Issues in Proxy Reporting in the Consumer Expenditure Survey

Nora Cate Schaeffer
University of Wisconsin – Madison
16 November 2010

Proxy reporting about expenditures

Although proxy reporting is ubiquitous in surveys – for example in completing household rosters, reporting labor force participation, and detailing the health status of spouses and children – the use of proxy reporting for detailed reporting of expenditures in the Consumer Expenditure (CE) Survey entails challenges that differ from those for other topics and other surveys. Moore's (1988) review of studies about proxy reporting concluded that the studies available at that time did not point to a clear advantage for self reporters, but that the available evidence was not definitive. However, even though differences between proxy reports and self reports about participation in the labor market or major illnesses may have been modest (see Moore 1988), these findings cannot be extrapolated to reporting about expenditures by a consumer unit. More recent research suggests that proxy reports are most likely to be comparable to those given by the person if the proxy has an opportunity to have direct knowledge (either by participation or conversation) of the event, and the opportunities for such knowledge can be expected to be greater if the event reported about is engaged in regularly over a long period of time and is salient or important for other reasons, such as emotional or financial reasons (see, e.g., Sudman et al. 1994 ; Schwarz and Wellens 1997; Kojetin and Miller 1993; Tucker and Miller 1993). Clearly, labor force participation and major illnesses may share some of these features in a way that many expenditures do not.

The somewhat limited literature that addresses issues relevant to assessing proxy reporting about expenditures has recently been reviewed by Mathiowetz (2010) and earlier by Kojetin and Jerstad (1997). There is evidence that proxy reporting about expenditures might be of poor quality. Kojetin and Jerstad (1997) compared answers of self and proxy reporters about expenditures in a task simpler than that in the CE Interview Survey. They concluded that proxy reports were incomplete, even when considering reports at the category level, and more detailed reports were of even poorer quality. The number of purchases and their cost were significantly underreported by proxy reporters, and a comparison of reports of specific purchases indicated that proxy reports were incomplete. Because the task that Kojetin and Jerstad studied was simpler than that in the CE Interview Survey (e.g., one-month recall, reporting at grosser levels for some purchases) one would expect quality in the CE Interview Survey and CE Diary Survey to be even lower than the low quality they describe. Furthermore, it is useful to bear in mind that this negative assessment of the likely quality of proxy reports is based on a comparison to self reports – which are also flawed -- rather than on a comparison with an accurate criterion. Despite their conclusion that "the completeness and accuracy of proxy reports of these expenditures was distressingly low" (1991, p. 16), Kojetin and Jerstad also concluded that proxy reports are the only "realistic operational alternative" to self reports for expenditures. In the intervening years, however, there are reasons to suspect that the quality of self reports may be in decline, some features of the way data are collected that supported the view that proxy reports are the only realistic alternative to self reports have eroded, and other opportunities to collect expenditure data have become available.

Issues in redesigning the CE Interview Survey and CE Diary Survey

Ideally, in designing a study, a researcher begins with the questions: What do you want to know, and where is that information? In the case of a project as massive as the CE Survey, another question must be added, where will the information be by the time the redesigned CE is implemented?

The CE Survey's reliance on self and proxy reports for information about expenditures originated in a data collection environment very different from that found in many consumer units today. Self and proxy reports can only achieve even the limited level of accuracy they are capable of if the values being reported are actually encoded and if recall can be aided by records when that is needed. Several features of contemporary domestic information are particularly relevant for the quality of both self and proxy reports in the CE Survey:

1. Electronic and automatic payments may provide consumers little opportunity to observe and encode the values we will later ask them about, and use of these forms of payments can only be expected to increase both within and between consumer units.
2. The records that might supplement memory during an interview may be scattered over various websites – one for telephone service, one for gas and electric, and so on – stored in electronic images of bank transactions in an electronic bank statement, or delivered by e-mail for those who have gone "paperless."
3. Electronic records may be precise, but uninformative for some requirements of the CE Interview or Diary: Records may be difficult to interpret – for example, the entry on the bank statement may note that a payment was made, but it may not be obvious to the consumer which credit card the payment was for. Or electronic records may give a precise total because the bank statement notes the amount of a debit card transaction, but exactly what groceries were paid for is not indicated.

These changes are likely to compromise the quality of both self and proxy reports. Furthermore, current developments in electronic records and developments to come, will not be universally adopted by consumers. Consumer units are now, and probably will be increasingly, heterogeneous in the types of records of expenditures they are exposed to and maintain. The heterogeneity within and between consumer units in the types of records they maintain and the ways in which members of a consumer unit can access those records places demands on CE Survey interviewers and instruments, and the way that interviewers and survey instruments interact with these records affects how much the CE burdens respondents.

Just as the organization of information about expenditures in consumer units has changed since the CE Survey was designed, so has the internal organization of the records themselves changed. As a simple, but relevant, example, even a few years ago, a receipt was an uninformative list of prices, with no indication of what had been purchased. Currently, most receipts include detail about each item purchased, whether coupons were used, where the purchase took place, and how payment was made. This increase in the level of detail in receipts and the persistence of electronic records present both opportunities and challenges for the CE Survey.

Use of records in CE

As these remarks suggest, an alternative to reliance on proxy – and self – reports is greater and more systematic use of records. If encoding of transactions by respondents deteriorates, respondents will either rely more heavily on records or on estimation and guessing. Use of records could result in more complete data and reduce bias and presumably variable error as well. For example, Safir and Goldenberg (2008), analyzing CEQ data collected between April 2003 and March 2007, reported that aggregate total expenditures were higher when the interviewer reported that the respondent consults records and receipts. In an earlier study of record use in the SIPP cognitive research, Moore, Marquis, and Bogen (1994) found that although record use did not reduce omissions of sources of income, for sources of income that were reported, consulting records reduced underreporting of amounts. In an analysis of CEQ data collected between April 2006 and March 2008, Edgar and Gonzalez (2009) concluded that the odds that editing would be required was lower for those who consulted utility bills and other records.

Attempting to increase and systematize the use of records has different implications for the CE Interview and CE Diary because of their different reference periods and the different types of purchases they focus on. For the CE Diary, a design in which respondents made diary entries only for purchases for which they did not have detailed receipts would reduce burden substantially, but would require that individual purchases recorded on detailed receipts be classified or coded, either in consultation with the interviewer or centrally. It is possible that rewarding respondents for keeping receipts by not requiring that those purchases also have diary entries could motivate respondents to be more comprehensive in saving receipts (though it could also lead to less complete data if respondents did not record purchases for which they did not have receipts). For some respondents, the longer reference period used by the CE Interview and the types of purchases the CE Interview focuses on would require that the respondent consult records from a number of online and paper sources, and instruments and interviewer training would have to accommodate the use of records in multiple modes. By the time the redesigned CE Survey is fielded, it is possible that records could be scanned in the field and that optical character recognition could make some scanned records more fully digital.

Just as conclusions about whether or not proxy reports are adequate will necessarily be conditional (on the topic, the reference period, the relationship between self and proxy, etc.), so must be conclusions about the use of records. In addition to the accessibility, relevance, clarity, and portability of records, one must consider how easily they can be translated into the information sought by the survey. In an extreme situation in which the translation is so complex that the respondent and interviewer cannot manage the translation, the records can serve as a way of identifying another system (such as energy providers) from which information of higher quality can be obtained (O'Brien 2010), but this step is unlikely to be integrated into the CE Survey.

One reason that consumer units might be delegated the burden of transcribing self and proxy reports about purchases onto data collection forms (either directly in the diary or via the interviewer in the interview) is to remove the cost of dealing with records from the CE Survey. There is reason to believe that the current survey procedures pay a price in accuracy; it is also not clear how long consumer units will be bear the burden of translation and transcription given them by the CE Survey. However, the alternative, the use of records, is one of those components of a survey design that is appealing in principle but can become unwieldy as a large production survey attempts to deal with the variety and other details of the records respondents might produce. One appropriate comparison for the variability in the condition of records is the variability in the quality of the diary records that respondents keep. In addition, record use is a component of survey design that needs to be reassessed regularly because the form that records take and the methods available for dealing with those records keeps changing.

Roles of (or for?) technology

One of the challenges in redesigning the CE Survey is envisioning the state of technology at the time the redesigned survey enters the field. Some changes seem particularly important because of their implications for how respondents can be contacted, where their records might reside, and how diaries or other reports will be collected. Examples of changes that could be important for the CE Survey include reduced frequency of landlines and substitution of increasingly powerful smartphones (capable of making digital audio and digital visual records) for laptops and desktop computers. Implications of these changes for the CE Survey include:

1. Contacts with respondents become both less expensive and less intrusive. For example, automatically generated text messages could remind respondents to keep receipts or record purchases for which they do not have receipts, either at regular or random intervals. Such text reminders, even though they are as impersonal as a robocall with a digital voice, may not be experienced as impersonal (and irritating) in the same way, because an automatic text message would betray its non-human origins in fewer ways than would a digitized voice. Similarly, email with links to web-based diaries and recording forms could result in more contemporaneous reporting for some consumer units.
2. Consumer units could be encouraged to track purchases using technology in flexible and varied ways. For some respondents the record of a purchase could be an audio memo on a phone (which some future reliable voice recognition software could reformat as text), a photograph of a receipt (possibly with a camera provided by the study), or an entry in a web-based diary made from a smartphone at the time of purchase that could supplement (or substitute for) a paper diary kept at home.
3. Methods for reporting information about purchases for the CE Survey could also be tailored to the consumer unit’s technology, substituting for current paper technology. For consumer units with appropriate equipment and access, web-based diaries or data collection forms could be developed to support the use of records and the concurrent recording of expenditure information for both the CE Interview and CE Diary.

4. Collection of (supplementary) documentation of expenditures could be digitized. Records of respondents that could not be turned over (or e-mailed) to the interviewer might be scanned in the field (for both the CE Interview and CE Diary).

Segmentation of respondents and differentiation in field approaches

Regardless of whether a redesigned CE Survey relies more on records than the current design, persists in using proxy reports, or adopts an individual diary model, implementation will probably need to segment consumer units with respect to their composition, likely patterns of expenditures, and their technological profile. For some types of events (in this case, some types of expenditures), the relationship between the proxy and the person reported about (e.g., the relationship of spouse or of parent and child) serves as a proxy for some of these features of events that facilitate agreement between self and proxy reports.

In consumer units with only a single adult (some portion of 44% of consumer units are "single person and other consumer units") proxy reporting is not an issue (see Bureau of Labor Statistics 2009, Table 59). In consumer units with only husband and wife (22% of consumer units plus presumably those among the "other consumer units" in which there are only two adult partners), a single Main Record Collector might shoulder the burden of assembling the records needed for the CE Interview or the receipts needed for the CE Diary. Using a single Main Record Collector for these consumer units could result in little loss of data quality as compared to a design in which there was a Main Record Collector and an Individual Record Collector (Kojetin and Jerstad 1997; Edgar et al. 2006). In the remaining (approximately 30% of) consumer units with children over 6, the test of augmenting a Main Diary Keeper with Individual Diaries for consumer units with older children, suggests that using this approach is likely to improve the quality of reporting (Edgar et al. 2006).

A second type of segmentation with potential to improve data quality is segmentation based on likely spending patterns. Research would be needed to develop models to predict likely spending patterns based on information provided at the initial interview. Such information could be used to target follow-up (such as text or email reminders) aimed at most frequent expenditures, expenditures most likely to be overlooked, or consumer units most likely to need reminding.

A third potentially useful segmentation is based on the technologies used by the consumer unit, an inventory of which would be taken in the first interview and include all potential record collectors and reporters. Although some consumer units will continue to require paper, that number will rapidly decline – and at some point it will become cheaper to provide such consumer units with devices for responding than to maintain a paper data collection stream.

Although segmenting consumer units in various ways and tailoring approaches based on this segmentation requires substantial changes in design, it can be dismissed only if one calculates that consumer units will continue to accept current levels of burden, levels that will soon begin to appear even more burdensome as consumer units become less acclimated to working with paper and perceive it as clumsy.

Research in support of redesign

The proposals put forward above are, obviously, sketchy and superficial; some are easy to dismiss (and some probably deserve that fate). But redesigning the CE Survey would seem to require at an early phase background research of several kinds using in-depth field studies:

1. Describing the technology profile of consumer units of different compositions and income levels. In addition to describing hardware and software, such a profile would examine how consumer units use technology – smartphones, laptops, and desktops -- in their spending at home, work, and while traveling.
2. Complementary studies of what consumer units buy where, and where expenditures can be tracked. Such studies would document what types of purchases are bought or paid for over the web using the technologies just mentioned; what bills consumer units pay electronically; what payments are made automatically; how debit cards are used; what payments are reimbursed by

reimbursement accounts or insurance. In addition, field studies would explore where records of those transactions reside: what information is download into account management software; whether credit card, utility, and other bills are on paper, in e-mail, or scattered over vendors' websites.

3. Studies of receipts and records of expenditures. One daunting aspect of contemplating a greater reliance on records is the variety of forms records can take, the challenges that interviewers might face interpreting them, and the problems of turning those receipts into survey records. The goal of these field studies would be to collect records from a sample of consumer units to investigate how the CE survey could manage and code them.

Additional research with the existing CE Interview and CE Diary data that could inform a redesign include examining how records are currently used in the survey and by which consumer units. In addition, several types of profiles of consumer units that seem promising could be developed and tested, particularly characterizing consumer units with different patterns of item nonresponse and with different patterns of expenditure. One goal would be to predict which consumer units might need more extensive support to provide data, so that interviewing resources are deployed where they are likely to do the most good. Currently field organizations are attempting to develop and apply models that predict the likelihood of participation in order to guide the investment of scarce field resources where they are most likely to be effective; in an effort as massive and demanding as the CE Survey, such an approach could be extended to guide and tailor the application of such interviewing resources as reminders, personal telephone calls, visits from the interviewer, additional training for respondents, specialized instruments, and such.

Research about proxy reporting

If the redesign effort continues in a more traditional direction, and further studies of proxy reporters are seen as desirable, supporting studies should, if possible, move beyond using someone who reports for themselves as the criterion and to consider variable response errors as well as bias. In addition to such criterion-based studies, self and proxy reporters can be compared with respect to the types and number of items reported, the level of detail at which items are reported, level of spending reported, proportion of items reported that need editing, as well as the use of records. Although recent studies of proxy reporting consider the relationship of the proxy to the events being reported about, examinations of the accuracy of "proxy" reporters also need to consider important features of how a proxy is embedded in the overall survey design, which affects how well a proxy reporter performs. These include:

1. What is being recalled: Proxy reporting about expenditures is different from proxy reporting about health conditions or participation in the labor market, and proxy reporting may be more adequate for some types of expenditures than for others (see for example, Kojetin and Jerstad (1997)). In part, this is because the relationship of a proxy to an event varies depending on the event, as already noted. But it is also because some events (for example, some purchases) are associated with a collectivity (e.g., a family car), while other events are associated only with the individual (e.g., buying a new scarf). Some purchases may be individual in the sense that they are made by a person or made with that person's "own" funds. Some purchases may be made on behalf of the consumer unit or with shared resources.
2. The length of the reference period. For example, over three months, proxies may have more opportunity to observe (or to forget about) items that have been purchased than they do in a week.
3. The details of the task and the medium for reporting. As already suggested, proxy reporting for a three-month recall period is different from proxy reporting for two one-week diaries; this may partly be because the events asked about using a three-month recall period are different from those asked about in one-week diaries. But it may also be that details of the survey design, such as reminder calls from an interviewer, have different implications for self and proxy reports.

Conclusion

Although the assignment for this discussion was proxy reporting, the available literature suggests several conclusions that required expanding the discussion:

1. A comparison between proxy reporters and self reports is a weak starting point for assessing the advantages and disadvantages of proxy reporters. Self reports may be the criterion available in most research designs, but that criterion will overestimate the success of proxy report, and distract attention from simultaneously considering the impact of errors in self reports.

2. The existing literature is of limited use because the quality of proxy reports cannot be generalized from one domain to another. The literature suggests that the quality of proxy reports about expenditures appears to be poor.

3. The ways in which consumer units spend have become more varied, and records that store the information sought by the CE Survey have become more detailed and potentially useful for data collection.

4. The technologies available for collecting and recording consumer purchases have expanded since the CE Survey was designed and that expansion will only continue.

5. The population's tolerance of burden can be expected to continue to decrease, and considerations of both nonresponse and response error require designing for reduced burden.

6. In evaluating the contributions of proxy reporting – or increased reliance on records – an attempt should be made to consider total survey error, in particular variable and fixed response error and item nonresponse in addition to unit nonresponse and sampling error. Obviously, precise estimates of all these sources of error cannot be made. Nevertheless, discussions should make an effort to consider the likely order of magnitude of all these sources of errors, so that decisions about one element in the survey design, such as proxy reports, are placed in context.

References

Bureau of Labor Statistics. 2009. Consumer Table 59. Composition of consumer unit: Shares of annual aggregate expenditures and sources of income, Consumer Expenditure Survey, 2009. Linked: http://www.bls.gov/cex/2009/aggregate/cucomp.pdf

Edgar, J., J. Davis, S. Spell, R. Verlander, and G. Wetzel (2006). Individual Diary Feasibility Test Final Report. Bureau of Labor Statistics unpublished internal report.

Filippucci, C. and M.R. Ferrante (1997). Individual Diaries and Expense Documents in the Italian Consumer Expenditure Survey. *Journal of Official Statistics*, Vol.13, No.2, pp. 113-121

Kojetin, B. and S. Jerstad (1997). The Quality of Proxy Reports on the Consumer Expenditure Survey.

Kojetin, B. and L. Miller (1993). The Intrahousehold Communications Study: Estimating the Accuracy of Proxy Responses at the Dyadic Level. ASA.

Marquis, K., J. C. Moore, and K. Bogen (1994). Effects of a Cognitive Interviewing Approach on Response Quality in a Pretest for the SIPP. SRMS.

Mathiowetz, N. (2010). Self and Proxy Reporting In the Consumer Expenditure Survey Program. Bureau of Labor Statistics Consumer Expenditure Survey Methods Workshop. 8 December, Suitland, Maryland.

Mathiowetz, N. and R. Groves (1985). The Effects of Respondent Rules on Health Survey Reports. *American Journal of Public Health*. Vol. 75, pp. 633-639.

Miller, L. and C. Tucker (1993). The Intrahousehold Communications Study: A Typology of Family Cohesion. ASA.

Moore, J.C. (1988). Self/Proxy Response Status and Survey Response Quality: A Review of Literature. *Journal of Official Statistics*, Vol. 4. No. 2, pp 155-172.

Moore, J.C., K. Marquis, K. Bogen (1996). The SIPP Cognitive Research Evaluation Experiment: Basic Results and Documentation. RESEARCH REPORT SERIES, Survey Methodology #2010-04. http://www.census.gov/srd/papers/pdf/rsm2010-04.p*df*

O'Brien, E. (2010). For Clueless Consumers, Can Records Improve Data Quality? Paper presented at the annual meetings of the American Association for Public Opinion Research, Chicago, May.

Schwarz, N.and T.Wellens. 1997. "Cognitive Dynamics of Proxy Responding: The Diverging Perspectives of Actors and Observers." *Journal of Official Statistics* 13(2):159-79.

Sudman, S., B. Bickart, J.Blair, and G. Menon (1994). "The Effect of Participation Level on Reports of Behavior and Attitudes by Proxy Reporters." Pp. 251-66 in *Autobiographical Memory and the Validity of Retrospective Reports*, edited by Norbert Schwarz and Seymour Sudman. New York: Springer-Verlag.

Tucker, C. and L. Miller (1993). The Intrahousehold Communications Study: Family Cohesion and the Level of Knowledge. ASA.