Learning To Say No: Conditioned Underreporting in an Expenditure Survey

Jennifer Shields and Nhien To
Bureau of Labor Statistics

Abstract

The Consumer Expenditure Interview Survey (CEQ) is a comprehensive federal expenditure panel survey. Respondents are asked to provide detailed information about their expenditures over the course of five waves. For each reported expenditure, the respondent is asked a series of detailed follow-up questions; the more expenditures a respondent reports, the longer the interview becomes. Survey methodologists suggest respondents realize answering a question a certain way can lead to additional questions (Fowler, 1983; Sudman and Bradburn, 1982). Previous research has shown respondents may become conditioned to provide negative responses to reduce interview length (Jensen & Edelbrock, 1999; Lehnen and Reiss, 1978). A study looking specifically at CEQ data found evidence in some commodities of a decline in mean expenditures across interview waves (Silberstein & Jacobs, 1989).

Our study hypothesizes that respondents may become 'conditioned to underreport,' that is, they learn to say 'no' to questions about purchases after discovering that responding 'yes' increases interview length. To find evidence that interview length may condition underreporting, we propose to examine rates at which respondents report no expenditures. Focusing on these rates, as opposed to expenditure values, allows inspection of reporting patterns across waves without consideration of the amount of purchase reported.

This study will use respondents who started the CEQ over a one year period, approximately 8,000 cases. Using respondents who reported a purchase in a previous interview, the probability that they report the same type of purchase in subsequent interviews will be estimated. If the respondent is conditioned to underreport across interviews, the probability of reporting a purchase in subsequent waves will be lower if they reported the same type of purchase in a preceding wave.

Keywords: household survey, expenditure survey, underreporting

1. Introduction

The Consumer Expenditure Interview Survey (CEQ) is a comprehensive federal expenditure survey. The survey collects detailed information about household expenditures on a large number of items. Data from the survey are used for economic analysis and to provide weights for the Consumer Price Index (Jacobs, Jacobs & Dippo, 1989).

It is known that respondents to the CEQ do not report all their expenditures, or that they underreport. Comparisons to external sources such as the Personal Consumption Expenditures (PCE) (Garner, Janini, Passero, Paszkiewicz, Vendemia, 2003), and modeling of the data (Tucker, Biemer, Meekins, & Shields, 2004), has shown that, for whatever reason, respondents do not accurately provide all their expenditures. Many hypotheses have been developed as to the causes of this underreporting, including forgetting, not correctly understanding questions, rushing to complete the interview, and interviewer effects. This paper hypothesizes that a potential cause of underreporting is that respondents are conditioned to do so.

During the survey, respondents are asked to provide detailed information about their expenditures over the course of five waves. For each reported expenditure, the respondent is asked a series of detailed followed up questions; so the more expenditures a respondent reports, the longer the interview becomes. This paper is concerned with the rate at which a respondent gives a negative response following a positive response. Respondents may 'learn to say no,' or are conditioned to underreport, in order to to decrease interview length.

Survey methodologists have suggested that respondents recognize the patterns in a survey, and can identify the impact their responses have on the following questions (Fowler, 1993; Sudman & Bradburn, 1982). When a specific response prompts a series of additional questions, respondents may learn not to give that response, to reduce interview length.

Research has supported this hypothesis, showing that respondents may become conditioned to provide certain responses to reduce the length of the interview (Lehnen & Reiss, 1978). Even when reporting health disorders or symptoms, respondents were found to provide fewer positive answers later in the interview (Jensen & Edlebrock, 1999; Jenson, Watanabe & Richters, 1999). Looking specifically at the CEQ, Silberstein & Jacobs suggested that "respondents may report more selectively after the first or second interview in order

to shorten the interview" (1989, pp. 289). They, and others, found by looking at the CEQ expenditure data that there was evidence of some decline in mean expenditures across waves (Silberstein & Jacobs, 1989; Silberstein 1990).

Anecdotal evidence also supports this theory. Field interviewers have reported that they witness conditioned underreporting (Shields, 2004). They pointed out that the CEQ is a very long survey, and some respondents develop ways to make it shorter. As one interviewer pointed out: "Respondents learn that if they answer 'yes' to the main question, more questions follow, so they have learned to say 'no' to hurry the interview." (Shields, 2004, page 22). Combining theoretical and anecdotal support for the theory of conditioned underreporting provides a strong basis for which to explore the data for additional evidence of this theory.

2. CEQ

The Consumer Expenditure Quarterly Interview Survey (CEQ) is a nationwide survey of households used by the Bureau of Labor Statistics (BLS) to collect detailed expenditure information on the buying habits of American families. The CEQ collects larger items such as household appliances, furniture purchases, vehicle expenses, and travel expenses. Respondents are asked to recall expenditures over a three month period and report to an interviewer once a quarter for five consecutive waves. This study will be examining the section covering expenditures for Trips and Vacations collected in the CEQ. This section was selected for two reasons; it has clear screening question, a point where a yes response directly leads to a series of follow-up questions and a no response leads to none. The second reason is that field evidence suggests that the Trips and Vacation section is very burdensome to respondents, with the many questions asking about details often difficult to recall.

The nature of consumer behavior associated with trips and vacation expenditures needs to be considered prior to conducting analysis. Unlike many purchases, trips or vacations require time and additional money to be expended by the respondent. This makes these types of expenditures less likely to occur frequently than other types of purchases.

3. Identifying Conditioning Effects

3.1 Method

To gain a broad picture of the pattern of reporting across waves, the proportion of respondents reporting a trip in each wave was calculated. The first interview reporting was adjusted to account for the single month recall which differs from the three month recall period asked about in waves two through five. Research by Silberstein (1989) who found that large household expenditures (more than $100) recalled from first month accounted for approximately 53% of total expenditures reported for the three-month period. Of the purchases reported, only one-third of expenditures would be expected to reported in each month. The fact that more than half come from the first month can be attributed to telescoping and recall issues, and prevents wave 1 data from simply being multiplied by three.

Instead, the first wave interview data was reduced to account for the overrepresentation of the first month purchases; the difference between Silberstein's findings of first month purchases reported (53%) and the expected proportion (33%) was taken, and then multiplied by the wave one data. After this, the data is thought to show the actual trips taken during the first month, so multiplying the adjusted first interview data by three adjusts it to be similar to the other four waves of interview data. It is recognized that this broad type of analysis allows for many other interpretations, but will be used as a starting point.

The next step was to identify the frequency of respondents who fall into each possible response pattern. There are some pattern groups that would be susceptible to conditioning; those containing a 'yes' response followed by one or more 'no' responses would be indicative of conditioned underreporting. Groups such as YNNNN, NYNNN and NNYNN are groups that may have experienced conditioning. Within the groups, there is a very useful comparison available, the group who reported a single trip (or two or three) but was not eligible for conditioning (i.e. NNNNY). Using this group as the 'true' proportion of respondents who only took one trip, the other one-trip groups can be compared to determine if they occur at higher rates, which provides evidence for conditioning.

3.2 Limitations

The question of the occurrence of conditioned underreporting cannot be definitively answered by this series of analysis. The data cannot inform us as to the reasons a respondent provided a certain answer, nor can this type of exploratory analysis determine the 'true' response. However, evidence of conditioned underreporting can be found through this series of analysis. If evidence is found showing that this behavior may be occurring within this section of the CEQ, additional analysis techniques might be found to further investigate this theory.

Since the CEQ interviews households rather than individuals, it is possible that multiple respondents participated at each wave. This would mitigate the

effect of conditioned underreporting, as only respondents who reported a purchase previously would know the effect of such an answer. The available data on respondent for each wave is thought to be unreliable, so this confounding factor cannot be analyzed.

In addition, once each response pattern was identified, the cell sizes were quite small, even for the trip types most frequently reported. The small numbers of respondents falling into each group limits the types of analysis that can be done, and the conclusions that can be drawn.

3.2 Data

The data that will be used in this study will only include cases in which all five quarterly interviews were completed with start dates between April 2001 and March 2002. A complete year of data was used to control for seasonality effects, a potential problem with reports of trips, a type of purchase that varies greatly by season. In addition, only cases which had valid responses to the Trip screening questions will be included. The total number of cases included is 6,168.

It should be noted that the first interview is a bounding interview, different from the others as it has a one-month recall period rather than three months. The purpose of this interview is to control for telescoping effects, or the situation where respondents report trips they took earlier than the recall period. When comparing frequencies between waves, the frequencies reported for the first interview need to be adjusted to account for the shorter recall period. However, when examining patterns for evidence of conditioned underreporting, frequencies are not adjusted since it is only respondents who reported a trip that are susceptible to this.

In the section covering Trips and Vacations, respondents are asked six screener questions. Each screener question asks whether they or someone in their household has taken any specific type of trip. Trip types include any trips entirely paid for by anyone outside their household, such as a business, employer, or relative (OTSD); been away overnight or longer to visit relatives or friends (RELA); for business (BUSN); for recreation such as sightseeing, sports events, club or organizational meetings, or outdoor recreation (RECRE); any other kind of trip (OTHR); and trips in which they did not stay away overnight, but went somewhere at least 75 miles away from home (TRIP). If they answer "yes" to any of the screener questions, they are then asked a series of follow-up questions regarding that particular trip, such as "How many trips like this did you have?" and "Where did you (they) go on this trip?" If they answer "no" to any of the screener questions, they get to skip the follow-up questions and proceed to the next screener question.

Within the trip section, the majority of respondents never report a trip. More than 75% of respondents never reported a BUSN (94%), OTSD (80%), OTHR (91%) and TRIP (82%) trip types. The most trips were reported in RECRE (56% of respondents reported at least one) and RELA (38% of respondents reported at least one), and to take advantage of the larger cell sizes for these types, these were selected.

4. Results

To gain a broad picture of the pattern of reporting across waves, the proportion of respondents reporting a trip in each section across waves is shown in Figure 1. The first interview reporting was adjusted to account for the single month recall based on research done by Silberstein (1989) who found that large expenditures (more than 100%) recalled from first month accounted for approximately 53% of total expenditures reported for the three-month period.. Using this figure to reduce the number of trips reported in wave 1, and then multiplying it by three, the first interview figures were adjusted to estimate a three month recall period.

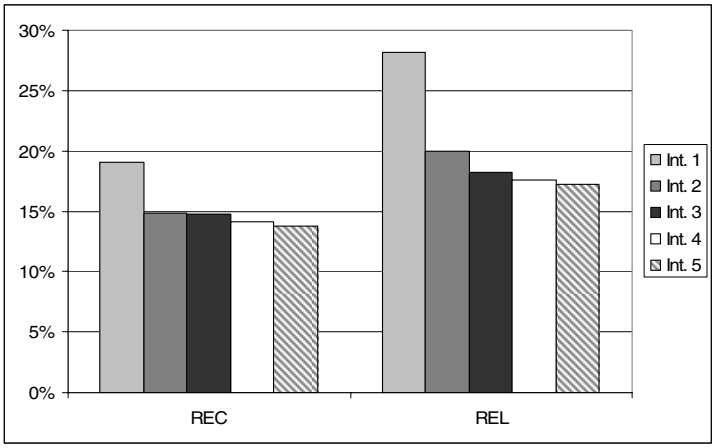

Figure 1. Percentage of Respondents Reporting Trips by Interview Wave

	RECRE			RELA		
Pattern	#	%	SE	#	%	SE
YYYYY	26	0%	0.1	98	2%	0.2
YYYYN	18	0%	0.1	30	1%	0.1
YYYNY	24	0%	0.1	24	0%	0.1
YYYNN	21	0%	0.1	29	1%	0.1
YYNYY	17	0%	0.1	30	1%	0.1
YYNYN	15	0%	0.1	23	0%	0.1

YYNNY	25	0%	0.1	35	1%	0.1
YYNNN	48	1%	0.1	74	1%	0.1
YNYYY	10	0%	0.1	30	1%	0.1
YNYYN	15	0%	0.1	22	0%	0.1
YNYNY	29	1%	0.1	22	0%	0.1
YNYNN	35	1%	0.1	67	1%	0.1
YNNYY	19	0%	0.1	31	1%	0.1
YNNYN	50	1%	0.1	52	1%	0.1
YNNNY	53	1%	0.1	55	1%	0.1
YNNNN	227	4%	0.2	312	5%	0.3
NYYYY	26	0%	0.1	51	1%	0.2
NYYYN	42	1%	0.1	56	1%	0.2
NYYNY	36	1%	0.1	45	1%	0.1
NYYNN	81	1%	0.1	92	2%	0.1
NYNYY	38	1%	0.1	44	1%	0.1
NYNYN	65	1%	0.1	97	2%	0.2
NYNNY	68	1%	0.2	84	1%	0.1
NYNNN	288	5%	0.3	322	5%	0.3
NNYYY	34	1%	0.1	56	1%	0.2
NNYYN	73	1%	0.2	68	1%	0.1
NNYNY	66	1%	0.1	70	1%	0.2
NNYNN	307	5%	0.3	277	5%	0.3
NNNYY	73	1%	0.2	67	1%	0.2
NNNYN	286	5%	0.3	244	4%	0.3
NNNNY	234	4%	0.3	233	4%	0.3
NNNNN	3819	62%	0.7	3428	56%	0.8

Table 1. Distribution of Responses by Pattern

Looking at the distribution of reporting by wave, a slight decline can be seen through later waves in both trip types, but the differences are not large. Overestimation made by the rough adjustment made to Interview 1 data may account for the large difference between wave 1 and wave 2; however it is suspected that reporting is higher for the first interview than the others, although the exact amount cannot be known because of the differing recall periods.

The next step was to identify the frequency of respondents who fall into each possible response pattern. For example, a respondent who reported one trip in the first interview and not in the last four would fall into group YNNNN. Table 1 shows the distribution of respondents between the patterns. This table also clearly shows the small number of respondents who fall into each pattern group. (The standard errors were calculated using weighted estimates.)

Within the distribution of response patterns, there are groups that are candidates for conditioned underreporting, that is there are response patterns which would be expected if this type of learning is occurring. Patterns containing a 'yes' response followed by one or more 'no' responses would be indicative of conditioned underreporting. These groups include YNNNN, or the group that after reporting one trip learned to say no on the second wave; NYNNN, the group that learned to say no after a yes during the second wave, etc. The list of response patterns which would be candidates for conditioned underreporting, and the frequency which respondents fell into the groups are shown in Table 2.

	RECRE			RELA		
Pattern	#	%	SE	#	%	SE
YNNNN	227	3.7%	0.2	312	5.1%	0.3
NYNNN	288	4.7%	0.3	322	5.2%	0.3
NNYNN	307	5.0%	0.3	277	4.5%	0.3
NNNYN	286	4.6%	0.3	244	4.0%	0.3
NNNNY†	234	3.8%	0.3	233	3.8%	0.3

†: 'true' value of one-trip respondents

Table 2. Distribution of One-Trip Respondents Eligible for Conditioning

There are two possible explanations for the proportions of respondents falling into these groups. The first is that they are conditioned to underreport, and the second is that they actually only took one trip during the interviewing period. To determine which explanation is more likely, the proportion falling into each group can be compared to an estimate of the 'true' value for that group. This 'true' value can be found by looking at the group who reported a single trip but was not susceptible to conditioning effects, or the group that reported a trip during their last interview (NNNNY). This value does not represent the gold-standard value for the group, as there is still underreporting of trips expected within this group for non-conditioning reasons, such as forgetting or proxy reporting.

Since we know that this group is not experiencing conditioning effects, we can use this number as the 'true' proportion of respondents who only took one trip. The difference between this value and the observed value in the other groups can then be to evaluate the proportion of respondents in the group who were conditioned. These differences are shown in Table 3.

RECRE RELA

Pattern	%	difference	%	difference
YNNNN	3.7%	-0.1	5.1%	1.3*
NYNNN	4.7%	0.9*	5.2%	1.4*
NNYNN	5.0%	1.2*	4.5%	0.7*
NNNYN	4.6%	0.8*	4.0%	0.2
NNNNY†	3.8%	--	3.8%	--

†: 'true' value of one-trip respondents; * significant at .01 level

Table 3. Differences between observed percentages and 'true' percentages

Comparing the observed percentages in each of the possible conditioning groups and the one-trip group where conditioning was not possible shows that there is evidence in the data that respondents are conditioned to underreport. For both trip types, for all but one group (YNNNN in RECRE and NNNYN in RELA), the difference is almost 20% the total size of the group.

For six of the eight comparisons, there is a statistically significant difference between the proportion of respondents in the groups susceptible to conditioning and the 'true' value group. This shows that there is evidence of conditioned underreporting in the data.

Respondents who took more than one trip also may have been conditioned to underreport. The patterns for these respondents are shown in Table 4.

	RECRE			RELA		
Pattern	#	%	SE	#	%	SE
YYNNN	48	0.8%	0.1	74	1.2%	0.1
NYYNN	81	1.3%	0.1	92	1.5%	0.2
NNYYN	73	1.2%	0.2	68	1.1%	0.1
YYYNN	21	0.3%	0.1	29	0.5%	0.1
NYYYN	42	0.7%	0.1	56	0.9%	0.2
NNNYY††	73	1.2%	0.2	67	1.1%	0.2
NNYYY†††	34	0.0%	0.1	56	0.9%	0.2

††: 'true' value of two-trip respondents; †††: 'true' value of three-trip respondents

Table 4. Distribution of Respondents Reporting More than One Trip Eligible for Conditioning

The same comparison can be done as that for respondents who reported one trip. The proportion of respondents reporting two trips not susceptible to conditioning (NNNYY) can be compared to the groups reporting two trips who are susceptible to conditioning (YYNNN and NNYYN), and the group reporting three trips as a 'true' value (NNYYY) can be compared to the group reporting three trips with possible conditioning effects (YYYNN). This comparison is shown in Table 5.

	RECRE		RELA	
Pattern	%	difference	%	difference
YYNNN	0.8%	-0.4%	1.2%	0.1%
NYYNN	1.3%	0.1%	1.5%	0.4%
NNYYN	1.2%	0.0%	1.1%	0.0%
NNNYY††	1.2%	--	1.1%	--
YYYNN	0.3%	-0.3%	0.5%	-0.4%
NYYYN	0.7%	0.1%	0.9%	0.0%
NNYYY†††	0.6%	--	0.9%	--

††: 'true' value of two-trip respondents; †††: 'true' value of three-trip respondents

Table 5 Differences between observed percentages and 'true' percentages

Unlike respondents who reported only one trip, those who reported more than one trip do not show the same evidence of conditioning. The differences between the groups eligible for conditioning and the groups not (NNNNYY and NNYYY) are very small, or are in the wrong direction. This lack of difference may suggest that if condition underreporting is to occur, it only occurs after the first reported trip. It could also be the case that after reporting two or three trips, respondents appear unlikely to become conditioned, possibly because they feel obligated to maintain consistent reporting.

5. Conclusions

Multiple approaches were taken to examine the CEQ trips and vacation data for evidence that respondents 'learn to say no.' The first method, looking at the overall trend of reported trip by wave revealed a slight decline in later waves, suggesting that respondents may be conditioned. The first interview data was adjusted to account for the shorter recall period, and may have overestimated the number of trips, exaggerating the downward trend. The results from this analysis do not provide conclusive evidence for conditioned underreporting.

The second approach did find evidence that respondents become conditioned. By comparing the expected 'true' proportion of respondents who took one trip to the proportion of respondents who reported one trip but may have been conditioned, it was found that more respondents fell in the groups susceptible for conditioning. This suggests that respondents may 'learn to say no' after reporting a trip. Consistent results were not found for respondents reporting more than one trip however; the proportion of respondents falling into the 'true' groups for two and three groups were not

significantly different from the groups reporting two or three groups but susceptible for conditioning.

Despite mixed results, it can be concluded that there is evidence of conditioned underreporting in the CEQ trips and vacation section. Although not consistently found using multiple methods, the statistically significant differences between the reporting pattern groups provides the evidence that this exploratory study was seeking.

6. Future Research

This study was an exploratory study to determine if the data supported the theory that respondents learn to say no to reduce burden. Now that there is at least initial evidence, additional work can be done to further examine the issue. The number of trips reported by respondents should be used as a factor, as it is possible that the rate of conditioned underreporting varies depending on the number of trips. Additional CEQ sections can be analyzed to determine if similar results are found across sections. Demographic groups, particularly education, can be identified to see if certain groups are more likely to have this type of evidence.

Qualitative work can also be done, interviews with respondents and field interviewers, to gain insight about the thought processes respondents have during the CEQ, and the situations interviewers believe conditioned underreporting occurs.

References

Fowler, F. J. (1993). Survey Research Methods. Newbury Park, CA: Sage Publications.

Garner, T. I., Janini, G., Passero, W., Paszkiewicz, L & Vendemia, M. (2003). The Consumer Expenditure Survey in Comparison: with a Focus on Personal Consumer Expenditures. Prepared for the Federal Economic Statistics Advisory Committee Meeting, March 21, 2003.

Jacobs, E., Jacobs, C., & Dippo, C. (1989). "The US Consumer Expenditure Survey." *Proceedings of the 47th Session of the International Statistical Institute.* Paris, pp. 123-142.

Jensen, P.S., & Edelbrock, C. (1999) Special Section: Subject and Interview Characteristics Affecting the Reliability of the DISC. Journal of Abnormal Child Psychology, 27(6), pg 413-415.

Jensen, P.S., Watanabe, H.K., & Richters, J.E. (1999). Who's up first? Testing for order effects in structured interviews using a counterbalanced experimental design. Journal of Abnormal Child Psychology, 27(6), 439-445

Lehnen, R.G. and Reiss, A.J. (1978). Response Effects in the National Crime Survey. Victimology, 3, (1-2), pp 110-160.

Moehrle, T.G. (1994). Seasonal Adjustment of Quarterly Consumer Expenditure Series. Monthly Labor Review Online, 117(12).

Shields, J. (2004). 2004 Consumer Expenditure Survey Field Representative Focus Groups.

Silberstein, A. R. (1989). Recall Effects in the CE Interview Survey. Journal of Official Statistics, 5(2), 125-142.

Silberstein, A. R. (1990). First Wave Effects in the US Consumer Expenditure Survey. Statistical Methods, Canada, 16(2), 293-304.

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

Sudman, S. & Bradburn, N. (1982). Asking Questions: A Practical Guide to Questionnaire Design. San Francisco, CA: Jossey-Bass Publishers.

Tucker, C., Biemer, P., Meekins, B., & Shields, J. (2004). Estimating the Level of Underreporting of Expenditures among Expenditure Reporters: A Micro-Level Latent Class Analysis. Presented at the Joint Statistical Meetings, 2004. Toronto Canada, August 8 – 12.