

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

Volume Title: The Flow-of Funds Approach to Social Accounting

Volume Author/Editor:

Volume Publisher: UMI

Volume ISBN: 0-870-14182-1

Volume URL: <http://www.nber.org/books/unkn62-1>

Publication Date: 1962

Chapter Title: Consumer Surveys as a Source of Information for Social Accounting: The Prospects

Chapter Author: Robert Ferber

Chapter URL: <http://www.nber.org/chapters/c2043>

Chapter pages in book: (p. 383 - 410)

Consumer Surveys as a Source of Information for Social Accounting: The Prospects

ROBERT FERBER

UNIVERSITY OF ILLINOIS

BEFORE ways and means of improving anything can be considered, information has to be available on what is wrong, preferably on where the main sources of difficulty lie. The preceding paper by Arthur Broida is very useful in this respect, insofar as it discusses sources of error in consumer survey financial data and makes an attempt to evaluate the effect of reporting errors on averages and on distributional data as related to auto purchases.

In the present paper, we shall cover some of the same material discussed in the Broida paper, though from a different perspective. As one objective, we shall attempt to pinpoint the principal sources of error in consumer survey financial data from a methodological viewpoint. Here, we shall draw partly on Broida's results, and partly on the results of other research in this general area, principally on the Ford Foundation-financed Consumer Savings Project sponsored by the Inter-University Committee for Research on Consumer Behavior. Partly for lack of data and partly for other reasons, the relative importance of the different sources of error is indicated only approximately; at a future time, it is hoped that more precise estimates can be presented.

The second objective of this paper is to propose various courses of action for coping with these main sources of error. In the course of doing so, some preliminary results from the Consumer Savings Project and from the Federal Reserve car-buyer study will be presented which indicate the nature of the problems and of the type of improvements that may be expected.

Because of the orientation of this paper, relatively little will be said about the application of these financial data to social accounting. Detailed discussion of this question seems hardly necessary in view of the space already devoted to it in Broida's paper.

General Considerations

Improvement of the quality of consumer survey data poses the economic problem of resource allocation. The data can be improved

by additional expenditures in at least as many different directions as there are sources of error. For example, sampling errors can be reduced by increasing the size of the sample, or by seeking still more efficient sampling designs. The problem of skewness can be minimized by disproportionately heavy sampling from the relevant population groups.

In similar ways, additional expenditures can be used to reduce nonsampling errors. Thus, errors due to failure to contact some respondents can be reduced by scheduling additional callbacks. Errors introduced by reluctance of certain types of respondents to cooperate can be reduced by devoting additional sums to securing better cooperation, such as through more elaborate personal call and mail explanations, soliciting support from third parties, etc. The effectiveness of personal explanations can invariably be improved by increased expenditures for the selection, training, and supervision of interviewers. Errors introduced by misunderstandings can be reduced by devoting additional time and money to proper question wording and order, to more thorough pretesting, and to improving interviewer abilities. Errors due to inadequacy of memory can be reduced by additional efforts to convince respondents of the need to refer to records, perhaps even by offering various monetary incentives for doing so. Finally, many of these errors can be minimized by devoting additional resources to developing more appropriate estimation methods for item nonresponse, for nonresponse in general, for checking coding and editing procedures, and for over-all improvement of the data processing operation.

Theoretically, in any given situation, the appropriate allocation of resources among these many possibilities would be determined in accordance with established economic principles, relying primarily on the equality of marginal outlay with marginal returns. In actual practice, however, such an approach is not feasible because the appropriate functional relations are hardly ever known. Although substantial strides have been made in the past few years in solving the problem of resource allocation, there is little doubt that still greater progress will be needed before a semblance of any scientific approach is applicable to allocating a survey budget in such a manner as to maximize the over-all reliability of the data in any given situation per dollar spent.

Thus, it is possible to specify very closely the probable error reduction associated with given increases in sample size for different designs. It is also possible to specify closely the cost of samples of different sizes under given collection procedures. With somewhat less precision, we can often judge, for example, that additional

expenditures on improving interviewer abilities or on improving respondent cooperation will be much more productive than increasing sample size. In the latter case, however, we are not able as yet to measure the magnitude of this effect. Furthermore, in certain situations we cannot even judge in advance whether an expenditure for a given purpose may be productive. For example, it is often difficult to say whether more elaborately written explanations of survey objectives will increase or reduce cooperation.

The problem in any given situation is complicated by the fact that many individual surveys or survey programs may have more than one objective. Hence, the allocation of resources among different aspects of an operation may aid in improving one set of data while serving of relatively little value in improving another set of data. The fact that individual sources of error may have different impacts on different variables and on different methods of data collection only serves to complicate the problem still further.

Nevertheless, the problem of resource allocation is not insuperable, and it can be expected that ultimately efficient information will be available so that an approach approximating this technique may be used in allocating the survey budget. Knowledge about the functional relations between particular types of expenditures and about the reliability of the resulting data undoubtedly will grow with time, partly through operating experience but primarily through methodological research. Methodological research is also likely to be most productive in determining which survey techniques are most reliable for given types of data and of consumers and, hence, for developing new techniques of improved reliability.

Segmentation of Survey Errors

The efficiency of methodological research, and of survey analysis in general, can be improved considerably if prior thought is devoted to the segmentation of survey errors. The framework thereby provided for ascertaining the relative importance of different types of errors in a survey operation can be of immense value, not only for its aid in resource allocation but also for the greater ease of pinpointing possible causes of error and of specifying research that might be undertaken to reduce or eliminate such errors. Lip service has been paid to this approach in the past, but relatively little has been done in putting it to use. To some extent this is not surprising, since classification of errors by type requires not only a workable classification scheme but also a means of detecting and classifying errors when they do occur. Such devices are hardly feasible in the ordinary

SOURCES OF INFORMATION

data collection operation because of the special features that have to be incorporated in the survey for the detection and classification of error.

A basic prerequisite for this approach is some means of measuring the accuracy of consumer interview data. This is possible either through internal consistency checks or through the use of external information, as was done in the 1956 Federal Reserve survey of new-car buyers. Neither is particularly easy, but recent work has demonstrated that each is feasible, depending on the purpose at hand, and that immensely useful results can be obtained thereby.

THEORETICAL FRAMEWORK

It behooves us therefore to consider a possible framework for survey errors. For such a framework to be most useful, it must satisfy a number of criteria. Thus, it must be comprehensive relative to the types of errors under study; it must be of potential analytical value; it must be workable; and it should, if possible, be of use in discriminating among different reasons for particular errors.

The essence of such a general analytical framework has already been provided in part II of Broida's paper. This classification is reproduced below, adding two other sources of error for the sake of completeness:

- I. Errors relating to the sampling distribution
- II. Errors associated with noncontacts
- III. Errors relating to willingness and ability to cooperate
 - A. Attitude—the motivation problem
 - B. Understanding—the communication problem
 - C. Knowledge—the data availability problem
 - D. Circumstances of the interview
- IV. Data processing errors
- V. Errors due to faulty aggregation

From the viewpoint of social accounting, there is little doubt that the main avenue for improvement lies in types II and III. Errors in the first category are controllable by proper sampling designs, and those in the fourth category are amenable to control by various checking procedures.¹ Errors in the fifth category are also largely at the analyst's control.

¹ See, for example, S. L. Payne and W. D. Rugg, "A Sampling Plan for Verifying Punching Work," *Public Opinion Quarterly*, Summer 1948, pp. 328-330; R. B. Voight and M. Kriesberg, "Some Principles of Processing Census and Survey Data," *Journal of the American Statistical Association*, June 1952, pp. 222-236.

It is, therefore, in the second and third categories that research is likely to be most fruitful—the errors made in the field interviews. The subdivisions listed under these categories provide one framework for classifying such errors. (For the sake of completeness, one more subdivision might be added, namely, faulty selection of sample members.) From an operational point of view, a more detailed classification is likely to be needed, particularly with regard to pinpointing the nature of response errors. For this purpose, the following conceptual scheme may serve as a starting point, relating accuracy of data supplied to certain relevant characteristics of the interview situation. This classification presents the more common situations (and for this reason is not logically complete):

- I. The respondent knows the answer
 - A. He supplies the data as he knows it
 - B. He deliberately supplies inaccurate data—perhaps for mistrust of the interviewer or of the nature of the survey; perhaps because of a contrary disposition, etc. (a motivation problem)
 - C. He refuses to supply the data for one reason or another (a motivation problem)
- II. The respondent does not know the answer and is unaware of it (data availability problem)
 - A. He supplies data without realizing they are inaccurate—he may not have checked his sources or records recently; may not keep any records at all; or he is not well informed on the matter (as a parent, on an adult son's saving out of wage receipts)
 - B. He deliberately tries to supply inaccurate data (interaction between data availability and motivation problems)
 - C. He fails to supply data which should have been supplied—such as omitting to mention certain interest payments or dividend receipts
 - D. He refuses to supply the data (another interaction situation)
- III. He is aware of his lack of knowledge; in this case, he may either:
 - A. Offer to cooperate in obtaining the data
 - B. Make a guess not to show his lack of knowledge (a motivation problem)
 - C. Refuse to supply the data on one or more grounds (a motivation problem)
- IV. The respondent misinterprets the question and supplies the wrong data (the communication problem)

Clearly the answers of any particular respondent will not always fall in one category, though from a practical point of view some are likely to be mutually exclusive, e.g., I-B and III-A.

The notations in parentheses show how the different subdivisions under this classification can be combined to yield the more general subdivisions of response error presented previously.

This theoretical framework for response (including nonresponse) error serves to bring out several general points apart from its practical applicability. One is that interaction between response errors is possible: note subdivisions II-B and II-D. Another is that the problem of response error may be said to be essentially one of discriminating among five types of data:

1. Accurate data (I-A, III-A, potentially also III-B)
2. Refusal of data (I-C, II-D, III-C)
3. Deliberately inaccurate data (I-B, II-B, possibly III-B)
4. Accidentally inaccurate data, due to respondent ignorance or memory lapse (II-A and II-C)
5. Accidentally inaccurate data, due to misinterpretation (IV)

This classification is more feasible from an operational point of view than the preceding one.² At the same time, to the extent that certain types of families or types of subject matter or combinations of both can be segregated by these five categories, the problem of securing accurate information can be simplified considerably. This is because the best means of obtaining a given item of information will vary with the type of inaccuracy to which it is subject. Thus, if interest payments on savings accounts are often omitted, as may well be the case, one approach to this item would be to place heavy emphasis on memory recall and records; if income (and losses) from gambling are deliberately underreported, a different, perhaps more impersonal approach, may be needed; and so on.

By varying the method of approach or of asking for the data in accordance with the type of error likely to occur in the data, improvement in accuracy may well result. This assumes, of course, that operational criteria are established for identifying these different types of error and that these distinctions are meaningful in a practical sense.

² There still remains the knotty problem of distinguishing between errors brought about accidentally and errors brought about by deliberate intent. Difficult as this problem may sound, it does appear possible to make this distinction in practice in a great majority of cases in a panel operation. Nevertheless, mistakes are undoubtedly made in the process and for this reason considerable leeway has to be allowed in interpreting the results.

The third general point highlighted by this framework is that the problem of response error cannot be considered in a vacuum, but depends to some extent on the methods used to secure the data. For the willingness of a respondent to furnish data and the care with which he does so clearly depends on the method of approach used to solicit the data. If the method of approach is altogether unsuccessful, outright refusal will result—situations I-C, II-D, or III-C. If the method succeeds in obtaining some cooperation but at the same time engenders mistrust, situation I-B will be encountered; whereas if cooperation results to the extent that the respondent is willing to supply the data but not spend too much time doing so, situation II-A may arise.

The applicability of this particular framework, or any other, in a practical situation depends on the nature of the problem and the means available for detecting and classifying errors. The framework presented above is detailed enough so that the combination of error categories not easily distinguished should still leave, to a large extent, a meaningful and analytically useful system. However, this is not the only possible framework for measuring response error.

SOME EMPIRICAL JUDGMENTS

A major portion of the Consumer Savings Project has been concerned with the detection of errors in consumer reports of financial data and with identification by type of error. The framework presented in the preceding section has served as a general guide in this work. Since this work is currently in progress, and relatively few findings have yet been obtained, only fragmentary details can be supplied at this time.

The findings relate to errors in reports of consumer nonmortgage debt and of two asset holdings, one a fixed-dollar, relatively insensitive asset and the other a fixed-dollar asset about which consumers are fairly sensitive. The results of these pilot studies point to various inferences regarding the relative importance of different types of survey errors. These inferences will have to be taken at face value for the time being. They are therefore advanced in the form of hypotheses for future investigation.

1. *The relative importance of different survey errors varies with respondent sensitivity to the type of holding for which information is being requested.*

Thus, refusal of data on nonmortgage debt has been very low, less than 3 per cent in pilot studies in both Chicago and St. Louis, accounting for roughly corresponding proportions of total holdings

of this debt. Refusal of asset data tends to be much higher, rising with apparent respondent sensitivity to questions about particular assets to as high as 15 to 20 per cent.

In a similar fashion, the frequency of reporting of deliberately inaccurate data appears to vary with sensitivity to the particular holding. Deliberate withholding of debt information—either by refusing to acknowledge the existence of debt or by substantial underreporting—is relatively infrequent, seemingly occurring on less than 5 per cent of all interviews and accounting for an even smaller proportion of total debt. On the more sensitive assets, however, such withholding tends to rise to substantial proportions, accounting for roughly 20 to 30 per cent of total holdings of one sensitive asset in one pilot operation.

To a lesser extent, accidentally inaccurate data due to respondent ignorance appear to be more frequent as sensitivity to the request increases, occasioned by the greater reluctance of the respondent in these instances to consult records.

These findings lead to the following corollary hypotheses:

2. *For debts and relatively insensitive assets, accidentally inaccurate data (both Types 4 and 5) are the primary cause of response error, principally errors due to respondent ignorance or misinterpretation.*

3. *For sensitive assets, the primary causes of response errors are refusal of data and deliberately inaccurate data.*

Segregation of response errors due to misrepresentation from those due to accidents of one type or another is admittedly not an easy task and is itself subject to substantial error. Such errors notwithstanding, these two hypotheses find strong support in the results obtained to date, as noted above. The more sensitive a person is to reporting holdings of a certain asset, the more likely he is to refuse or to substantially understate the true amount. Errors due to honest accident or misinterpretation appear to be present more or less in all asset and debt reports. Perhaps for this reason, the latter assume primary importance in the reporting of less sensitive financial holdings.

4. *Reported small holdings of an asset or debt tend to be overstated, while large holdings tend to be understated. At the same time, omissions of small holdings are likely to be both accidental and frequent, whereas omissions of large holdings are less frequent and less likely to be accidental.*

This hypothesis is roughly in accord with the empirical findings reported by Broida, though at the lower end of a distribution of

holdings omissions might well more than offset the overstatement of those holdings that are reported.

5. Relatively little relationship is apparent between reporting errors and respondent socio-economic characteristics, though there is some tendency for refusals to possess more of a given holding than respondents.

Sensitivity appears to be largely a personality characteristic cutting across all socio-economic groups. To some extent, sensitivity to reporting different holdings varies by holding as well as by respondent. Thus, some respondents will report holdings of stock, but not of savings accounts while others will do the opposite. At the same time, more respondents are sensitive to reporting certain types of holdings than others. It is these differential frequencies, brought about in part by personality characteristics conditioned perhaps by social custom and by actual holdings, that account for the observed differential sensitivities—and hence, survey errors—in obtaining data on different assets and debts.

6. The relative importance of different survey errors varies with the form in which data are requested.

Very few refusals are encountered in asking for changes in holdings rather than for levels. Perhaps partly because of this fact, the frequency of accidentally inaccurate data, both Types 4 and 5, tends to rise sharply in reporting changes. This is particularly true of a holding in which a large number of relatively small transactions are made during a period, such as charge accounts or checking accounts. The tendency in such instances is to gloss over changes that are small in the opinion of the respondent as well as to “correct” the data for “unusual” events.

These tendencies appear to be even more pronounced when changes are requested by mail instead of by personal interview. There is, then, a strong inclination to rationalize that “small” changes are of no interest and to minimize one’s effort by reporting no change.

7. The relative importance of survey errors is influenced by the interaction between the respondent and the interviewer.

Preliminary indications are that aggressive interviewers are less likely to obtain accidentally inaccurate data than other interviewers. The reason is that aggressive interviewers are more likely to insist on the use of records and tend to be more successful in achieving this objective.

It is clearly evident that outright refusal is less likely when rapport is good. Deliberately inaccurate data are also less likely to be

obtained in such situations. Instances of confusion do not seem to be less frequent, however, unless the interviewer takes special pains to make himself clear.

With any given respondent, there is little doubt that the size and nature of reporting errors are also influenced by the personality of the interviewer and by the interviewing approach.

Approaches to Improvement

A general conclusion that can be drawn from the foregoing material is that the reduction of survey errors requires research on many different facets of the survey operation. Different types of errors involve different types of interviewing problems, as noted in the detailed theoretical framework for response error, above; and the relative importance of the different types of error varies with the subject of the interview as well as with the interviewing situation.

Accordingly, a number of possible courses of action are indicated. In this section we shall highlight areas in which research on the collection of financial data by consumer surveys may be particularly fruitful. No attempt is made to indicate the relative importance of the need for work in these different areas—in fact, there is some overlapping of areas, as indicated at appropriate points—but there is little question that all of them warrant considerable attention.

RESPONDENT MOTIVATION

Motivating a sample member to cooperate in a survey and provide the information requested is in many ways the heart of the data collection problem, relating particularly to the refusal and deliberate misstatement of information. As noted by Broida, the problem is especially difficult in surveying consumers on matters likely to be considered confidential. If consumers could be motivated properly, immense improvement in data reliability would undoubtedly ensue.

From an analytical point of view, the motivation problem may be subdivided as follows:

1. Inducing the sample member to grant an interview
2. Persuading him to provide complete and accurate information
3. Securing further interviews, if this is the plan of the study

Initial Cooperation

The type of approach to take to an individual—what appeals or motives to use, the advisability of advance notice, whether to press for an immediate interview, and related matters—is a crucial aspect

of the interview situation about which much is unknown. Virtually nothing is known about the effect of different appeals in motivating individuals to grant interviews on their finances. Preliminary results from the Consumer Savings Project indicate that

the appropriate appeal can vary substantially with the personality and outlook of the individual. An appeal based on the social value of savings information will work very well on one person, but will antagonize completely another individual who may interpret this as just another step toward a fully regimented society. Reference to the need for savings information for research purposes is highly effective with some, but only serves to irritate others who assert that such research is a social waste. To be sure, it is apparent that some of the people who object to particular appeals are essentially looking for an excuse not to cooperate and would object to any appeal made to them. However, most objectors do not appear to be of this type, as is clear from the success frequently obtained by interviewers once appeals are switched.³

Evidence obtained from various operations of the project indicates conclusively that, at the least, printed materials and reports dealing with the survey or panel operation are definitely helpful in making people more willing to supply financial information. Sending the people photostats of newspaper articles about the study and periodic, nontechnical reports dealing with substantive findings appear to serve two distinct purposes. On the one hand, they help convince the respondent of the authenticity of the study. For this purpose, there seems to be no substitute for the printed word, especially when the printed word comes from a newspaper. Second, this material, particularly the reports, serves to show the individual how the data that he supplies are being used; and they serve as partial proof for the statement used frequently by the interviewers that the data obtained from different individuals are combined for analytical purposes so that anonymity is assured.

Even with newspaper releases and reports, the numerical effects of these factors have yet to be measured and optimum policies formulated. In addition, it remains to be determined whether or not it pays to use gifts even then, and if so, of what form, at what interval, and with what type of people.

The prior selection of the most effective means of motivating people

³ Robert Ferber, *Collecting Financial Data by Consumer Panel Techniques: A Case Study*, Studies in Consumer Savings, No. 1, Bureau of Economic and Business Research, University of Illinois, Champaign, Illinois, 1959, p. 67.

to grant such interviews would contribute immeasurably to the efficiency of consumer financial surveys.

Consulting Records

Another major advance lies in inducing respondents to give financial data after consulting records, and not from memory. Convincing people to cooperate is, of course, a necessary prerequisite; but it is not sufficient because, for one reason or another, many cooperative people do not want, or feel it necessary, to consult records. Yet, such little work as has been done indicates that use of records may contribute substantially to accuracy of reporting.

Panel members in the St. Louis operation of the Consumer Savings Project who consulted savings passbooks were asked in one round of interviews to sign a request to their savings institution for verification of their balance. Twelve did so: most of the 170 had been interviewed by mail, some did not consult passbooks, and some refused. The sample for analysis is even less, for in five instances the correct procedure was not followed. The sample is, therefore, quite small and is also clearly not representative, but the uniformity of the results nevertheless seems to be significant: in one instance the bank report differed by 5 per cent from the panel member's report (but the balance in the account was only \$6), in two other instances (of balances of \$300-\$400) the difference was less than 2 per cent, and in four instances (three- and four-figure balances) the two figures corresponded exactly.

Admittedly, this was not a controlled experiment; but it does serve to reinforce other indications that use of records is a major, if not the only, means of ensuring accuracy of consumer financial reports. How to induce consumers to do so, however, is as yet largely an unanswered question.

Continuing Interviews

The unique contributions that panels can make to the study of consumer behavior have been recognized for some time. In the course of a panel operation, rapport can be established to a far greater extent than is possible in a single interview. Partly because of this fact it appears that the procedure of repetitive questioning made possible by the panel technique contributes to the securing of more complete and accurate financial information. Thus, Table 1 indicates the extent to which additional information on holdings was picked up in two initial panel operations.⁴ To be sure, panel opera-

⁴ Additional details on the Chicago study will be found in Ferber, *Collecting Financial Data*, Chap. 3.

CONSUMER SURVEYS AS A SOURCE

TABLE I

PICK-UP OF SELECTED HOLDINGS ON PANEL OPERATIONS IN CHICAGO AND ST. LOUIS

<i> Holding</i>	<i> Pick-up as % of Total Savings Units^a Reporting Given Holding</i>		<i> Pick-up as % of Total Amount Owned of That Holding</i>	
	Chicago ^b	St. Louis ^c	Chicago ^b	St. Louis ^c
Checking accounts	13	7	9	8
Savings accounts	31	13	10	4
Mortgages lent	0	28	0	74
Life insurance ^d	35	18	9	11
Real estate ^e	11	3	†	2
Stocks	29	19	2	2
Government securities	20	20	21	3

^a Defined as one or more related individuals living in the same dwelling unit and pooling their income and savings.

^b Pick-up based on five waves of interviews, 1957-58.

^c Pick-up based on three waves of interviews, 1958-59.

^d Face value.

^e Equity, including own home.

† Less than 0.5 per cent.

tions are not without disadvantages, not the least of which are cost and the maintenance of representativeness. Nevertheless, to the extent that panel techniques do provide more complete information, it becomes all the more desirable to explore their uses for collecting financial information and the means of motivating people to participate in a panel operation.

Motivating a respondent to participate in a panel operation is often a more difficult task than securing a single interview. Some respondents seem to have the feeling that by having granted one interview they "have done their part," and there is no need for them to participate any further. In addition, some very busy people are willing to grant one interview but not to go any further. Here again, there is need for experimentation with different motivational approaches to ascertain their relative effectiveness on different groups of people and in different types of situations.

THE USE OF INCENTIVES

How effective are incentives in inducing people to cooperate in financial interviews? The problem is complicated by the fact that on the one hand there are many different types of incentives to be considered, ranging from payment of currency to mailing the respondent copies of reports. On the other hand, any particular incentive can be used under many different circumstances, as is

SOURCES OF INFORMATION

evident from the preceding section. In the field of marketing research, and at times in social research, incentives in the form of either money or gifts have proven successful in securing cooperation from people who otherwise would be most reluctant, particularly among low-income families.

No comparable experiments seem to have been made in the past with regard to the collection of financial data, and it is not at all clear whether the generalizations drawn from these other areas are applicable to the present case. The experience of the Consumer Savings Project with some uncontrolled experiments has been that a gift is more likely to be effective among lower-income groups. At the same time, when higher-income people are offered gifts, the interviewer is often rebuked for doing so. The unnecessary nature of the gift was particularly stressed by most people receiving such a gift in the course of the Project's Chicago pilot panel operation. At the same time, objective evidence pointed to a noticeable improvement in rapport just after the gift was sent out.

QUESTIONNAIRE CONSTRUCTION

There is little doubt that the form and organization of the questionnaire affect not only the accuracy and completeness of the data that are obtained but also the willingness of the respondent to cooperate. Thus, is it desirable to lead into questions on financial holdings only gradually, after a long series of "ice-breaking" attitudinal and other questions, or should an attempt be made to take up the financial data questions virtually at the outset? Presumably, the long introduction serves to relax the person and enables the interviewer to establish sufficient rapport, so that by the time the financial data questions are asked a respondent feels that he knows and can trust the interviewer. On the other hand, there is the possibility that the introduction gets to be very long, or that the person is very busy anyway and can only give about ten or fifteen minutes of his time. As a result, by the time the interviewer reaches the questions dealing with financial data, the respondent is in no mood to spend much more time on the interview, and either refuses outright or in one way or another speeds things up; so incomplete information on financial holdings is obtained.

Clearly this is a matter which depends partly on the interaction between the respondent and the interviewer as well as on the interviewing situation generally. If the respondent is convinced of the authenticity of the project and the need for providing accurate information, there would seem to be no reason why an interviewer could not ask the financial data questions at the outset. Thus, various interviewers have reported considerable success with an

CONSUMER SURVEYS AS A SOURCE

approach to the respondent such as the following: "Now let's first get out of the way these figures I have to fill in on the characteristics of your family and on your financial holdings. Then, we can sit down and talk about your savings practices and attitudes."

To be sure, this is not an approach which would work with every respondent, and it is not an approach which could be used by every interviewer. However, there seems no question that this approach is more effective under certain conditions, and it remains for these conditions to be identified and for the margin of relative efficiency to be established.

Whether to use a structured or an unstructured questionnaire is another basic question on which much research remains to be done. Recent experience suggests that unstructured questionnaires may be more effective in securing rapport. Thus, Table 2 presents response

TABLE 2
RESPONSE RATES BY INTERVIEWING APPROACH AND REQUEST FOR DOLLAR FIGURES,
ST. LOUIS PANEL OPERATION, WAVES 1-3

<i>Approach</i>	<i>Request for Figures</i>	<i>Per Cent of Eligibles Interviewed</i>			<i>Base (No. of Eligible Savings Units)</i>
		Wave 1	Wave 2	Wave 3	
Structured	Wave 1	76%	65%	60%	94
	Later	72	64	57	88
Unstructured	Wave 1	81	70	71	49
	Later	91	80	71	45
Total structured		74	65	59	189
Total unstructured		86	74	71	94
Total figures requested, Wave 1		77	67	65	143
Total figures requested, later		78	69	62	133
Total		78	68	63	276

rates obtained in the first three waves of the Consumer Savings Project's St. Louis panel operation by type of questionnaire approach and by whether or not figures on dollar holdings were requested in the first wave or later on.

The results suggest that postponement of the request for dollar figures did produce better response initially with an unstructured approach but not with a structured approach. On the other hand, the absence of a formal questionnaire appears to have produced higher response whether or not figures were requested in the very first interview, and to have had longer-run effects as well.

These results are not conclusive, particularly since a relatively small number of interviewers were involved (about sixteen), who could not be

randomized by questionnaire approach. Randomization was effected by sample member and by request for figures. Further research is being carried out on this subject in other operations of this project.

If an unstructured approach is found to be superior, means of handling various problems distinctive to this approach will be needed. Thus, this approach increases the possibility that various details of financial ownership may be omitted precisely because of the absence of a structured questionnaire. To the extent that this is true, it would seem worthwhile to experiment with modifications of an unstructured questionnaire approach to remedy this deficiency. At the same time, this raises the question of the most desirable means of interviewer selection and training, a subject which is reviewed briefly in the next section.

There are many other questions relating to questionnaire organization on which research is needed. Thus, there is the question of whether to obtain information on debts by asking for it directly or by asking for it with regard to the various possible reasons for debt, or sources for debt, or in some other way. The optimum length of a questionnaire is another question that remains virtually unanswered. Should all assets and debts be covered in a single questionnaire, or should an attempt be made to subdivide the various financial holdings among two or more questionnaires, and if so, how?

INTERVIEWER SELECTION AND TRAINING

Despite all the work that has been done on interviewer effects and bias, we still are a long way from being able to select successful interviewers for a particular survey. Aside from some of the most obvious criteria—ability to make oneself understood, literacy, pleasant disposition—the current practice is much the same as it was years ago, namely, to select people who seem as if they would make competent interviewers, send them out into the field, and hope for the best. If an interviewer does poorly, this would presumably come to the supervisor's attention sooner or later and that interviewer is dropped—though often not before damage has been done. In this way, by the time the survey is ended, only the best interviewers are left. Nevertheless, this is a small consolation, considering that the survey is already finished and that the knowledge gained of who actually are the best interviewers can hardly be put to any further use, at least in that survey.

In any survey dealing with a socially sensitive subject—and the collection of financial data certainly falls within this category—the interviewer becomes doubly important. Since the sample member is likely to be especially suspicious in a survey of this type, any mistake

CONSUMER SURVEYS AS A SOURCE

or questionable action on the part of the interviewer can have serious effects on the degree of cooperation obtained. It becomes therefore all the more necessary to investigate new means and new approaches to the subject of interviewer selection and training.

This is an area in which new and more imaginative techniques need to be introduced. It might do well to re-examine the standard clichés of the past regarding what makes a good interviewer, at least as applied to the special problem of collecting financial data from consumers.

For example, the evidence of the Federal Reserve survey of new-car buyers indicates that discrepancies between purchaser and lender reports were, if anything, somewhat less for inexperienced interviewers of buyers than for experienced interviewers, for those who attended a training conference than for those who did not, and that later interviews tended to produce larger discrepancies than earlier calls (Table 3).

TABLE 3
DISCREPANCY BETWEEN PURCHASER AND LENDER REPORTS OF CAR PURCHASE LOANS,
BY SELECTED INTERVIEWER CHARACTERISTICS, FEDERAL RESERVE
SURVEY OF NEW-CAR BUYERS

<i>Characteristic</i>	<i>Category</i>	<i>Ratio of Average Absolute Discrepancy in</i>	
		<i>Principal Loan</i>	<i>Total Loan</i>
Attended training conference	Yes	10.3	7.6
	No	11.2	7.9
	No answer	12.6	8.9
Number of interviews prior to this survey	None	10.7	6.7
	Less than 100	9.9	8.0
	100 or more	11.8	8.6
	No answer	12.7	8.9
Date of interview	Before June 15	7.9	8.4
	June 15-30	10.4	7.7
	July 1-15	11.0	7.5
	July 16-31	12.6	8.9
	August 1 or later	18.2 ^a	22.4 ^b
Total		11.1	7.9

SOURCE: Unpublished tabulations provided by the Federal Reserve Board.

^a Based on two observations.

^b Based on three observations.

Along a similar line, in the Consumer Savings Project inexperienced interviewers were found to be better able to handle an unstructured

SOURCES OF INFORMATION

approach than either experienced structured interviewers of the marketing research variety or experienced unstructured interviewers of the social research variety; and early interviews were found to be more complete and easier to carry out than later interviews. Such evidence indicates that the training and qualifications for an interviewer for this type of survey may in some ways be very different from those qualifications that would be sought for an interviewer in a more usual type of survey. Exactly what these qualifications should be, however, remains to be answered.

Research is also needed on the effect of different forms of interviewer training on interviewer efficiency. Exactly how much training should an interviewer have for this type of work? How much additional accuracy or efficiency can be expected per dollar spent for this purpose? These are questions which remain unanswered not only for this type of survey but also for virtually all other surveys as well.

It might also be worth noting that continual supervision is needed if interviewers are to be employed for a long period of time either on a number of different surveys or in connection with a panel operation. The fact that an interviewer has done a trustworthy and fully reliable job at one time does not mean that he will continue to do so. Techniques are needed and are currently being tested for evaluating interviewer effectiveness over time.

COST

The cost of consumer financial interviews is considerably higher for the amount of data received than is the cost of comparable interviews with alternative sources of similar information. The sample members are invariably questioned by personal interviews, which tend to be rather lengthy and for which considerable interviewer training is required. As a result, the cost per field interview in surveys of this type can range anywhere from \$15 to \$30, and may at times be even higher. These figures are high enough to serve as a major deterrent to consumer survey operations, and means of reducing these costs would clearly be most welcome.

Two means of reducing field costs would seem to warrant serious consideration. One approach is to experiment with compensating interviewers by a method other than the usual hourly pay, portal-to-portal procedure with expenses additional. The bulk of interviewing expenses are incurred in travel time and in travel expense. It is evident that the hourly pay method of compensation does not provide much inducement to the interviewer to economize.

Though the textbooks on survey operations advise against it, payment by the interview or payment by some combination of a flat

fee for travel plus a high rate per unit of time spent in the interview situation might well be worth further consideration. The circumstances under which these interviews are conducted would seem to invalidate many of the usual criteria governing the determination of interviewer compensation.

Second, and potentially the most effective approach, is to experiment with obtaining these data either by mail or by telephone. Admittedly, this may not work with many people and in many situations. However, to the extent that such an approach is effective, if only in a minority of the interviews, substantial cost reductions may nevertheless be achieved.

The mail questionnaire approach may be particularly useful in a panel operation. Thus, after three waves of personal interviews with panel members in the St. Louis operation of the Consumer Savings Project, more than 60 per cent were induced to fill out a questionnaire on the fourth wave by mail. Those who did not fill out the questionnaire by mail were followed up by personal interview. Substantial savings were achieved in this manner, though there remains the possibility that the reliability of the data supplied by mail may not have been as high as that obtained by personal interview.

Reducing the frequency of interviews and altering the size and format of the questionnaire are further means of cost reduction which merit investigation.

WHO COUNTS?

Information seems to be accumulating on the concentration of errors in survey reports among particular groups of families, especially among those families who have a lot of savings. Substantial reduction of survey errors is, therefore, most likely to be obtained through experimentation on means of securing more accurate and more complete data from these families. It may well be, for example, that special questionnaires are needed for different types of families in these groups, and that a different type of interviewer and a different type of approach would be more effective than the more or less standard approaches being currently used on all sample members in a single survey operation. To some extent, this may lead to increases in the cost of obtaining the data, but there would seem to be little doubt that effective techniques of this type would yield more reliable data per dollar spent.

What these techniques are, however, is a question which is also unanswered. Some work along this line is presently being carried out by the Consumer Savings Project.

SOURCES OF INFORMATION

USE OF SELECTIVE SEGMENTS

The evidence accumulated so far indicates that the variance due to response and nonresponse errors far exceeds any sampling variance that may be present in most data. At the same time, it seems clear that in probability sampling, sizeable proportions of those interviewed will either refuse to give financial information or will supply such information on a very incomplete basis.

In effect, the results of these surveys suggest that sampling errors constitute a relatively small portion of the errors encountered in financial surveys. It is evident, also, that procedures are badly needed for using techniques that will in one way or another take into account biases and nonsampling errors encountered in these surveys. A possible approach to dealing with such biases lies in what we may call the use of "selective segments." This is not the place to consider this approach in detail, particularly since it is still being investigated. In essence, however, it involves the stratification of a probability sample of consumers on the basis of their willingness to provide accurate and complete financial information. If a sample can be stratified in this manner, representation can be assured of all different segments in terms of the highly relevant characteristic, accuracy of report. Data obtained from each of the subsamples would be adjusted to correct for any biases known to exist in that particular stratum, and only then would the subsamples be recombined, and probably reweighted, to provide a picture of conditions in the entire population.

It is conceivable that in the process some or all of the subsamples will lose their probabilistic nature. Considering the relative magnitude of the errors involved, however, this would still undoubtedly increase substantially the accuracy of data on consumer savings.

Nonprobability samples have yielded accurate estimates of other aggregates at various times in the past. Study of the extent to which such samples might be adapted for the present purposes seems warranted, and concomitantly may reveal the extent to which nonprobability samples may possess virtues for estimation hitherto overlooked.

Concluding Comments

In view of the preceding discussion, what can we say about the prospects for improvement in the quality of consumer survey data? We cannot predict how soon improvements will come, or how extensive they will be; but we can predict that improvements are inevitable,

considering the need for such data and the increasing attention that is being given to this question.

It is also clear that the rapidity with which consumer survey data for social accounting is improved will depend largely on the amount of effort put into, and the results obtained from, methodological research. Specifically, it will depend on the progress that is obtained in improving the quality of consumer survey data and in reducing the cost of collecting data of a given quality.

The improvements that are effected in collection techniques for consumer survey data over the next few years may not be substantial enough to warrant substitution of consumer surveys for surveys of business or institutional sources if securing aggregate data is the only, or the principal, objective—especially considering the much lower costs of the alternative approaches. The fact is, however, that aggregates need rarely be the sole purpose of surveys of this type. The surveys can generally be designed not only to obtain data for aggregates and distributions but also to supply supplementary information on financial behavior. As time goes on, the latter objectives will in all likelihood tend to dictate the selection of the population to be surveyed.

C O M M E N T on Broida and on Ferber

JAMES N. MORGAN, University of Michigan

I should like to supplement my comments on these two papers with some ideas of my own as to the potential contributions of surveys to flow-of-funds analysis, taking the commentator's privilege of suggesting that the authors need not have restricted themselves as they did. Robert Ferber and Arthur Broida concentrate their attention on errors and potential sources of errors in financial magnitudes estimated from surveys, and assume that the main contribution of surveys to flow-of-funds analysis is in providing estimates of totals and distributions of assets and debts.

Both papers start with a theoretical and conjectural treatment of sources of errors. One then discusses some consequences of various error distributions, and then each paper provides some new data from recent validity studies.

As to the discussions of sources of errors in surveys, both authors begin by considering this problem in theoretical and systematic light. My own reaction is that their systems are nowhere near so appealing nor so useful as others which have been proposed. For those who are interested, I should recommend books by Hyman *et al.*, Moser, and Kahn and Cannell, particularly Chapter 7 of the book

Dynamics of Interviewing by the last two.¹ In particular, the distinction made by the papers under discussion between conscious and unconscious (accidental or purposeful) distortion does not strike me as being either operational nor very useful. On the other hand, a more detailed analysis of the positive and negative motivations of the respondent, and of the factors that may make some data more "accessible" to memory search than others might have been rewarding.

A number of particular statements are made which appear to me improbable; others, untestable.

The effectiveness of personal explanations can invariably be improved by increased expenditures for the selection, training, and supervision of interviewers.

In this case, the same author later avers:

. . . it is often difficult to say whether more elaborately written statements of survey objectives will increase or reduce cooperation.

Why should marginal productivity of additional resources put into interviewer training always be positive, but that put into designing written materials sometimes become negative? I might add as illustrative material that extra effort put into a booklet sent to half the sampling units in the 1958 Survey of Consumer Finances produced no appreciable change in either response rate, the financial information, or the frequency with which respondents consulted their financial records.

Preliminary indications are that aggressive interviewers are less likely to obtain accidentally inaccurate data than other interviewers.

This appears difficult to prove or disprove. What does it mean to conclude:

. . . outright refusal is less likely when rapport is good.

Or

. . . use of records . . . contribute[s] substantially to accuracy of reporting.

¹ Herbert Hyman *et al.*, *Interviewing for Social Research*, Chicago, 1954; Robert Kahn and Charles Cannell, *The Dynamics of Interviewing*, New York, 1957; C. A. Moser, *Survey Methods in Social Investigation*, New York, 1958.

(A more realistic statement would be that strong positive motivation of the respondent to cooperate is likely to lead to the use of records by the respondent.)

Even under some pressure to do so, only one in six looks up any records on income or income tax in the Survey of Consumer Finances; and only one in ten, any records on liquid assets. If anyone hopes to increase those fractions, it is necessary for him to answer the question: "Why should a respondent be so enthusiastic or interested as to go to such trouble? What appeals would induce even the most educated and research-minded among you to search for passbooks, etc.? Indeed, why should anyone believe that his checking account balance now in his check book is more important than its recent high, low, or average?" To do so requires understanding of and belief in the sampling properties of averages. Many a respondent's reaction is, "Why talk with me? I'm not average or representative at all."

Similar comments as to what causes what, might be made about the conclusion that those who continue to cooperate on a panel provide additional information on holdings of assets.²

Without describing the nature of the differences between his two questionnaires, Mr. Ferber concludes that unstructured questionnaires were more effective in securing rapport. "Unstructured" does not mean the same thing to everyone, probably. Does rapport mean better data? Do interviewer reports on success (Table 1) mean good data? We have findings from other validity studies which indicate that interviewer impressions of the accuracy of the data are frequently quite erroneous. Perhaps this is because some people don't like the interview but still tell the truth, while others may lie sweetly or merely misremember. I should like also to see some further support for the statement (page 358) by Mr. Broida that random errors which increase the variance will also increase estimates of inequality.

Finally, a number of statements, in Mr. Ferber's paper especially, imply that almost nothing is known about the effects of different appeals in motivating respondents, and that interviewers are generally selected almost at random. Actually, a great deal is known, and a substantial amount has even been published in scattered places. And while there may not be any good mechanical ways to select interviewers, it is easy to screen and retain only those whose performance is adequate. I might add one qualitative remark: much of the discussion deals with training interviewers, i.e. providing them with more information and training, when the main problem may well be

² See Marion Sobol, "Panel Mortality and Panel Bias," *Journal of the American Statistical Association*, March 1959.

motivating the interviewers as well as the respondents. The interview is an interaction between two persons, with a great deal of communication, where the enthusiasm and confidence of the interviewer mean more than her fluency or the pieces of paper she carries with her.

Mr. Broida indicates that data from the Surveys of Consumer Finances have not been used in flow-of-funds analysis because analysts felt that the “. . . quality . . . is not adequate.” He then proceeds to indicate a number of difficulties in assessing the adequacy of such data, including the skewness of the distributions and the existence of personal trust holdings of savings bonds. The impressive fact is that after many years and very large expenditures in assembling and analyzing both types of data, so little has been done to reconcile the two sources and types of estimates.³ Such reconciliation cannot be done at a desk, nor by validity studies of Surveys alone, but requires a separate set of reconciliation studies.

Let me turn now to the empirical findings reported. Those reported by Mr. Broida in particular I find interesting. He provides some hitherto unpublished data on discrepancies between borrowers' and lenders' reports on car debts from the 1956 National Analysts-Federal Reserve study of the financing of new-car purchases. Not only are the discrepancies small—though they are statistically significant—but the resulting biases in over-all estimates turn out to be substantially reduced by making simple weighting adjustments for differential nonresponse.

I feel compelled, however, to add a word of warning. In a survey which attempted to cover more material—and a major advantage of surveys is the possibility of studying interrelations—it might be that the discrepancies would be larger. Secondly, in a study like the Survey of Consumer Finances, where the nonresponse rate is around 16 per cent, instead of the 28 per cent in the new-car financing study, weighting for differential nonresponse probably does not make so much difference. Indeed, our investigations of the impact of the weighting we do in the Survey of Consumer Finances indicate substantial effects of weighting for different sampling rates, of course, but little apparent effect of weighting for differential rates of nonresponse. This is so in spite of the fact that a somewhat more complex scheme is used in weighting the S.C.F. data.⁴ One explanation, in addition to the smaller amount of nonresponse, might be that in a many-faceted study, nonresponse might result less from the particular

³ Except, of course, by Selma F. Goldsmith, in presenting annual data on the distribution of income by size (see *Survey of Current Business*, April 1959, for the last in this series).

⁴ See John Lansing and Thomas Eapen, “Dealing with Missing Information in Sample Surveys,” *Journal of Marketing*, October 1959, pp. 21–27.

subject matter than from the personality of the respondent and interviewer and the situation of the moment.

The effect of weighting for differential nonresponse on some means in the Survey of Consumer Finance is as follows:⁵

	Weighted	Unweighted
Mean net outlay for		
All cars	\$1,112	\$1,110
New cars	2,122	2,126
Used cars	659	658
Mean debt incurred by credit buyers of		
New cars	1,922	1,920
Used cars	704	702

Let me now say a few words about the general topic of these papers—the potential contribution of sample surveys to flow-of-funds analysis. The potential contribution will not come from either better estimates of aggregates, nor even of their distributions by age groups by size, etc. The main contribution of sample surveys lies in two directions:

1. Interrelations among items within the same family, i.e. between income and assets, debt and purchases, total assets and ownership of each component. Are those individuals who account for one flow (into common stocks, for instance) the same as those accounting for another (out of United States bonds)?
2. Inferences as to why decisions affecting flows of funds are made the way they are. Such inferences come both from the interrelationships above, and from data on the information and attitudes of respondents. In some cases they may even come from simply asking, "Why?"

If we want to know more about who is behind the flows we study, and why these people behave the way they do, then surveys are an essential tool. Instead of talking about the behavior of aggregates, we can talk explicitly about the behavior of individuals as it affects aggregates. Instead of inferring purpose from the form in which a flow appears in the aggregate, we can infer it more directly from the

⁵ The 1959 Survey utilized a uniform combined sampling fraction; so weighting was entirely an adjustment for differential nonresponse. However, a few families where the related secondary could not be interviewed were completed by substituting a secondary from a matched family, and these thirty-five units were left in the spending unit decks as well. Hence some delicate adjustments for nonresponse had already been made even before the differential weighting was imposed.

pattern of action by individuals and from what people directly responsible tell us.⁶

The papers at this conference are rich with statements of the "should" or "may" or "might" variety. Upon examination, many of them prove to be hypotheses about the behavior and even the motivation of businessmen, bankers, or consumers. The great contribution that surveys could make to the flow-of-funds analysis is in selecting among alternative explanations and in providing a motivational interpretation of the "behavior" of the flows.

If, since the war, there has been a shift in interest rates, and a shift also out of demand deposits into time deposits and securities, who has done the shifting? Has it been a few large holders or many small ones? Has it even been the same people? Perhaps the dis-savers have used up demand deposits, and the savers have accumulated time deposits and securities. The interest rate changes may have been causal in either case, but certainly it is important to know just how they have operated.

And if we discover just who in the population is responsible for certain shifts, isn't it important to know:

- what alternatives were really open to these people
- what alternatives they were aware of among those which were open to them (information, insight)
- what else they were doing at the same time
- what were their sources of information and advice
- what else do we know about them from which we might infer why they did what they did (age, unemployment experience, attitudes, contractual commitments)

It is also my impression that most of the hypotheses about motivation and behavior being made at this conference have to do with the behavior of people with substantial assets and/or incomes. One of the problems, then, is how a survey, which usually involves interviewing a lot of poor people, can be of use. We are currently working with the Federal Reserve Board on a project to concentrate a survey among high-income, high-asset families (with a few others for comparison). It behooves those interested in either national income analysis or flow of funds, to see whether such a survey couldn't answer some of their questions about which behavior assumption is best.

Another substantial group of hypotheses about behavior deal with behavior that is quite rare, i.e. engaged in by only a very small part of the population. The small part may not even necessarily be those with the highest incomes. In this case, I like the suggestion of Irwin

⁶ See George Katona, *The Powerful Consumer*, New York, 1960.

Friend that we sample the financial records, find a sample of those responsible for the account or the transaction in which we are interested, and interview these people. If we have doubts about bank reports on the proportion of their savings accounts owned by businessmen, farmers, and ordinary consumers, why not visit a small, carefully selected sample and find what proportion are engaged in farming, have a business, etc. One need not even know the size of the account: if the sample is drawn proportionate to the size of the account, then 10 per cent of the addresses represent 10 per cent of all savings accounts. Some gains in insight and efficiency are possible if one has information from bank or respondent about the size of his account, but the possibilities of real, insightful information without violating any confidences are great. This differs from Mr. Friend's approach in that he suggests trying to cover all assets this way, and all types of financial institutions, asking only the most rudimentary things from respondents, and ending up with simple distributions of financial magnitudes by age, occupation, income, etc., individually. I'm suggesting use of financial institutions as sample sources in order to (1) pinpoint the sample on the most interesting groups and (2) have one correct financial figure, at least, on a sample basis. If we are to go to the expense of interviewing people, we can find out a great deal about them that will throw light on why the financial aggregates have been moving as they have.

Further, it should be noted that even if the respondents' reports on their assets are erroneous, if one has a sample drawn from asset records, the only thing one needs with any accuracy from the respondent is the *other* information to which we want to relate, i.e. demographic background, attitudes, information level, explanations (which may or may not be rationalizations), and some other rough financial information. It is, after all, the interrelationships among these variables which provide us the maximum insight, rather than simple distributional data.

In conclusion, I suggest, first, that we think more broadly about the possible sample designs which might be most efficient for the purposes at hand and, secondly, that we concentrate the content of surveys where they are most likely to prove fruitful: combinations of assets or actions; rough measures of dollar amounts; attitudes and information; and the relations of demographic facts, attitudes, and knowledge to economic behavior. We may be able to make minor improvements in the actual estimates in the flow-of-funds data, and obtain some approximate information about their distributions; but the major contribution of surveys must be in understanding the *meaning* of the flow-of-funds data.

