr 842: No apparel reported	1.49	−3.41	3.28
No recall effects	0.00	1.68	−1.90
Moderate	0.03	4.00	−4.33
Great	−2.33	0.37	1.89
Interaction of recall effects and previous interview week			
Second half of month	2.33	−0.41	−2.13
Interaction of recall effects and current interview week			
Second half of month	−2.89	0.80	1.87
	Test values	DF	
Jackknifed chi-square test			
Previous interview week	1.8	2	
Current interview week	2.3	2	
Recall effects in Qtr 842	4.8	6	

Table 7. Apparel and house furnishings recall effects

Variables: Apparel effects (dependent), house furnishings effects, week of previous and current interview. (Qtr 843)

	Apparel recall effects		
	None	Moderate	Great
	Standardized values		
Interaction of apparel house furnishings			
No house furnishings reported	1.30	−5.02	5.96
House furnishings recall effects: None	1.59	2.09	−3.34
Moderate	−0.31	1.85	−2.07
Great	−1.97	−0.04	1.87
Jackknifed chi-square test	Test values	DF	
Previous interview week	2.2	2	
Current interview week	2.8	2	
House furnishings recall effects	5.8	6	

grouped in two levels, first half and second half of the month (values shown for the second). CUs interviewed later in the month in the previous wave exhibit no recall effects in the current wave ($z = 2.33$). By contrast, great recall effects are exhibited when the previous interview was in the first half of the month. Opposite results can be seen for the interview week of the current wave.

4.2. *Apparel and house furnishings effects*

The relationship of recall effects in two parts of the questionnaire, apparel (Section 9) and house furnishings (Sections 6 to 8), was tested in a model that included previous and current week of interview, as controls. (See Table 7.)

It was expected that recall effects would be less for house furnishings than for apparel, since that category comprises several salient large expenditure items. The interaction between the two categories shows a consistent direction: moderate apparel recall effects are exhibited when house furnishings are reported with moderate or no recall effects. When no house furnishing expenses are reported, greater recall effects for apparel are observed.

4.3. *Respondent characteristics*

Significant relationships with recall effects were found for age and education of respondent, but not for sex or race of respondent. Respondents between the ages of 25 and 44 exhibit moderate effects, whereas older respondents (65 and over) exhibit either no effects or extreme effects. No significant relationships were noted for other age groups (less than 25, 45 to 54, and 55 to 64). Respondents with less than high school education are the most likely to exhibit great recall effects, respondents with more than high school education are the least likely to exhibit these effects.

Age and education were tested simultaneously, and their effects were both significant, as shown in Table 8. The interaction of age and education indicates that less than high school education is highly correlated with older respondents, as reported in other studies. This relationship is reflected in the results: no recall effects are evident for older respondents when the model controls for education. The noted results for education, instead, are retained in this model.

The respondent's family setting is important in relationship to recall effects; size and type of the consumer unit are especially noteworthy. The findings suggest what seems logical: it is easier to report when fewer family members are involved, and more difficult when there are more members. No effects are exhibited by CUs with one or two members, moderate effects are exhibited by CUs with three or four members, and moderate to great effects can be observed for larger families (Table 9).

Moderate effects are likely in reports by husband–wife families with children or other members. No recall effects are more likely for single–person and husband–wife CUs without any other members. These relationships are maintained when the expenditure level and interview week are included as control variables (Table 10).

Other CU characteristics are associated to recall effects. The relationship to home tenure suggests that renters exhibit greater effects than home owners do. Recall effects appear to differ by family income. As Table 11 shows, greater recall effects are experienced by CUs with lower incomes. CUs with incomplete income reporting are also associated with greater effects; these reporters refuse to disclose major sources of

Table 8. Age and education of main respondent

Variables: Recall effects (dependent), age and education of main respondent.

	Apparel recall effects		
	None	Moderate	Great
	Standardized values		
Interaction of recall effects and age of main respondent			
Less than 40 years old	−4.05	3.46	−1.37
40 to 64 years old	−2.43	0.53	1.51
65 years old and over	4.19	−2.68	0.01
Interaction of recall effects and education of main respondent			
Less than high school	−1.32	−1.62	2.99
High school	0.49	0.62	−1.39
More than high school	1.04	1.45	−3.49
Jackknifed chi-square test	Test values	DF	
Age of main respondent	4.2	4	
Education of main respondent	1.8	4	

Table 9. Consumer unit size

Variables: Recall effects (dependent), CU size and record usage during the interview.
Note: Restricted to CUs that reported two or more apparel records.

	Apparel recall effects			
	None	Moderate	Great	Extreme
	Standardized values			
Interaction of recall effects and CU size				
One person	3.50	−3.45	−2.14	1.63
Two persons	2.29	−2.07	0.07	0.16
Three or four persons	−1.27	3.02	0.77	−1.95
Five or more persons	−3.52	2.40	2.46	−1.14
Interaction of recall effects and record usage				
Some records used	1.37	6.47	−1.02	−4.99
Jackknifed chi-square test	Test values	DF		
CU size	5.0	9		
Record usage	7.2	3		

Table 10. Type of consumer unit

Variables: Recall effects (dependent), type of CU, week of previous interview and expenditure group.

	Apparel recall effects		
	None	Moderate	Great
	Standardized values		
Interaction of recall effects and type of CU			
Husband–wife with children or others	−2.58	2.13	0.28
Single or husband–wife without others	2.45	−1.33	1.42
Other CUs	−0.18	−0.89	1.13
Interaction of recall effects and expenditure group			
Five or more apparel records	−0.74	12.47	−11.19
Jackknifed chi-square test	Test values	DF	
Previous interview week	1.8	2	
Type of CU	2.2	4	
Expenditure group	20.6	2	

Table 11. Consumer unit income

Variables: Recall effects in Qtr 843 and income.

	Apparel recall effects		
	None	Moderate	Great
	Standardized values		
CU income before taxes			
Less than U.S.$15 000	1.43	−4.97	2.81
U.S.$15 000–30 000	0.34	0.99	−1.65
U.S.$30 000 or more	−0.88	5.10	−4.18
Reported incompletely	−0.46	−1.71	1.97
Jackknifed chi-square test	Test value	DF	
CU income before taxes	6.1	6	

income, and this is often considered an indication of lower cooperation to the survey.

4.4. *Interview aspects*

Specific subsets of respondents were selected to facilitate the analysis of variables dealing with interview aspects. Husband–wife consumer units were analyzed in terms of who participates as respondent, whether the husband, the wife, or both; (the few cases that did not have either spouse as the main respondent were excluded). This model takes into account the respondent's age, collapsed in two levels: less than 40 years old, and 40 and over (values shown for this level). The presence or absence of other members in the CU is another variable in the model. (See Table 12.)

The effect of CUs with old respondents and scant reports, which accentuates the extremes of the recall variable, is reduced in this model, due to the specific set of families selected. A significant relationship shows that respondents in the 40 and over age group tend to exhibit greater recall effects than respondents in the younger age group.

Husband–wife CUs exhibit moderate recall effects if both husband and wife participate, and no effects if only the husband does. The interactions of the independent variables (values not shown) indicate the wife alone (more often in the younger age group) is more likely to be the respondent for CUs with children or others, whereas both spouses are more likely to answer in husband–wife CUs with no other members.

The number of respondents and the use of records were tested simultaneously in another model for CUs (of any type) with more than one member (Table 13). Reports with more than one respondent tend to be associated with relatively higher (moderate) effects when compared to reports by a single respondent; moderate recall effects are also associated with the use of records. Other interactions in this model (values not shown) indicate that records are more likely to be used by CUs with two people when compared to CUs of other sizes, and when more than one respondent participates.

5. Conclusions

The findings confirm the fact that recall effects in surveys are not a sole function of recency; the type of respondent, the family setting and the resultant level of interest affect the quality of responses. It is also evident that in order to have an effective interview in reporting family expenses, a

Table 12. Type of respondent

Variables: Recall effects (dependent), type of CU, who is the respondent and age of respondent.
Note: Restricted to husband–wife CUs with main respondent either husband or wife.

	Apparel recall effects		
	None	Moderate	Great
	Standardized values		
Interaction of recall effects and who is the respondent			
Both husband and wife	−2.75	2.81	−1.11
Wife only	0.59	−1.30	0.42
Husband only	2.13	−1.61	0.52
Interaction of recall effects and age of main respondent			
40 years old and over	0.82	−3.29	2.44
Interaction of recall effects and type of CU			
CUs with children or others	−2.13	3.62	−0.35
Jackknifed chi-square test	Test values	DF	
Who is the respondent	2.0	4	
Age of main respondent	2.8	2	
Type of CU	2.8	2	

Table 13. Record usage and number of respondents

Variables: Recall effects (dependent), record usage and number of respondents.
Note: Restricted to CUs with more than one member.

	Apparel recall effects		
	None	Moderate	Great
	Standardized values		
Interaction of recall effects and record usage			
Some records used	0.27	4.49	−5.31
Interaction of recall effects and no. of respondents			
More than one respondent	−2.33	2.48	−0.25
	Test values	DF	
Jackknifed chi-square test			
Record usage	6.0	2	
Number of respondents	2.1	2	

considerable effort is required, through record consultation and by participation of more than one family member. The analysis is suggestive of the potential effects on estimates derived from recall data.

Reports with most of the expenses relating to the first recall month are symptomatic of greater degrees of underreporting. An example of a family characteristic associated with this problem is the size of the family: greater recall effects are found for larger families. By contrast, no recall effects are apparent for single and two member CUs, and this can be interpreted as a consequence of an easier recall process. Education of the respondent tends to affect reporting quality, separately from age of respondent; the least educated respondents tend to report with greater recall effects.

Underreporting appears to be present even under the best interview conditions, e.g., when more than one respondent participates and some records are used during the interview. These reports show moderate recall effects, rather than no recall effects; however, reports with moderate effects display the highest mean expenditure levels. Husband–wife families with children, home owners, and respondents between the ages of 25 and 44 are associated with these response patterns. This seems to imply that respondents with relatively more (apparel) expenses attempt to report more completely than respondents with less involvement in these expenses, while still underreporting for the early portion of the reference period.

6. References

Andrews, F.M. and Herzog, R. (1986): The Quality of Survey Data as Related to Age of Respondent. Journal of the American Statistical Association, 81, pp. 403–410.

Biderman, A.D. and Linch, J.P. (1981): Recency Bias in Data on Self-Reporting Victimization. American Statistical Association, Proceedings of the Social Statistics Section, pp. 31–40.

Bishop, Y.M.M., Fienberg, S.E., and Holland, P.W. (1975): Discrete Multivariate Analysis. MIT Press, Cambridge, MA.

Cannel, C.F., Marquis, K.H., and Laurent, A. (1977): A Summary of Studies of Interviewing Methodology. National Center for Health Statistics, Vital and Health Statistics, 2, 69.

Goodman, L.A. (1978): Analyzing Qualitative/Categorical Data. MIT Press, Cambridge, MA.

Fay, R.E. (1985): A Jackknifed Chi-Square Test for Complex Sample Surveys. Journal of the American Statistical Association, 80, pp. 148–157.

Fay, R.E. (1987): Contingency Table Analysis for Complex Surveys (CPLX), Program Documentation. Washington Statistical Society Short Course, Washington, D.C.

Fienberg, S.E. (1980): The Analysis of Cross-Classified Data. MIT Press, Cambridge, MA.

Moss, L. and Goldstein, H. (1979): The Recall Method in Social Surveys. University of London, Institute of Education, NFER Publishing Co., Great Britain.

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

Sudman, S. and Bradburn, N.M. (1977): Response Effects in Surveys: A Review and Synthesis. Aldine, Chicago, IL, pp. 81, 92.

Tanur, J .M. (1981): Advances in Methods for Large-Scale Surveys and Experiments. In "The 5-Year Outlook on

Science and Technology", Source Materials, Vol. 2. National Science Foundation. Washington, D.C.

Tucker, C. (1986): An Analysis of the Dynamics in the CE Diary Survey. American Statistical Association, Proceedings of the Survey Research Methods Section, pp. 18–27.

Turner, C.F. and Martin, E. (1984): Surveying Subjective Phenomena. Russel Sage Foundation, New York.

U.S. Bureau of Labor Statistics (BLS) (1986): Consumer Expenditure Survey: Interview Survey, 1984. Bulletin 2267.

Received November 1987
Revised February 1989