

This PDF is a selection from an out-of-print volume from the National Bureau of Economic Research

Volume Title: Annals of Economic and Social Measurement, Volume 3, number 2

Volume Author/Editor: Sanford V. Berg, editor

Volume Publisher: NBER

Volume URL: <http://www.nber.org/books/aesm74-2>

Publication Date: April 1974

Chapter Title: Consumer Expenditure and Price Data: An Overview

Chapter Author: Robert Ferber

Chapter URL: <http://www.nber.org/chapters/c10110>

Chapter pages in book: (p. 299 - 305)

## CONSUMER EXPENDITURE AND PRICE DATA: AN OVERVIEW

BY ROBERT FERBER

This introductory article provides a summary of papers presented at the Conference on Consumer Expenditures and Price Data, May 3-4, 1973. The Conference, sponsored by the NBER Conference on the Computer in Economic and Social Measurement, examined sample design, data collection, respondent motivation, and error identification and corrections.

The six papers and numerous comments presented at the first three sessions of The Conference on Consumer Expenditure and Price Data<sup>1</sup> can be classified under four general headings:

1. How to design samples that collect consumer expenditure and price data.
2. What data collection instruments should be used.
3. How to motivate respondents to cooperate and to give accurate information.
4. How to pinpoint errors in the data.

The discussion at these sessions is summarized under these four headings, and an effort is made here to bring together discussion on the same topic occurring in different sessions. Since not all of the papers were able to be revised in time to meet publication deadlines, this introduction also provides a summary of all papers.

### 1. SAMPLE DESIGN

In their paper, "Sample Design Alternatives," Waksberg, Hansen and Steinberg provided a clear summary of the design of the 1972-73 survey of consumer expenditures and of some of the decision problems involved. Particular attention was given to the rationale for the diary approach and for the division of the data collection operations into use of a diary and use of a recall questionnaire on a panel basis. This type of survey design yields two separate sets of data on consumer expenditures, one from the rotating two-week diary samples with continuous one-week overlap, and the other set from the quarterly recall, using 2 subsamples, one for 1972-73 and the other for 1973-74.

Ways of designing a sample for the collection of the price data are also discussed in this paper, with particular attention to the sampling of the highly important rent component. Both for this component and more generally, the authors note that the techniques for selecting outlets still present problems. The two principal possibilities are to use the Census of Business data to select outlets with probability proportionate to the volume of activity, or to ascertain from consumers the outlets where they shop and to use this information as a basis for sample selection.

In his discussion of this paper, William Madow pointed to the need for investigating the extent to which the coverage of the survey of consumer expenditures corresponds with that of the sample for the price index. If at all possible, he

<sup>1</sup> A follow up conference has been organized by L. Taylor and J. Popkin for April 1974.

indicates that it would be highly desirable for the two samples to be comparable with each other and for data from these two samples to be combined. Ideally, Madow notes, there should be some overall design to incorporate both the consumer expenditure survey and the price survey within a single system, something that Waksberg indicated had indeed been given a fair amount of attention.

Another sort of conflict was brought out by Martin Frankel, who observed that frequently the goals of studies of this type at the national level may indicate a very different approach than if goals at the local level were to be achieved. The problem of reconciling these different goals within a single survey design is exceedingly difficult.

Regardless of the goals, Frankel noted that question remains in a study of this sort of the best allocation of funds to reduce overall survey errors.<sup>2</sup> The experience of the past would suggest that not enough has been allocated to coping with response errors and various other forms of nonsampling errors, and more attention needs to be given to this aspect in the future. As part of this overall approach, it would be interesting to consider the application of some of the newer devices of modern technology, such as the installation of recorders in the home to make it easier for people to report expenditures.<sup>3</sup> As several participants pointed out, a great deal of attention has been given to response errors in the current survey of consumer expenditures. This was the reason for the use of diaries to supplement the recall questionnaire, and it is hoped that nonsampling errors in the present survey will be greatly reduced.

## 2. DATA COLLECTION INSTRUMENTS

Much valuable information on the effectiveness of diaries versus recall for obtaining consumer expenditure data is provided in the paper by McWhinney and Champion reporting on the Canadian experience in testing these approaches (pp. 411-437 below). Overall, these results suggest that the diary approach yields more accurate information than the recall approach on food expenditures and about the same accuracy on other expenditures. The authors also find that a two-week period is about optimum for the use of diaries in a single household, though in fact results for up to four weeks were not too different. As in other diary studies, McWhinney and Champion find that expenditures in the first week tend to be highest of all, and the reason for this remains a mystery.

All things considered, the Canadian experience suggests that from a field point of view, diaries may be easier to use than a recall approach, principally because of interviewing problems in the latter case. However, for data reduction purposes diaries are more difficult to process and involve more expense than a recall questionnaire.

<sup>2</sup> In fact, the Census statisticians have given a great deal of attention to this question over the years, as have other survey statisticians, but the resolution of this problem in a particular situation can be highly controversial.

<sup>3</sup> An experiment of this type was conducted by Sudman and Ferber in a prior study for the BLS, with mixed results. See Sudman, Seymour, and Ferber, Robert, *Experiments in Obtaining Consumer Expenditures of Durable Goods by Recall Procedures*. Urbana: Survey Research Laboratory, University of Illinois, February, 1971.

Are these two approaches really competitive? As Robert Pearl pointed out, for most present purposes the diary approach and the recall approach may be complementary rather than competitive, with diaries being preferable for small, frequent purchases, and a recall questionnaire for other purchases. But for what period of time diaries should be employed is essentially unanswered. As Margaret Reid noted, what we need are methods that yield data accurate enough to meet our needs and with available resources.

How to collect data on expenditures out of the home is a problem that neither diaries nor conventional recall methods have yet solved. Some ingenious ways of doing this were suggested in the paper by Lester Frankel, "Food Expenditures, Out of the Home," such as by dividing the day into "occasions" on some meaningful criterion relative to what is being measured, and asking for expenditures for each occasion separately. As an example, in the case of food consumption, people might be asked what they ate before breakfast, what they ate during breakfast, what they ate between breakfast and lunch, and so on. This is likely to yield a more meaningful set of data because of the close linkage between food consumption and the occasion which the respondent is being asked to remember.

Also pointed out very clearly and illustrated in this paper are the pitfalls involved in extrapolating data from such a procedure any distance into the past. In addition to the biases resulting from increased recall periods, Lester Frankel points out that the responses should be analyzed as clusters, with the result that increases in the number of recall periods may become less efficient statistically. Observations from five past occasions do not necessarily represent five independent observations, and Frankel provides one means of estimating the likely true number of independent observations that such data may represent.

### 3. MOTIVATING RESPONDENTS

While the primary focus of this session (actually the first of the three sessions) was on the effect of incentives on respondent cooperation, Charles Cannell in his paper (pp. 307-317 below) very usefully placed this problem in the broader context of how to motivate people to give information, in the light of their natural tendency to forget and also of their reluctance to report certain types of information.

What is the answer? According to Cannell, we must use the interviewer and various other available devices through the medium of the interviewer to create leverage, and hence motivation, to overcome this reluctance. With more highly educated, verbal reinforcement may be quite effective and compensation may not be effective, while with the more poorly educated the reverse is likely to be true. While Cannell advanced this explanation in terms of social psychology, it can also be presented in economic terms. Thus, an offer of a given amount of compensation is much more likely to approximate the cost of time of a lower educated person than of a higher educated person. Hence, the compensation offer is more likely to be attractive to, and to be effective with, the lower educated and the lower income groups.

This result is, in fact, precisely what Sudman and Ferber (pp. 319-331 below) find in their experimental diary studies. Whatever effect compensation seems to

have is greatest in the low income areas and is hardly noticeable in the high income areas. The effect is quite noticeable in terms of cooperation, is not too clear in terms of data quality, and tends to increase with the number of diaries a family is asked to keep.

On the other hand, Barbara Bailar reported that preliminary results from a Bureau of the Census experiment with remuneration on the 1972 survey of consumer expenditures do not support this finding of increased cooperation on diary-keeping. However, the comparability of the two sets of results is not clear, since in the case of the Census experiment, the response rate was already very high (about 90 percent), the results pertained only to one weekly diary, and interviewers were told to either offer or not offer remuneration to all their sample households, depending on the group in which they (the interviewers) were assigned, unlike the Sudman-Ferber experiments. Further analysis is yet to be done on these Census data.

Even if remuneration has no effect on data quality, Sudman and Ferber suggest that incentives may still pay on purely economic grounds because of their tendency to reduce overall interview costs, especially given the high cost of personal interviews and of diary placement relative to the much lower expense for remuneration. Moreover, within a fairly wide range of alternatives, remuneration effects seem to be insensitive to the specific form of incentive as well as to whether the auspices is a government agency or another nonprofit type of institution.

The role of incentives was placed in a very different context in the case of commercial diary operations, by Jack Abrams of the Market Research Corporation of America. In such cases, where cooperation is sought by mail, response rates tend to be much lower, usually of the order of 10-15 percent. As a result, the principal source of error in such survey operations is likely to be nonresponse bias rather than response bias. Remuneration then becomes essential in order to reduce bias as much as possible.

The idea of remuneration is best broadened, Abrams noted, to include non-material items and rewards to groups other than the respondent's household. As an example, he cites a procedure used by MRCA to offer remuneration to churches for providing low income respondents for their panel. (An opposite sort of example is that recently used in a mail questionnaire by American Express offering, essentially to high income groups, to donate an amount of money to charities designated by the respondent, the amount increasing with the overall response rate to the questionnaire.)

In effect, Abrams reiterated the point made earlier by Cannell that the basic problem is to motivate respondents to cooperate, a point also expressed by Barbara Bailar when she urged that more studies be undertaken on motivating influences. In other words, as Fowler observed in his formal comments, remuneration is an additional tool to induce cooperation but its effect will depend to a very great extent on how this tool is handled and on how the interviewers are trained to use it.<sup>4</sup>

<sup>4</sup> It might be noted that the focus in these sessions was largely on personal interview situations. However, the evidence is that the same sort of problems hold for telephone interviews with minor variations, and the literature is replete with examples of favorable effects of remuneration in commercial work. Telephone interviewing is especially important for the future since it is likely to be used much more on consumer expenditure surveys both as an adjunct to the diary approach and as a method in itself.

#### 4. SPOTTING ERRORS IN THE DATA

Discussion of the detection and correction of errors in the expenditure and income data recorded on the interview schedules was begun prematurely during the discussion of the McWhinney-Champion paper when a number of comments were made on the pros and cons of using the so-called balancing difference on the schedules.<sup>5</sup> From a practical point of view an alternative to the balancing difference is development of an operational procedure for detecting errors through use of outlier methods of choosing questionnaires and variables for further examination. This was the topic of Philip Musgrove's paper (pp. 333-345 below) in which he compares three approaches for detecting outliers, and tests two of them with mixed results. In particular, the use of confidence intervals on univariate distributions, taking one variable at a time, yielded results in his simulation tests not much worse than the use of multiple regression to correct each of the key expenditure variables in turn for the effects of socio-economic factors that would be expected to influence the values of those variables. As Musgrove pointed out, however, his test data contained so many obvious outliers that a sophisticated test procedure was hardly necessary to spot them.

In his comments on the Musgrove paper, William Madow suggested that many of these outliers can be identified by first-stage processing of the data, using not only the variables themselves but also simple functions of them; thus, errors not causing outliers on the basic variables are found because they cause outliers on ratios, as well as in the residuals about regression lines by the Musgrove technique. Such processing can be done as part of on-line editing of interview schedules and, in general, of automating the entire data record preparation process, including imputation. Conceivably, some of the new devices suggested by Martin Frankel would contribute to the feasibility of automating the editing process. However, Musgrove noted, such procedures may have potential in a country like the United States but were hardly likely to be feasible in the expenditure surveys in the less developed countries in which he has been engaged.

#### 5. TOPICS FOR FURTHER INVESTIGATION

One of the most useful aspects of all three sessions was the discussion of controversial issues and the clarification of types of questions that require further study. A useful means of rounding out this summary would seem to be to outline briefly the principal such issues.

##### 1. On sample design:

- a. Is it a good idea in an expenditure study to separate diary households from recall households?
- b. How does recall of expenditures for an entire quarter decay when the recall is made one month, two months or three months after the end of the quarter, as in the case of the 1972-73 survey of consumer expenditures?

<sup>5</sup> The balancing differences refers to the practice of comparing the total of the different types of income reported with the totals of the expenditure and saving categories, and inferring that the data contain errors if the two sums are not within  $X$  percent of each other. The value of  $X$  may vary anywhere from 5 to 25 percent, depending on one's opinion of the completeness of the interview schedule and of the data provided.

- c. What are the efficiencies of different sample designs for a price sample, especially of a sample of outlets?
  - d. How can automation be introduced into new approaches to sample design to reduce survey errors?
  - e. How does one deal with the problem of estimation when, as in the case of the 1972-73 survey of consumer expenditures, one has two or more different sets of data on the same variables?<sup>6</sup>
2. On data collection instruments:
    - a. What is the best way to combine the diary method and the recall approach for collecting required expenditure and income data with sufficient accuracy, given limited resources?
    - b. Why do reported expenditures always seem to be higher for the first few days on a diary study?
    - c. How can one best design data collection instruments and develop interviewing procedures for obtaining accurate data on expenditures out of the home?
    - d. What is the shape of the memory decay curve in using a recall questionnaire? Does the shape of this curve vary with the type of commodity?
  3. On respondent motivation:
    - a. How can experimental studies be designed to yield a better understanding of the influences that motivate respondents?
    - b. How can such studies be designed to yield a better understanding of the influences that also motivate interviewers?
    - c. How is the effect of respondent incentives affected by interviewer training and the nature of the interview?
    - d. What is the effect of respondent incentives on data quality? How can "hard" data on this question be obtained in an expenditure survey?
  4. Spotting errors in the data:
    - a. What is the efficiency of different test procedures for spotting outliers?
    - b. Is a multivariate test for outliers feasible and practicable?
    - c. How should outliers be adjusted when they are spotted? Under what conditions should they be adjusted at all?
    - d. What is the feasibility of the use of real time systems for the screening of outliers while questionnaires are still in the field?

## 6. TWO CONCLUDING COMMENTS

By hindsight, most of the focus in these sessions was on consumer expenditure surveys. Indeed even more time could have been fruitfully devoted to this topic. The fact remains, however, that relatively little attention was given to the design of surveys for the collection of prices. This topic could almost constitute a conference in itself, except that so little work has been done in the area and so much needs to be done.

<sup>6</sup> The problem is further compounded in this particular survey by the fact that the diaries request actual expenditures whereas the recall questionnaires request "usual" expenditures for many of the same items.

As noted by a number of the participants, an especially useful feature of the conference was the opportunity to exchange ideas among people in both the public and private sectors working on different facets of the same question or who were working on the same questions. The interchange of ideas and of different views made possible by the conference seemed to help everybody's understanding of the techniques and problems of consumer expenditure surveys and provided all concerned with a much more complete picture of what was going on in this field.<sup>7</sup>

*University of Illinois*

<sup>7</sup> It should be noted that a new journal, *Journal of Consumer Research*, begins publication in 1974. Jointly sponsored by ten scholarly associations, it represents a unique interdisciplinary venture in this area.