Individual Diaries and Expense Documents in the Italian Consumer Expenditure Survey

Carlo Filippucci and Maria Rosaria Ferrante[1]

This article deals with the effectiveness of the diary in the Italian consumer expenditure survey. Two experiments are presented: one using individual diaries and the other the compilation of a household diary based on expense documents. Households had no problem in dealing with individual diaries even though their effectiveness clearly depends on household socio-economic level. Individual diaries allowed us to find underreporting, especially for some important items. On the contrary, expense documents were not found to be effective in recording, but the experiment proved useful for understanding recording behaviour. The results give some suggestions for improving the Italian survey.

Key words: Data collection instruments; underreporting; proxy response; aided recall.

1. Introduction

Underreporting is a source of concern in the Italian Consumer Expenditure Survey (ICE) as it is in many other countries. In Italy national account estimates of consumption are on the average about 20 per cent higher than those of ICE and the difference rises to 40 per cent for some items. The degree of underreporting is directly related to the adequacy and functionality of data collection instruments with regard to the phenomenon under observation. Based on this consideration, this article deals with the effectiveness of using diaries in the ICE. Specifically, two experiments are presented: one using individual diaries and the other the compilation of a household diary on the basis of expense documents (receipts, slips, etc.).

The diary is used for stimulating and facilitating expense recording and can be addressed either to households or to their members. The household diary is justified by the indivisibility of certain expenses, by limiting the number of household members to be interviewed, and by costs, which increase with the sample size (Sudman and Ferber 1971). However, the household diary can lead to underreporting because: i) household members neglect to give some information out of reservedness, forgetfulness or distraction, since they are not directly involved; ii) the proxy respondent has his or her own view which does not necessarily match the actual behaviour of household members (Silberstein and

Scott 1991). Consequences of these effects are contradictory (Moore 1988; Blair et al. 1991).

Individual diaries for each member able to keep records help avoid the problems we mentioned. This solution, adopted by some statistical agencies (Innocenzi 1992) is also the most consistent with the aim of providing data for studying individual behaviour. Studies on the subject (Kemsley and Nicholson 1960; Sudman and Ferber 1971; Grootaert 1986) suggest that individual diaries are effective as far as the completeness of recording is concerned, especially for personal expenses. However, problems of cooperation arise on the part of households or their members. Grootaert relates that the rate of cooperation depends significantly on household characteristics.

In consumer expenditure surveys the respondent's performance is conditioned by his or her capacity to remember, which, in turn, depends partially on the context of the survey (Lessler 1989). The diary recall period is short, but since events are very frequent, but not very notable, they are easily forgotten. Methods generally used to aid memory, like psychological cues are not sufficient for collecting current expenditures, so solutions must be found to facilitate the recording of events as they happen. Expense documents seem to be a useful device to aid diary keeping.

In the second section of this article two experiments are presented. The results of the experiments appear in Section 3. Finally, conclusions and some suggestions for improving ICE are presented.

2. The Diary in the Italian Consumer Expenditure Survey: Two Experiments

ICE is based on a two-stage sample with stratification of the primary sampling units (psu's) (municipalities) in each region and systematic selection of second-stage units (ssu's) (38,500 households each year). Some psu's are self-representative (SR). Final estimates are based on poststratification according to household size. Three different questionnaires are used: i) a household ten-day diary for current expenditures; ii) a household ten-day diary for self-consumptions (i.e., consumptions not purchased); iii) a final interview questionnaire, concerning household characteristics and non-current expenditures. The diary (i) is semi-open and includes 23 expense categories and about 150 purchase items are listed. It must be compiled by one household member, who reports the expenditures and, in the case of food, quantities as well.

In the first experiment each household member, 14 years or older, was asked to keep records in a personal diary which had the same characteristics as the household diary used by Istituto Nazionale di Statistica (Istat, (i)). This solution was adopted essentially to isolate the effect of individual reporting.

By relying on individual diaries, the aim of this experiment was to evaluate how and to what extent proxy responses influence the number and the amount of expenditures reported. The experiment also allowed us to evaluate the degree of household participation and the capacity of individuals to deal with a more onerous survey.

In the second experiment households were asked to compile the diary referring to documents attesting to the purchases made. This procedure was implemented in order to improve recall, recording, and respondent reliability. Households were invited to collect and keep receipts, invoices or any other documents in special envelopes (one for each

day of the survey). Whenever an expense document was not available, households were asked to record the purchase using a document facsimile.

The experiments were limited to one region (Emilia-Romagna) and both referred to half the region. They were concurrent with the survey being carried out in November 1990, which constituted the control sample of the experiment.

The individual diary experiment involved six SR municipalities and three non-SR. In total 65 experimental households and 74 control households were observed. The actual sample size was 77, including single-member households, but this kind of household had to be excluded reducing the actual sample size.

The experiment of expense documents, in turn, involved four SR municipalities and three non-SR, in total 81 experimental households and 80 controls. The two experimental samples were independent of the control samples.

The decision to restrict the experiment to one region due to problems of time and resources, led us to observe a limited number of households (Filippucci and Marliani 1992). For this reason generalization of results should be done with caution. For each experiment the Istat sample was considered as control. The experimental sample had an equivalent size and was selected from the same municipalities. The experiment made use of the field organization adopted for the ordinary survey. In particular, the same interviewers were used and incentives were given neither to the interviewers nor to the households.

Since the experiment burden could have led to a high degree of partial nonresponse (Sudman and Ferber 1971; Tucker 1992), the possible influence of missing data in the samples had to be assessed. Some indicators of data quality were calculated for the experimental and control samples by matching the non-edited file with the edited one (automated editing) for each household. This strategy was proposed in Italy by Masselli (Istat 1989). A cross-classification of each sample unit was carried out according to the values it assumed in the two files. On this basis, indicators on the quality of the collected data, and on incoherent responses and refusals were calculated. In general, for the four samples the indicators differ only slightly from optimum values.

One final remark concerns the estimation procedures. Each experiment dealt with only half the region and one month of the quarter of the survey; therefore the sample is smaller than usual. Hence, due to the dimensions of the observed samples, it was necessary to collapse some strata. In practice this would have led to stratifying in two strata only, SR and non-SR municipalities, which would have greatly reduced the effectiveness of the adjustments. Besides, because of the dimension of the samples only one psu for non-SR strata was found, preventing us from calculating variability.

Because of these difficulties the comparability was checked by comparing households' socio-demographic profiles stemming from experimental and control samples. For this purpose multiple correspondence analysis was applied to the following variables: (i) household: number of members, monthly income, home ownership, possession of durables; (ii) both the head of household and the spouse: sex, age, education, working conditions, and professional position. If statistical units belonging to the experimental or control samples do not determine the description of principal axes, expenditure differences observed in the two samples can be attributed to the effect of the experiments. This condition was fulfilled. Several of the household variables contribute to the axes characterization. The professional position of the

head of household and the possession of durable goods can be eliminated, since these factors show little variation. The same results were obtained for two subsets of randomly selected individuals. Disturbances did not influence the above findings, confirming the stability of the factors (Lebart, Morineau, and Tabard 1977). In summary, the results show that the socio-demographic profiles which characterize experimental and control samples in the two experiments are similar.

3. The Results of Experiments

3.1. The effects of individual diaries

First of all it is important to stress that recording more diaries in each household did not produce a "fatigue effect." In 95.4 per cent of the households each eligible member recorded expenses in the individual diary (and 98 per cent of the eligible members). Looking at the individual diary efficacy in reporting, results did not completely meet our expectations. In fact, they are not statistically significant. Nevertheless the ten-day household expenditures and the number of records are higher when the individual diary is used (Table 1). It should be stressed that both household expenses and the number of records from individual diaries give the same indication and the expenditure and the number of records observed in the control sample used for the second experiment are even lower than the control sample utilized for the individual diary. This last aspect is an indirect further indication supporting the statement that individual diaries lead to reporting a greater number of records and amount of expenditures.

Last, but not least, the individual diary improves reporting for various important items. Contrary to what has been shown before, the number of non-food records is significantly greater and in particular is noticeably higher for "Newspapers, stationery, and education," "Transportation," and "Post and public telephones." Considering food, it must be noted that spending on "Meals and snacks away from home" is almost double compared to the control sample. It should be stressed that such expenditures are generally made by the head of the household.

Additional findings emerge from examining the number of records: the *t* test is generally not significant with the sole exception of "Meat," which is higher in the control sample. The last finding can be explained by taking into account that in the ordinary survey the diary keeper can easily be induced to group together the same type of purchases made

Table 1.   Ten-day household expenditure and number of records

	n	Number of records		Expenditure 000 (Lit.)*	
		mean	std dev	mean	std dev
Individual diary	65	90	38.8	1,033	947
Control	74	85	37.7	995	714
		$t = 0.75$ $df = 137$	Pr = 0.4570	$t = 0.27$ $df = 137$	Pr = 0.7896

*Lit. 10,000 are equivalent to about 6.5 U.S. dollars (August, 1997).

at different times and/or by different members of the household. In fact, grouping expenses leads to distortion of spending pattern. Hence, we must not disregard the information stemming from the "number of records" and it is relevant to our task that reports from individual diaries present a higher number of records than household diaries do.

These findings are confirmed when household members are considered separately. Spouses recorded a much higher number of expenses than the other members (almost double compared to heads of household and triple compared to children), but they spend a bit less than heads of household. Besides, for all the above listed items showing a higher average number of records in the experimental sample, the head of household and, frequently, the children record a higher amount and number of expenditures compared to the spouse. This finding shows that some typical personal purchases are easily captured using individual diary.

It can be concluded then that for certain expenditures, non-foods in particular, surveying by individual diaries can reduce the underreporting attributable to the proxy responses.

3.2.   The effects of expense documents

Recording aided by expense documents did not disclose underreporting or missed recording entries enough to affect the estimation of total expenditure. Moreover, relying on this experiment produced a negative effect on both the number and the amount of expenditures reported (Table 2). In the various expense categories the findings did not differ from what was seen for aggregate expenses.

Although the efficacy of expense documents in reporting is not supported by data, the experiment is still helpful for understanding respondent behaviour and for improving the survey. Firstly, the lower number of records in the experiment concerns mainly goods for which it is presumably more common to lose or not receive any document. In fact, if we compare the information obtained by documents, with that recorded in the diary (only in the experiment), we found that, on the average, 70 per cent of the records match a document, even though households were expressly invited to collect the documents or to keep note of each undocumented expenditure. At disaggregated level, the relationship between the average number of records in the diary and the average number of those documented was 76 per cent for food expenses, and the ratio decreased to 55 per cent for non-foods. The lowest values of the ratio were, notably: "newspapers, writing supplies, and education," and "transportation," and "amusement, tourism, and entertainment" as well as the obvious ones of "tobacco," "public telephones and postal expenses," and "children's allowances,"

Table 2.   *Ten-day household expenditure and number of records*

		Number of records		Expenditure 000 (Lit.)*	
	n	mean	std dev	mean	std dev
Expense documents	81	56	29.7	607	645
Control	80	67	28.9	813	748
		$t = -2.38\,df = 159$ Pr $= 0.0182$		$t = -1.87\,df = 159$ Pr $= 0.0635$	

*Lit. 10,000 are equivalent to about 6.5 U.S. dollars (August, 1997).

which did not reach 30 per cent. Considering these results, we feel that respondents took only the available documents into account even if they did not cover all the purchases made.

A second aspect can be traced to the difficulty which certain respondents have in reading and understanding documents, and taking advantage of them to fill in the diary. Most of the documents collected are cash register receipts, mostly from small shops (66 per cent) and the total amount of expenditure documented is usually rather low. In fact 63 per cent of the documents report an expenditure no greater than Lit. 10,000. In general, it was found that the majority of documents (60 per cent) contains only the amount spent, but no useful information as to the category. Since the type of goods purchased can rarely be identified through expense documents, they prove to be of little use as a support for recording. Using receipts of purchases made in supermarkets and department stores was not helpful either; in this case the great number of items listed probably discouraged recording.

It seems that the diary keeper, because of the burden of recording, makes a subjective evaluation about what and how much to record. In this way the respondents simplify their duties and reduce the respondent burden. Therefore, as in our experiment, when respondents are requested to refer to precise documentation and then to report exact values, the burden is not accepted and they disregard some receipts.

3.3   The effect of the experiments on spending profiles

In this section all the expense categories (22) are considered simultaneously, making it possible to study the effect of experiments on the "household's spending profile."

For this purpose nonparametric discriminant analysis (Hand 1982) was implemented separately for the two experiments comparing experimental and control samples. Specifically, if the experimented instrument influences the households' spending profiles, households will be attributed to the samples to which they effectively belonged at the time of the experiment (experimental and control). Hence, a small number of misclassified expresses a strong relationship between the instrument used and the spending profile.

The percentages of misclassified units, identified by the "leave-one-out" method, (Lachenbruch and Michey 1968), are 36 per cent and 25 per cent, in the case of individual diaries and of expense documents, respectively. Most households therefore appear to be affected by the instrument adopted.

As the number of misclassified should not be ignored, it is useful to understand whether those units which are not affected by the adopted instrument present any distinguishing characteristics. In particular, we checked if, dealing with data collection instruments, a household's behaviour is conditioned by its socio-economic level.

Comparison of the average monthly spending (per capita) between the misclassified and non-misclassified groups was performed separately for the two experiments. Non-misclassified expenditure was higher in both experiments but it was more apparent in the case of individual diaries. Furthermore it was found that households belonging to the two samples are those characterized by a lower average spending level. This could simply reveal that the households with a low spending level (which could also indicate households which do not record carefully) are also those less affected by the instrument

adopted. In these cases better information could only be obtained by relying on a more accurate surveying strategy.

3.4.   The interaction between instruments and household characteristics

So far we have considered only the effects of instruments on the entire sample. However, efficacy of questionnaires is not independent from respondents' characteristics. More information can be gained by examining the interaction between instrument and household type.

The available information for checking this aspect is the number of records and the amount of expenditures. We suggest evaluating the importance of errors associated with different surveying techniques by assuming that the greater the number of records and the higher the amount of expenditure, the more reliable the records.

Two-factor analysis of variance was applied to the number of records and expenditure, considering separately each of the household characteristics listed in Section 2. This means that each characteristic in turn represented the first factor, while the experimental-control variable constituted the second one.

The analysis confirmed that household characteristics, in particular the socio-economic level, substantially affected the quality of records in both experiments, even though in opposite ways. This aspect was even more evident when the interaction of the two factors was considered. Using individual diaries, the lowest number of records was reported by households of high socio-economic level. On the contrary, diaries improved the recording of households of low social level, unless the spouse was also employed in a medium-low position. Using expense documents proved more efficient when the head of household had a medium-high level of education or was self-employed or employed at a managerial level.

4.   Some Conclusions

The data analysis supplied us with various concordant pieces of evidence which can be useful in improving data collection, in particular in ICE.

No evidence emerged in the experiment with expense documents regarding their effectiveness as an aid to recording. The results, however, are on the whole useful for understanding some problems of the current survey and for taking corrective measures. The very low ratio of documented expenditures to recorded ones, indicates that fiscal receipts in Italy are anything but systematic. The discrepancy is a worrisome signal of how households receive and carry out the task assigned to them. It seems, in fact, that the households' collaboration, seemingly high, is actually subordinate to a sort of tacit agreement with the interviewers: the households collaborate as long as they are not required to be accurate. While they report the spending event, they do not worry too much about remembering the actual amount spent. Then, when asked to document the expenditure, they prefer not even to record the event approximately. Although this can partially be explained by the difficulty in tracing back to the specific purchases from receipts, it does not reduce the scope of the problem. If the above interpretation is correct, this means that aids to recall are not sufficient for reducing underreporting of current expenditures. Hence the problem has to be dealt with through a more general improvement of survey instruments and procedures.

As far as the experiment of the individual diaries is concerned, one result in particular must be stressed: the households showed no particular problem in dealing with the individual diaries. Neither the quality nor the number of records was reduced, nor did we find any refusals or complaints. This indicates that the instrument used did not encounter any substantial obstacles. However, using individual diaries is useful since it leads to improving expenditure estimates, as evidenced by various results of the experiments. First of all, the household expenditure for consumption reported by individual diaries was slightly higher. Secondly, the expenditure for some important items was underestimated when the household diary was used instead of individual ones. Thirdly, the average number of records for non-food items is significantly greater than those reported through the household diary. Finally, it was found that, although it is true that besides a principal purchaser in the households (the housewife), there are other important actors as well: the head of the household and at least one child. The former in particular makes most of the relevant expenses so his or her role is not negligible in the survey.

These results constitute important clues to the effectiveness of the individual diary in reducing underreporting. As a consequence, it seems highly advisable that all household members be involved in the survey, although we advise maintaining one main reference person in each household. For this reason we suggest an easy-to-use personal notebook, with simple and clear references to the items to be recorded, easy to carry around and whose attractiveness makes it inviting to use.

In these conclusions we have not yet taken into account one important aspect which was obvious in both experiments: the bulk of evidence showing a noticeable effect of socioeconomic and demographic characteristics on households' reactions to the surveys. This means that we should introduce more versions of the diary or the questionnaire. Obviously, from a practical point of view this approach presents some difficulties.

In conclusion, this study indicates that it is crucial to develop strategies for obtaining a better collaboration from households. This is not the only aspect of the problem: instruments used and data collection procedures are also relevant and have different effects on different types of households. Moreover, based on other studies on ICE (Filippucci and Zanelli 1993), it should be added that more attention needs to be paid to the interaction between interviewer and respondent. From the study cited above it can be clearly seen that the success of any instrument relies on the interviewer's behaviour and on his/her capacity to establish a good relationship with the households.

Each survey constitutes a case in itself where techniques, modalities and instruments must be reconciled with the subjects and the phenomenon to be investigated. Especially in measuring expenditures, the practical survey implementation has a crucial role. The success of a survey, like for any commodity, depends on its quality and cost, but especially on the producer's ability to pinpoint its specific recipient, communicating with him or her by means of adequate language and instruments.

5.   References

Blair, J., Menon, G., and Bickart, B. (1991). 'Measurement Effects in Self vs Proxy Response to Survey Questions: An Information-Processing Perspective. In Measurement Errors in Surveys, P. Biemer et al. (eds.). New York, John Wiley.

Filippucci, C. and Marliani, G. (1992). La misura dei consumi delle famiglie: una riflessione a partire dall'esperienza italiana. Conpri, Report n. 6, Bologna, Dip. Scienze Statistiche. [In Italian].

Filippucci, C. and Zanelli, E. (1993). Un esperimento di valutazione cognitiva del questionario nell'indagine sui consumi. Conpri, Report n. 16, Bologna, Dip. Scienze Statistiche. [In Italian].

Grootaert, C. (1986). The Use of Multiple Diaries in a Household Expenditure Survey in Hong Kong. Journal of the American Statistical Association, 81, 938–944.

Hand, D.J. (1982). Kernel Discriminant Analysis. New York, Research Studies Press.

Innocenzi, G. (1992). Principali aspetti dell'indagine ISTAT sui consumi delle famiglie. Conpri, Report n. 2, Bologna, Dip. Scienze Statistiche. [In Italian].

Istat (1989). Il sistema di controllo della qualità dei dati. Manuale di tecniche d'indagine. Note e relazioni. [In Italian].

Kemsley, W.F.F. and Nicholson, J.L. (1960). Some Experiments in Methods of Conducting Family Expenditures Surveys. Journal of the Royal Statistical Society, Series A, 307–328.

Lachenbruch, P.A. and Michey, M.R. (1968). Estimation of Error Rates in Discriminant Analysis. Technometrics, 10, 1–11.

Lébart, L., Morineau, A., and Tabard, M. (1977). Techniques de la description statistique. Dunod, Paris. [In French].

Lessler, J.T. (1989). Reduction of Memory Errors in Survey Research: A Research Agenda. Bulletin of the International Statistical Institute, 47th Session, Paris.

Moore, J.C. (1988). Self/Proxy Response Status and Survey Response Quality. A Review of the Literature. Journal of Official Statistics, 4, 155–172.

Silberstein, A.R. and Scott, S. (1991). Expenditure Diary Surveys and Their Associated Errors. In Measurement Errors in Surveys, P. Biemer et al. (eds.). New York, John Wiley.

Sudman, S. and Ferber, R. (1971). Experiments in Obtaining Consumer Expenditures by Diary Methods. Journal of the American Statistical Association, 66, 725–735.

Tucker, C. (1992). The Estimation of Instruments Effects on Data Quality in the CE Diary Survey. Journal of Official Statistics, 8, 41–61.