THE FAMILY EXPENDITURE SURVEY

- **An Experiment with Incentives**
  Håkan L. Lindström
- **Seasonal Variation and Response Behaviour in Swedish Households Expenditure**
  Peter Lundquist
- **Reducing Nonresponse Rates in Family Expenditure Surveys by Forming Ad Hoc Task Forces**
  Lars Lyberg
- **A Study of Errors in Swedish Consumption Data**
  Martin G. Ribe

SCB

R&D Report
Statistics Sweden
Research - Methods - Development
1991:10

INLEDNING

TILL

R & D report : research, methods, development / Statistics Sweden. – Stockholm : Statistiska centralbyrån, 1988-2004. – Nr. 1988:1-2004:2.
Häri ingår Abstracts : sammanfattningar av metodrapporter från SCB med egen numrering.

Föregångare:

Metodinformation : preliminär rapport från Statistiska centralbyrån. – Stockholm : Statistiska centralbyrån. – 1984-1986. – Nr 1984:1-1986:8.

U/ADB / Statistics Sweden. – Stockholm : Statistiska centralbyrån, 1986-1987. – Nr E24-E26

R & D report : research, methods, development, U/STM / Statistics Sweden. – Stockholm : Statistiska centralbyrån, 1987. – Nr 29-41.

Efterföljare:

Research and development : methodology reports from Statistics Sweden. – Stockholm : Statistiska centralbyrån. – 2006-. – Nr 2006:1-.

R & D Report 1991:10. The Family Expenditure Survey / Håkan L. Lindström m.fl.

Digitaliserad av Statistiska centralbyrån (SCB) 2016.

urn:nbn:se:scb-1991-X101OP9110

THE FAMILY EXPENDITURE SURVEY

- **An Experiment with Incentives**
  Håkan L. Lindström

- **Seasonal Variation and Response Behaviour in Swedish Households Expenditure**
  Peter Lundquist

- **Reducing Nonresponse Rates in Family Expenditure Surveys by Forming Ad Hoc Task Forces**
  Lars Lyberg

- **A Study of Errors in Swedish Consumption Data**
  Martin G. Ribe

SCB

R&D Report
Statistics Sweden
Research - Methods - Development
1991:10

Från trycket Augusti 1991
Producent Statistiska centralbyrån, utvecklingsavdelningen
Ansvarig utgivare Åke Lönnqvist
Förfrågningar Håkan Lindström, 019-17 66 94
Peter Lundquist, 08-783 49 18
Lars Lyberg, 08-783 41 79
Martin G. Ribe 08-783 48 54

ISSN 0283-8680
Printed in Sweden

The Family Expenditure Survey

An Experiment with Incentives

Håkan L. Lindström, U/STM

Seasonal Variation and Response Behaviour in Swedish Households Expenditure

Peter Lundquist, U/STM

Reducing Nonresponse Rates in Family Expenditure Surveys by Forming Ad Hoc Task Forces

Lars Lyberg, U/SFI

A Study of Errors in Swedish Consumption Data

Martin G. Ribe, I/MET

Preface

A **Workshop on Diary Surveys** was organized by Statistics Sweden in Stockholm 18-20 February 1991. Participants from many countries attended the meeting. The purpose of the Workshop was to meet colleagues from other countries and to exchange experiences about diary surveys.

In this report four papers from the Workshop about the Swedish Family Expenditure Survey are presented. Besides of this report, **A Methodological Review** written by Hans Näsholm, is also soon to be published in the R&D Report series. Unfortunately, it is not possible to reprint all the participants' papers. Those who are interested in the additional papers are referred to the authors of this report or to the staff at Statistics Sweden, who works with the Family Expenditure Survey.

Contents

An Experiment with Incentives

Håkan L. Lindström

1 Increased nonresponserate in 1988 Family Expenditure Survey

The 1985 and 1988 Swedish Family Expenditure Surveys used a combination of data collection methods. Individuals are drawn in a simple random sample without replacement. When an individual is drawn his/her family is included in the sample. The families are in that way chosen with a probability almost proportional to the number of members. The sample was split into 26 independent subsamples of equal size. Each subsample should start to keep diaries on a certain day and continue during four weeks. There is one starting day each second week. Before the diary period an introductory interview should be performed to establish family composition, housing conditions etc and some major expenditures. It should be a face to face interview but in some cases telephone interviewing had to be used. Interview and diary data were later to be matched with register information, mainly on income.

In the 1985 Family Expenditure Survey (FEX) the response rate was 72 per cent in the interview and diary part of the survey. As response rates had gone down 2-3 per cent in face to face and telephone interview surveys since then, we expected to reach only 70 per cent response rate in the 1988 survey.

After a good start of the data collection, very soon reports from the interviewers told us that the response rate was seriously in danger to fall far below the level we had expected. This was soon confirmed by the overviews the interviewing agency regularly puts together. In march 1988 the response rate was below 55 per cent. In May 1988 only 45 per cent of the sample had completed their diaries. This was 16 per cent below the comparable percentage in the 1985 Family Expenditure Survey.

The drop in response rate was a big surprise as the 1988 FEX was designed almost exactly as the 1985 FEX. There was no change in the design to explain for such a difference. Data was in both cases collected by the same combination of interviews, diaries and register information. In the 1985 FEX there was also a final interview done in the first month of 1986 so the respondents burden had not increased. The staff was mainly the same and about 60 per cent of the interviewers had experience from the 1985 survey.

2 Analysis of the reasons for increased nonresponse and the actions to reduce it

When it was obvious that the nonresponse rate in the 1988 FEX wouldn't improve without special efforts, it was decided to form an emergency task force. The members should analyze the distribution of nonresponse, try to find the reasons of increased nonresponse, suggest remedies and put them into action Lindström (1989).

Analysis of the distribution of nonrespondents up to May 1988, demonstrated that, within the limits of sampling variation, the increase compared with the 1985 survey was about the same in those main domains of study, that could be formed by use of the information in the register of the total population. Still there were a couple of differences to notice. The nonresponse increase was specially high in the Stockholm area and specially low among young households (head below 24).

We made three major conclusions:

(i) The interviewing agency must put aside more resources for this particular survey.
(ii) The increase of nonresponse was special to this survey and we could not blame it on a general change for the worse of survey climate.
(iii) The increase of nonresponse was on the whole equally distributed. General measures should be adopted and not measures applied to particular groups of the sample.

A number of extra efforts were suggested. Among these were:

* revision of the information to interviewers and families in order to adapt it to the current field experience.
* more contacts and encouragement to the interviewers from the central staff.
* more efforts of the central staff to give top priority to FEX
* delay of other surveys.
* temporary recruitment of three retired interviewers.
* concentration of work and reduction of the number of interviewers who worked on the survey.
* use of extra incentives

The rest of this report will concentrate on the last mentioned of the extra efforts - the extra incentives.

Anyhow a few words should be said on one main reason for nonresponse. Our conclusion was that the heavy workload both centrally and on the interviewers together with recruitment problems in the interviewing agency were an important explanations for the trouble. We believed that first of all a study of the organization of interviewing would provide the explanations. One important part of our work was to find more interviewer resources. We also found that one important reason for the increase of nonresponse was that the interviewers were overworked

during the data collection period. There had also been a large turnover the last few years among the interviewers specially in our three largest cities Stockholm, Göteborg and Malmö. We found that experienced interviewers reached about five per cent higher response rate than their unexperienced colleges. But there were also great individual variations in performance in these two groups.

There were also some indications of a changing attitude to private economy in the population. The interviewers told us that they were not as successful as before when they tried to persuade the sampled families that it would be in their own interest to se from their diaries how they spent their money. Many of them seemed to think that they had money enough and didn't care about the information they could get.

3 Choice of incentives

The Swedish FEX surveys have, since 1958, given some kind of compensation to the families that take part in the survey. The value of the compensation has gone down for each survey. One reason for this, was that according to our experience, it was the existence and not the level of the compensation was not very important. That is, if the value was not very high. There was also the problem that the families would have to pay tax already if the compensation was above 100 SEK and the need to reduce survey costs.

In 1985 the compensation was, some stationary, a lottery ticket and a subscription for a consumers journal. The cost was about 60 SEK in current prices for each responding household. The compensation was given partly when the family had agreed to participate, partly when it had fulfilled the data collection. As the response rate in the 1985 FEX was on *normal level* for Swedish diary surveys it was believed that the same level of compensation would do well also in the 1988 FEX.

Among the suggested methods to improve the response rate was the idea of prepaid incentives. This procedure had been used successfully before by others, but at the planning stage, our experience together with survey information did not tell us that it would be necessary to use. Berk et. al (1987) reports one experiment giving a ten to fourteen per cent increase in response rate with *incentives in advance* and tells about similar results in a couple of other surveys. When the nonresponse rate increased and we had to reconsider our methods it was apparent that these experiences should be tried. Although the results referred to interview surveys there was nothing to say that prepaid incentives would have entirely different effect in diary surveys.

What kind of *extra incentive* we should chose was not quite obvious. The incentive could not be too expensive as there were only small extra resources available. It should also be at hand pretty soon as we were in a hurry to improve the response rate of the 1988 FEX. There was no time for experiments. We wanted the incentive to be received as a pure gift, but still within the context of subject of the

survey. i.e. be useful if people liked to do some calculations of their expenses. After some discussion we decided for a pocket calculator with the logo of Statistics Sweden printed on its front. Its price was less then 20 SEK apiece. Although the price for each calculator is low there are also costs to distribute the incentive. We must also have in mind that the use of the incentive will make only about one of ten sampled families respond. This means that the incentive cost is at least 200 SEK for each of the *additional responding families*. The cost to distribute the calculator was fairly low as the calculator was sent to the families in the sample together with the introductory letter.

Although nonresponse reduction was the primary aim for our decision to use the incentive we believed that it would also make the job easier for the interviewers and increase the quality of the expenditure reporting. There are a couple of Swedish experiments demonstrating better response quality together with higher response rate Lindström (1985). Berk et al. (1987) are of the same opinion, saying: "However the net added cost may be far less than the value of the incentive payments, since a substantial part of the incentive costs is offset by savings in the follow-up activities." However we did not have time and resources to analyze such effects.

4 Planning an experiment

As the conclusion of the earlier experiments was that use of *incentives in advance* was favorable to the response rate and the evidence fully convincing, there was no reason to believe that the use of calculators as incentive would be anything but successful to increase the response rate. We did not hesitate to use it as soon as it was possible to buy and distribute the calculators. The first subsample to receive it was number 15 and after that all subsamples got it. The interviewers where also given a supply of calculators but only to use when they tried to persuade families who already had refused.

As use of incentives is expensive we were eager to establish how much the gift would influence the response rate of the 1988 FEX. As the data collection was spread evenly on all the year it was easy to build an experiment into the survey. It was obvious that taking away a number of families to the *control group - without calculators* - would reduce the overall response rate. We were however determined to accept this in order to make an estimate of the incentives effect.

The test should obviously be onesided. Calculations demonstrated that the sample size in each of experiment and control group should be about 500 observations for a observed difference of about five per cent to be significant. This size of the difference was the smallest we hoped for. If the difference was smaller it would be less important to establish its level. As the sample size of all subsamples should be the same it was possible to come close to 500 families if the experiment used four subsamples. There were also some considerations for the time it would take to collect the information and analyze the results, which made us want to complete the experiment as soon as possible.

The families who should start to keep diaries during a period of two months, (subsample 19-22) were randomly split into two groups of equal size - one who got the calculator and one who didn't. This gave us a sample size around 470 households in the gross sample, divided into four subsamples of equal size. The sequential design made a fast decision possible if the outcome had been totally different from what we had expected and it had been necessary to stop the experiment.

As the pocket calculators were sent from our central staff the interviewers had no possibility to interfere with the distribution procedure. But it was necessary that the interviewers knew in advance if the household had or had not been sent the calculator to know when they could refer to it their contact work.

5 Choice of testvariables

In the surveys reported by Berk et al the results were fairly easy to analyze as data collection was done on one occasion.

With the more complicated data collection procedure in the 1988 FEX it is not without difficulties to decide exactly which of all possible ratios that will constitute the proper test variable. As nominator we could take the number of performed introductory interviews or the number of families with introductory interviews and completed diaries. As the denominator we could take the gross sample or the net sample or the *net sample except those not contacted*. The third group consists of the households who had no opportunity to react to the incentive.

There is also a problem to decide how long period there should be between the start of the data collection period and the time when the ratio is measured. Should it be measured as soon as there is a report from the interviewers for each subsample or later? As time goes the effect of the incentive only may be weakened and mixed up with the effects of the efforts the interviewing agency makes to persuade nonrespondents. Both nominator and denominator of the ratio change during the period of data collection when the interviewers perform their job. As families were allowed to delay their diary period in about ten per cent of the cases some completed diaries may come in late. Already completed interviews and diaries might be rejected at a closer scrutiny. The denominator will also change as soon as we don't use the gross sample, which is less interesting when there is overcoverage[1] in the sampling frame. More and more of it is identified during the data collection.

[1] Units in the sampling frame but not belonging to the population under study.

The conclusion was that we had to use a set of complementary test variables to describe the effects of the incentive. In the denominator, N(t), we excluded those families that already were identified as part of the overcovarage or whom the interviewers still hadn't had time to attempt a real contact. This number grows as the data collection goes on. Two types of ratio were calculated:

$$\frac{\text{number of introductory interviews performed}}{\text{net sample minus not contacted families}} = R\{I_j(t)\} = I_j(t)/N_j(t) \quad \text{and}$$

$$\frac{\text{number of families having completed diaries}^2}{\text{net sample minus not contacted families}} = R\{I\&D_j(t)\} = I\&D_j(t)/N_j(t)$$

j = e for experimental group, c for control group;

t refers to the period elapsed since the planned day to start diary keeping. The reporting from our interviewing agency was as a routine made with two weeks intermission. It was convenient to calculate the ratios for t= 2, 4 6 and 8 weeks and at the end of all data collection in the survey.

The difference $D_{RI}(2) = R\{I_e(2)\} - R\{I_c(2)\}$ tells us about the effect of the incentive after two weeks. As time goes the incentive will have less and less effect. In the same time efforts are made to persuade refusers and find not at homes so the larger t the more are the effects of different efforts mixed up with each other. But our real interest is not to increase only the interview response rate but the rate of households with completed interviews and diaries. This is measured by $D_{I\&D}(t) = R\{I\&D_e(t)\} - R\{I\&D_c(t)\}$. As it took four weeks to complete the diary, the ratios $R\{I\&D_j(t)\}$ cannot take on any other value than 0 as long as $t<6$.

6 Results of the experiment

The results of the complete experiment are reported in table 1 (experiment group - with calculators) and table 2 (control group - without calculator). The result for each subsample is reported in appendix 1. It should be noted that the tables report the outcome of the experiment in the sample without weighting for varying sampling probabilities. The results should not be taken to represent population conditions.

[2] Only families who have given an introductory interview are asked to fill in their diaries.

Table 1 Response rates for experiment group

Time passed	I	I&D	N	I/N	I&D/N
2 weeks	276	0	394	0.70	0.00
4 weeks	299	0	433	0.69	0.00
6 weeks	306	150	446	0.69	0.34
8 weeks	309	243	454	0.68	0.54
Final	304	298	458	0.66	0.65

* I are those who answered in the introductionary interview
I&D are those who also filled in their diaries
N is net sample minus identified overcoverage and not contacted families

Table 2 Response rates for control Group

Time passed	I	I&D	N	I/N	I&D/N
2 weeks	240	0	413	0.58	0.00
4 weeks	266	0	440	0.60	0.00
6 weeks	275	128	454	0.61	0.28
8 weeks	278	203	462	0.60	0.44
Final	282	276	466	0.61	0.59

In order to evaluate the significance of the differences, critical values for onesided test have been approximated in table 3 for one panel with 115 observations in each of control and experiment group and the complete experiment with 460 observations in each group. The approximation is described in more detail in appendix two.

Table 3 Critical value for onesided significance test at 5% level

Response rate	Sample size in control and experiment	
P	115	460
50	0.108	0.054
55	0.108	0.054
60	0.106	0.053
65	0.103	0.052

For the complete experiment where the four subsamples are put together the differences in response rates are significantly larger than 0. $D_{RI}(2) = R\{I_e(2)\} - R\{I_c(2)\}$ is as large as 0.12. As time goes the difference is reduced to half its original size. If rounding errors are removed the final result will fall just above to the critical value of table 3.

$D_{I\&D}(t) = R\{I\&D_e(t)\} - R\{I\&D_c(t)\}$ is never so large as $D_{RI}(2)$ but stays at 0.06 which is above the critical value.

The variations in response rate between the rounds of the experimental group are within the error margin. The variations between subsamples of the control group are larger. In round 19 and 22 the differences $D_I(t) = R\{I_e(2)\} - R\{I_c(2)\}$ are so large that they exceed the critical value and in round 19 so does also $D_{I\&D}(t) = R\{I\&D_e(t)\} - R\{I\&D_c(t)\}$ when $t>6$. The conclusion must be that the use of the pocket calculator as incentive in advance was successful. The outcome was consistent with the results described by Berks et al but the final effects not quite so large in the 1988 FEX.

The results of this study give some interesting additional information to those presented by Berks et al. (1987). The large differences in the early stages of the data collection were reduced to about six per cent in both ratios when the data collection was finished. The experiment demonstrates that the incentives effect is not only temporary but persisting during a data collection period of one month as only a few families broke off their participation. A study by James and Bolstein (1990) supports the idea of lasting effects. In a mail survey the response rates were higher already after the first mailing and still after the forth when the incentives were worth 1 USD or more. Minor incentives seemed to have no effects.

7 References

Berk, M.L., Mathiowetz, N.A., Ward, E.P. and White, A.W. (1987): The Effect of Prepaid and Promised Incentives; Result of a Controlled Experiment. Journal of Official Statistics, 1987:4, pp 449-458.

James, J.M. and Bolstein, R. (1990): Monetary Incentives and Follow-Up Mailing. Public Opinion Quaterly, 54:3.

Lindström, H.L. (1985): Ersättningens och uppgiftslämnarbördans betydelse för kvaliteten i undersökningarna om hushållens utgifter, (Effect of incentives and respondent burden on the quality of the family expenditure surveys) Promemorior från P/STM nr 19.

Lindström, H.L. (1989): Krisgruppsarbetet och räknarexperimentet i HUT 88. Statistics Sweden, R&D Report 1989:14.

Lindström, H.L., Lindkvist, H. and Näsholm, H. (1989): Design and Quality of the Swedish Family Expenditure Survey, R&D Report 1989:7.

Lindström, H.L. and Lundquist, P. (1989): Application of generalized variance functions in the Swedish 1985 Family Expenditure Survey. To appear in R&D report, Statistics Sweden.

Appendix 1 Responserate rates for subsamples

Subsample 19	Response rates for control group				
Time passed	I	I&D	N	I/N	I&D/N
2 weeks	53	0	102	0.52	0.00
4 weeks	64	0	111	0.58	0.00
6 weeks	65	31	115	0.57	0.27
8 weeks	65	44	117	0.56	0.38
Final	63	62	119	0.53	0.52

Subsample 20	Response rates for control group				
Time passed	I	I&D	N	I/N	I&D/N
2 weeks	59	0	102	0.58	0.00
4 weeks	67	0	107	0.63	0.00
6 weeks	71	31	108	0.66	0.29
8 weeks	73	59	113	0.65	0.52
Final	75	75	113	0.66	0.66

Subsample 21	Response rates for control group				
Time passed	I	I&D	N	I/N	I&D/N
2 weeks	69	0	107	0.64	0.00
4 weeks	72	0	113	0.64	0.00
6 weeks	75	28	118	0.64	0.24
8 weeks	76	57	119	0.64	0.48
Final	76	74	119	0.64	0.62

Subsample 22	Response rates for control group				
Time passed	I	I&D	N	I/N	I&D/N
2 weeks	59	0	102	0.58	0.00
4 weeks	63	0	109	0.58	0.00
6 weeks	64	38	113	0.57	0.34
8 weeks	64	43	113	0.57	0.38
Final	68	65	115	0.59	0.57

Subsample 19	Response rates for experimental group				
Time passed	I	I&D	N	I/N	I&D/N
2 weeks	66	0	97	0.68	0.00
4 weeks	68	0	104	0.65	0.00
6 weeks	72	32	100	0.72	0.32
8 weeks	75	59	111	0.68	0.53
Final	74	73	112	0.66	0.65

Subsample 20	Response rates for experimental group				
Time passed	I	I&D	N	I/N	I&D/N
2 weeks	67	0	99	0.68	0.00
4 weeks	76	0	112	0.68	0.00
6 weeks	74	28	115	0.64	0.24
8 weeks	75	58	116	0.65	0.50
Final	73	73	116	0.63	0.63

Subsample 21	Response rates for experimental group				
Time passed	I	I&D	N	I/N	I&D/N
2 weeks	70	0	96	0.73	0.00
4 weeks	77	0	106	0.73	0.00
6 weeks	82	39	113	0.73	0.35
8 weeks	82	61	112	0.73	0.54
Final	79	76	113	0.70	0.67

Subsample 22	Response rates for experimental group				
Time passed	I	I&D	N	I/N	I&D/N
2 weeks	73	0	102	0.72	0.00
4 weeks	78	0	111	0.70	0.00
6 weeks	78	51	118	0.66	0.43
8 weeks	77	65	115	0.67	0.57
Final	78	76	117	0.67	0.65

Appendix 2 Approximate test variable

The test variable of the difference in response rate between experimental and control groups is:

$$z = \frac{p_e - p_c}{\sqrt{P(1 - P)(\frac{1}{n_e} + \frac{1}{n_c})}}$$

where e indicates experimental group and c control group. p is substituted with the proportion in case. P is estimated under the zero hypothesis and is the weighted proportions of the experimental and control groups. The test is onesided as the use of incentives is expected to increase the response rate.

The formula for the standard error assumes the same response probability for each household. As experience shows that response probabilities vary a lot - most of all with family size - there is good reason to believe that the standard error tends to overestimate the real standard error.

At independent experiments and response probability in experiment nr j is p_j and $\Sigma p_j = np$ the variance for the number of respondents would be

$$Var(\sum_{j=1}^{n} X_j) = \sum_{j=1}^{n} p_j(1 - p_j) = np(1 - p) - \sum_{j=1}^{n} (p_j - p)^2 < np(1 - p)$$

The denominator of the test variables are then too large and the test variable takes on to large values. Each significant differences would be still more so if we could calculate the correct standard error.

Response Behaviour and Seasonal Variation in Swedish Households Expenditure

Peter Lundquist

1 Background of the survey

Family Expenditure Surveys (FEX) have been carried out intermittently at Statistics Sweden; the three most recent ones were conducted in 1978, 1985 and 1988. The next survey is planned for 1992. After FEX 78 it was decided that the interval between the surveys should be shorter than before. The costs for conducting the surveys were reduced to one third of the cost of the FEX 78, which naturally has affected the design of the surveys. From FEX 85 the net sample was reduced to 6000 households. Less information was collected.

From 1985 the reporting of food purchases was simplified. Instead of noting all food items, the respondents report only one figure for all food they buy at one occasion. Food consumption has instead been studied in a separate survey, the Swedish Family Food Expenditure Survey 1989.

After a test in 1983 (Lindkvist and Lundgren (1984)), it was decided that diaries with preprinted headings should be used in the FEX 85. As this diary format, "product diary", gave a satisfactory result in FEX 85, it was also used in the FEX 88.

The sampling procedure in the surveys is essentially simple random sampling of individuals, from the Swedish Register of the Total Population (RTP). The sampling of individuals results in a sampling probability proportional to the household size. The survey population comprised all noninstitutional households with at least one family member not older than 74 years. The households were randomly divided into 26 equal-sized subsamples, each allocated to a specific period of the calendar year (in order to get good yearly estimates). The period of reporting was four weeks. The starting day could be chosen by the respondent, given that the reporting started in the selected week and to last for four weeks.

Further aspects of the design may be found in Näsholm et. al (1989), Ribe (1990) and Näsholm (1991).

2 Introduction to the study

This study is based on the FEX 85, and the purpose is to investigate how the permitted postponements of the diary period could affect the expenditure

estimates. Postponements have been allowed because the responserate otherwise would have been to low. The interviewers were thus instructed, that they could allow the household to delay their diary keeping up to four weeks. Afterwards the central staff analyzed the households that the interviewer failed to contact and those who refused to participate in the survey, in order to find if there were any households left which could be expected to participate in the survey.

Apart from the diaries, about 40 percent of the total expenditures were collected in an additional interwiev in FEX 85 at the beginning of 1986 for all the households and in the FEX 88 in the introductory interview. In this additional interview questions were made about larger expenditures such as rents and capital goods. The postponement does not affect these expenditures in the same way as the expenditures noted in the diary. The study deals with expenditures noted in the diary from the FEX 85.

The effects of nonresponse in the time dimension could introduce bias in the estimates. For example if families predominantly tend to postpone the diary period from the summer because of vacations to the more regular autumn, then the expenditures for typical holiday activities would be underestimated.

The reason why the study is done on the FEX 85, rather than the FEX 88, is that the diary had one couple of pages for each week. The diary used in the FEX 88 was designed in another way, and was because of that less suitable for this purpose.

3 Responding households over time

The total response rate in the FEX 85 was 72 percent, with 62 percent of the sampled households responding in the selected period and 10 percent postponing their diary-keeping to a later subsample. Table 1 presents distribution between the two response modes in the FEX 85, for different household sizes.

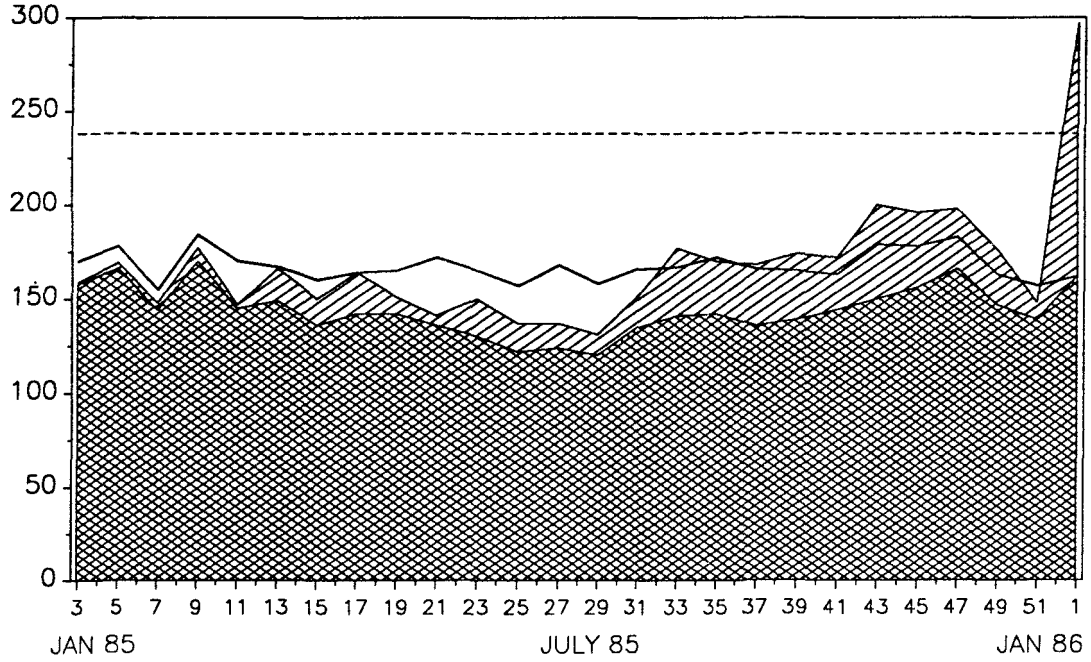
Table 1. Percentage of completed diaries by delay for different household sizes, FEX 85.

Household size according to sampling frame	1	2	3	4	≥5	All Households
No delay	57	58	65	68	73	62
2-4 weeks delay	5	4	5	5	4	5
6-12 weeks delay	2	2	2	2	2	2
14-50 weeks delay	3	3	3	3	3	3
Nonresponse	32	33	24	21	19	27
Number of households	1761	1549	950	1169	575	6004

If we look at the relation between response behaviour and the registered household size, there are no differences for different household sizes in the way they postpone their diary-keeping. The nonresponse rates follow the known manner, i.e. highest among single living persons, and decreasing for larger family sizes. Information on the nonresponding households is available only from the sampling frame (the Register of the Total Population). Unfortunately this information has limited validity regarding the actual household size. The actual household size is supposed to be generally greater for the households. Among the responding households, only 51 percent of those who were classified as single in the RTP were actually living alone. For families consisting of four persons the corresponding proportion is 86 percent correctly classified households. This calls for some care in making conclusions based on the RTP family size in Table 1.

Figure 1 presents how the response pattern varies over the survey period. The area enclosed by the crosshatched lines shows the number of households that participated in the survey during the selected period. The dotted line represents the number of households chosen for each of the different starting periods, constantly equal to 238 households. The area between this crosshatched area and the dotted line represents the "time nonresponse" i.e., those not reporting in the correct period. The solid line describes the number of households who postponed their diary start to a later period. The area consisting of the diagonal lines represents those who delayed their diary start to that actual starting period. The total response is thus the crosshatched area (responding on time) and the area consisting of the diagonal lines (delay to later subsample) together.

FIGURE 1
Completed Diaries per Week



In the estimation procedures of the FEX 85 there was no time for any analysis of the time nonresponse. We now afterwards can see that particularly those households that postponed their diary period are not equally spread over the survey period. Namely, the postponed reporting tends to be concentrated to the second half of the year. There was although a decrease in postponements to the subsamples in December. Among those households that responded in the selected samples one can note a decrease in response during the summer. That would be caused by the summer holidays. The capacity among the interwievers is also low at the summer and could have affected the response rates negatively.

The question is whether we should or should not include the households who delay their diary period in the estimation. If we do, the estimates may have to be compensated for the variation over time. Table 2 indicates that single persons are more disposed to delay their diary-keeping and larger families (5 or more) are more disposed to response on selected time. A Chi-2 test for homogeneity between the two response modes was rejected at the significance level of 5 percent, i.e. the distribution of respondent on the household sizes can not be seen as equal. This confirms the indication that people living alone are harder to get in contact with.

Table 2 *Comparison between respondents who filled in the diary on selected time versus those who postponed the diary-keeping. Percentage distributions with respect to household size, FEX 85.*

Household size according to sampling frame	1	2	3	4	≥5	Total Househ	No of
On selected time	27	24	16	21	11	100	3754
later	31	24	17	20	8	100	611
Net sample	29	26	16	19	10	100	6004

As mentioned before, the RTP family size tends to deviate from the actual family size. We are reduced to use the responding families to analyse the two different response modes. Table 3 confirms our statement that people living alone are more disposed to delay their diary period than families with three or more children. A similar test of homogeneity was also rejected on the 5 percent level for Table 3.

Table 3 Comparison between the two response modes for different family types. Percentage distributions with respect to household size, FEX 85.

	On selected time	Later
Single persons		
without children	14	18
with children	5	7
Cohabitant persons		
without children	22	20
with one child	11	11
with two children	21	20
with more children	10	7
Others		
without children	9	9
with children	9	9
Total	100	100
Number of households	3743	611

In table 3 we consider three different family types for further studies; 1) Single persons without children, 2) Cohabitant persons without children and 3) Cohabitant persons with two children. The aim is to study if there are any differences in response modes over the time for these three family types, and how their expenditure pattern differs over time.

In the FEX 85, the diary period was four weeks. In this study only the two first weeks have been used. Doing so we can form 13 four-week periods, covering the whole year and being (in principle) disjoint with respect to both time and respondents. The idea is that if there is any difference between the two response categories then it will show up already in the two first weeks. The study by Ribe (1990) indicated that there were no obvious differences between the respondents decrement of reporting rates over the diary period.

The distribution of the sample over the 13 periods composition of the 26 subsamples which seemed to be the best is shown in Table 4.

Table 4 Distribution of responding households in the FEX 85 by period.

Period	Subsample	Calendar time	Number of Reporting Households	
			In time	Delayed
1	1,26	Jan 85 + Jan 86	322	134
2	2-3	Feb	310	7
3	4-5	Feb - Mars	315	9
4	6-7	Mars - Apr	285	31
5	8-9	Apr - May	284	31
6	10-11	May - June	266	25
7	12-13	June - July	246	28
8	14-15	July - Aug	255	27
9	16-17	Aug - Sep	283	64
10	18-19	Sep - Oct	275	67
11	20-21	Oct - Nov	294	78
12	22-23	Nov	322	72
13	24-25	Dec	286	38
Σ			3743	611

The selection of the two first diary weeks of the households also made it possible to obtain 13 respondent groups with, in principle, no overlap in time. Because of the fact that the startday was not fixed there will though overlap in time among some households. The first period consists of the first and the last subsample in the survey, i.e. there is a time lag of one year between these two subsamples. The reason for aggregating subsample 1 and 26 together is that one could expect that the consumption pattern and response behaviour are more equal between these subsamples than between subsample 25 and 26.

In Figure 2 is the number of responding households plotted. The dashed line presents the households that participated in the selected period and the solid line presents the total number of households that kept their diary in that period, i.e. the solid line is the total response during the period. The peak at period one for the total response is caused by the last subsample in the survey. The figures of the number responding households are also given in Table 4.

A Chi-2 test was conducted on the households that participated at the selected time and the total response. The purpose was to test the hypothesis that the households were uniformly spread over the 13 four-week periods. The hypothesis was rejected for both the lines (at the 5 percent level), i.e. even the distribution of the households that participated on selected time varied over the 13 periods.

FIGURE 2
All households
over 13 four-week periods

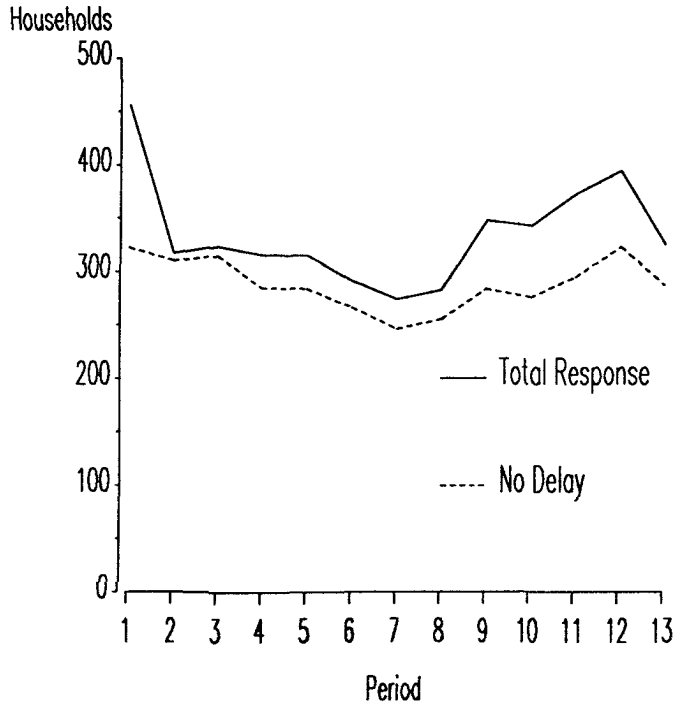


FIGURE 4
Cohabitant households without children
over 13 four-week periods

FIGURE 3
Single living persons
over 13 four-week periods

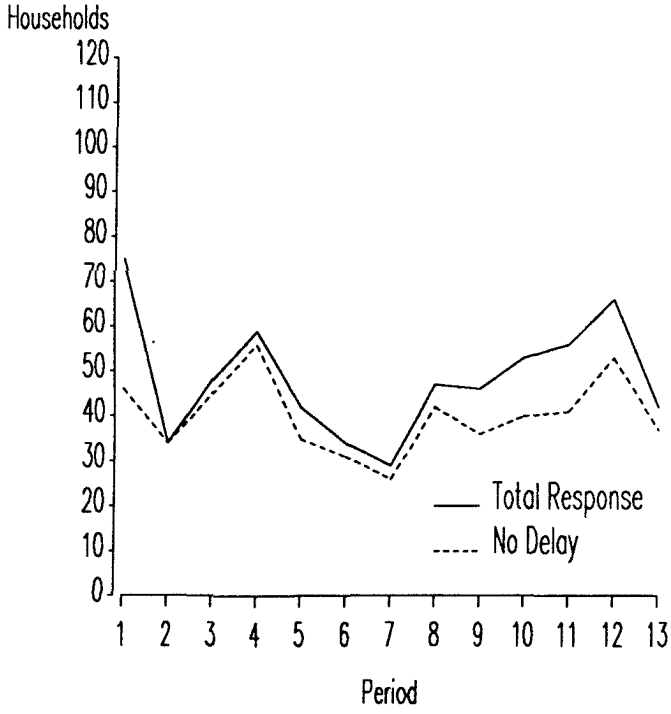
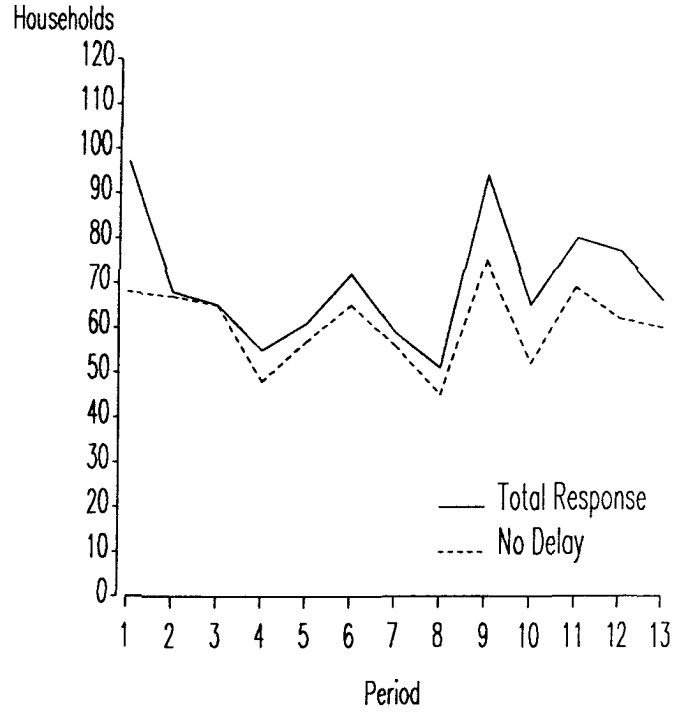


FIGURE 5
Cohabitant households with two children
over 13 four-week periods



The family types of Table 3, were also studied in the same way (presented in Figure 3-5). For the total response we found that there were significant differences in the frequency of diary keeping over the periods, for all the family types. But if the same hypothesis was tested on the households who participated at the selected time, only the single living persons' response pattern was significantly different (at the 5 percent level) over the 13 time-periods.

A Chi-2 test (at the 5 percent level) of only those who delayed their diary keeping among the three different family types, confirmed no differences in distribution among them.

The response pattern in Figure 5, cohabitant persons with two children, deviates most from the response pattern in Figure 2, all households. The fluctuation in Figure 3, single persons is stronger than that for cohabitant households without children, Figure 4. The major holidays were; during period 4 (Easter), period 6-8 the summer vacations, and period 13 (Christmas). The influence of holidays is different for the different household types, except for the Christmas, where they show the same pattern.

It is hard to know the reason for postponements among the households. The motivation among the interwievers in handling the FEX, and their ability to persuade the households to participate in the survey varies a lot. It is also known that the response rates decrease if the interwievers are overloaded with work; Lindström (1989). In the Swedish Time Budget Survey (TBS) I. Lyberg (1990) suggested that the interwievers should be instructed to note the reason for postponements among the households. In the TBS 90/91 this information is collected. The conditions are not the same in the two surveys, the respondent burden is larger in the FEX than in the TBS. Because of that it would be necessary to instruct the interwievers to record the reason for postponing in the succeeding FEX.

4 Expenditure pattern over the time

For the thirteen different time periods, the yearly expenditure was estimated. Two estimates were computed, the first for the households that kept their diary at the selected time and the second based on the total response. Among these two estimates we have to choose the one to be preferred.

The total expenditure and three different expenditures are estimated; 1) All expenditure 2) Food, 3) Eating out and 4) Entertainment. In Figure 6-9 the expenditure estimates for the thirteen period estimates are divided by the average of all the thirteen period estimates. This is done separately for total response and response with no delay. The estimates are presented for the two response categories, response in selected period (dashed line) and the total response (solid line).

FIGURE 6
All expenditure
Relative expenditure for the 13 four-week periods
relatively to the total means
All households

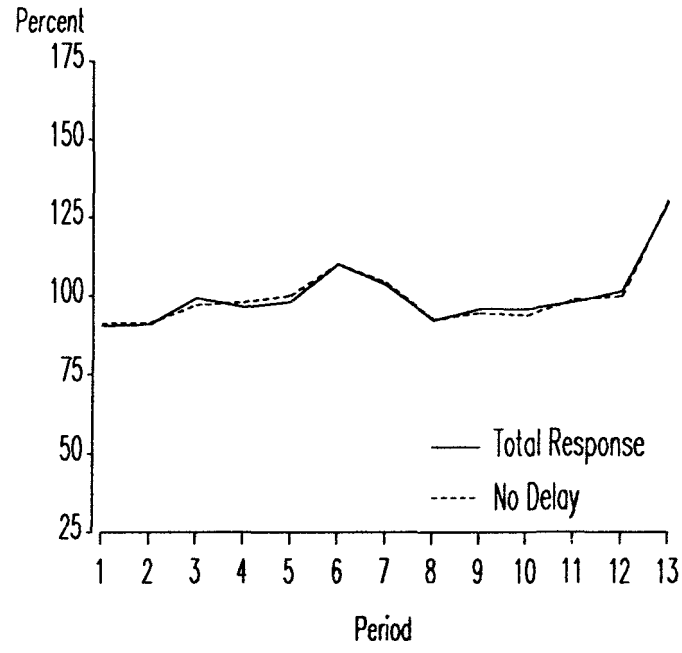


FIGURE 8
Eating out
Relative expenditure for the 13 four-week periods
relatively to the total means
All households

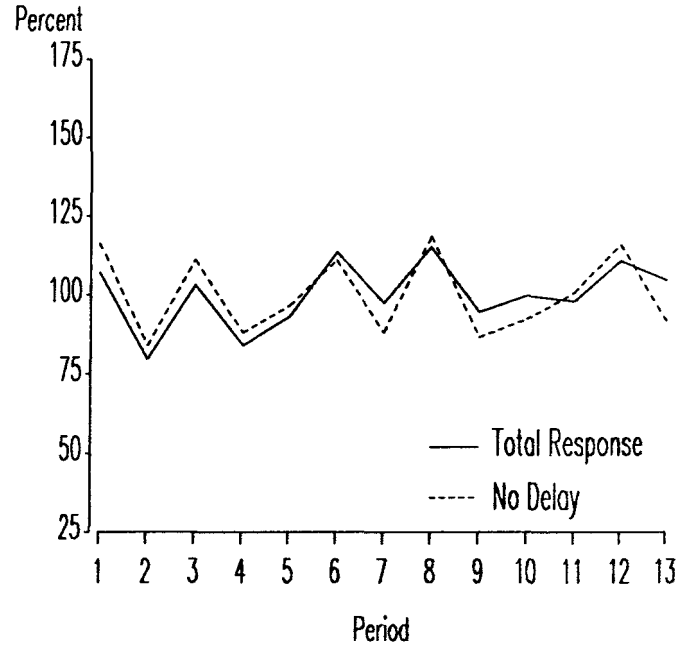


FIGURE 7
Food
Relative expenditure for the 13 four-week periods
relatively to the total means
All households

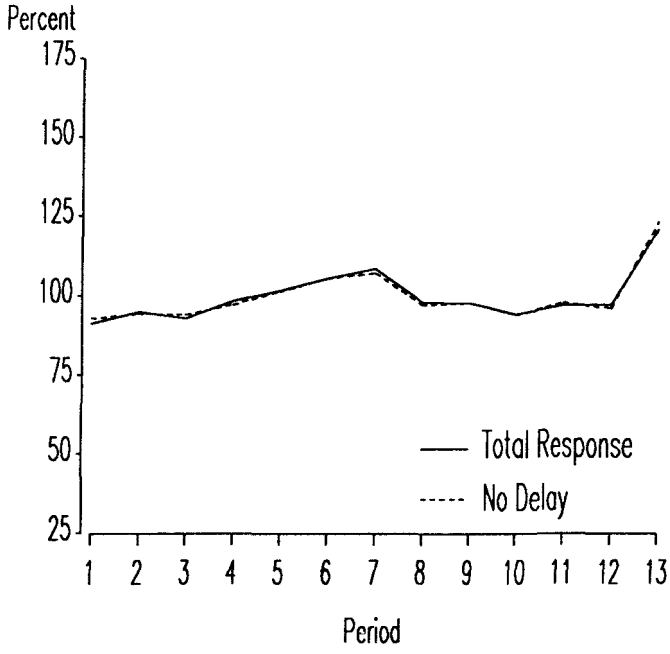
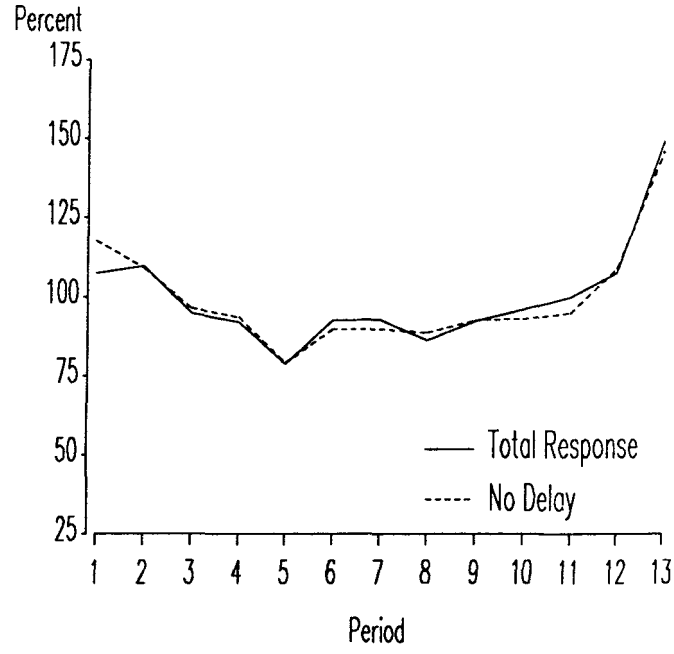


FIGURE 9
Entertainment
Relative expenditure for the 13 four-week periods
relatively to the total means
All households



The expenditures for Eating out in Figure 8 deviates most from Figure 6 All expenditures. The other expenditures follow the pattern of All expenditure with an increase in period 13, where the increase is caused by the Christmas. The same expenditures were plotted for the three different household types; Single persons without children, Cohabitant without children and Cohabitant with two children, and are presented in the Appendix.

For most kinds of goods and types of households, the period estimates of consumption were highly significantly nonconstant over the year. This was tested in the following way. If we normalize the period estimates by subtracting their average over the year and divide them by their estimated sampling errors, then the squares of those 13 normalized numbers would be approximately Chi-2 distributed with 12 degrees of freedom, under the null hypothesis of constant consumption. Table 5 presents the results of the test.

Table 5 Chi-2 test if the expenditure estimates are uniformly distributed over the 13 four-week periods, for both response modes.

	Chi-2	df = 12	Chi-2	df = 12
	Total response	p	No delay	p
All expenditures				
All households	49.14	0.000	44.20	0.000
Single persons	10.66	0.559	13.25	0.351
Cohabitant without children	51.17	0.000	51.08	0.000
Cohabitant with children	39.68	0.000	33.53	0.001
Food				
All households	45.02	0.000	38.97	0.000
Single persons	5.78	0.927	4.782	0.965
Cohabitant without children	31.21	0.002	35.48	0.000
Cohabitant with children	30.08	0.003	26.89	0.008
Eating out				
All households	16.64	0.164	18.38	0.105
Single persons	28.27	0.005	31.12	0.002
Cohabitant without children	7.83	0.798	9.15	0.690
Cohabitant with children	36.72	0.000	37.63	0.000
Entertainment				
All households	51.90	0.000	49.28	0.000
Single persons	11.26	0.507	9.87	0.628
Cohabitant without children	44.64	0.000	38.22	0.000
Cohabitant with children	44.33	0.000	43.35	0.000

It is seen that the variation over the year is mostly significant, with single persons as a notable exception. For single persons only Eating out had significant variation.

For comparing the estimated consumption between the two response modes, a paired comparison t-test of standard type was made, with a sample consisting of the 13 periods. This was done for all households and the three different family types. The results are presented in Table 6. The estimates for households who participated on selected time were subtracted from the estimate on the total response. A negative difference of the mean estimates on the thirteen periods indicates that the households that kept their diary on selected time had larger expenditures. For All expenditures the test indicates that there could be some differences especially for families with two children, and maybe for single living persons. Except of that there seems to be a difference in Eating out, but none of the family types in the study could confirm that.

Table 6 Comparison t-test between the two response modes

	Difference of Means	Std. Dev	t	Prob>| t|
All expenditures				
All households	601.66	1005.02	2.16	0.052
Single persons	1651.92	2761.39	2.16	0.052
Cohabitant without children	405.17	2700.68	0.54	0.599
Cohabitant with children	1448.27	2151.04	2.43	0.032
Food				
All households	-108.22	267.40	-1.45	0.170
Single persons	25.20	356.83	0.25	0.803
Cohabitant without children	-25.62	758.24	-0.12	0.905
Cohabitant with children	336.27	775.73	1.56	0.144
Eating out				
All households	145.30	221.66	2.36	0.036
Single persons	80.42	370.43	0.78	0.449
Cohabitant without children	343.33	746.46	1.65	0.123
Cohabitant with children	39.80	266.57	0.54	0.600
Entertainment				
All households	-6.45	207.82	-0.11	0.913
Single persons	-15.20	273.01	-0.20	0.844
Cohabitant without children	-73.18	359.60	-0.73	0.472
Cohabitant with children	-36.09	484.68	-0.27	0.793

5 Conclusion

There is no strong evidence that the households that postpone their diary period should be excluded in the computation of the survey estimates. There might be a difference between the two response modes but it can hardly be confirmed. It is plausible to think that the postponements are a combination of interwiever effects and different living circumstances among the households.

It is clear that the postponing households are overrepresented at the end of the year, except for Christmas. The expenditures also show a variation over time with an increase in the Christmas period because of that goods typical for the autumn might be overestimated, and conversely for goods typical for the spring and summer periods if the estimates are not compensated for the differences in response over time.

As mentioned by I Lyberg (1989) further information about the postponements has to be collected by the interwievers to see if there is an association between postponements and differences in the expenditures.

6 References

Johansson, G., Lundquist, P., Näsholm, H. and Säfström, V. (1991): Undersökningarna om hushållens utgifter 1985 och 1988. Teknisk rapport, Statistiska Centralbyrån, (In Swedish).

Lindkvist, H. and Lundgren, L. (1984): Hushållens utgifter - hösten 1983. Statistiska Centralbyrån, arbetsmaterial, 1984-04-13, (In Swedish).

Lindström, H.L. (1989): Krisgruppsarbetet och räknarexperimentet i HUT 88. Statistics Sweden, R&D Report 1989:14, (In Swedish).

Lyberg, I. (1989): Sampling, nonresponse, and measurement issues in the 1984/85 Swedish Time Budget Survey. Paper presented at US Bureau of the Census Fifth Annual Research Conference (ARC V) in Arlington, 19-22 March, 1989, pp 210-234. [Also R&D Report 1990:2, Statistics Sweden.]

Näsholm, H., Lindström, H.L. and Lindkvist, H. (1989): Response burden and data quality in the Swedish Family Expenditure Surveys. Paper presented at US Bureau of the Census Fifth Annual Research Conference (ARC V) in Arlington, 19-22 March, 1989, pp 501-514. [Also R&D Report 1989:7, Statistics Sweden.]

Näsholm, H. (1991); Household expenditure surveys. A methodologicalreview. To appear at the Workshop on Diary Surveys at Statistics Sweden, February 18-20, 1991.

Ribe, M. (1990): A study of errors in swedish consumption data. Paper presented at the International Conference on Measurement Errors in Surveys, Tucson, Arizona, November 11-14, 1990.

Silberstein, A.R. and Scott, S. (1990): Expenditure diary surveys and their associated errors. Paper presented at the International Conference on Measurement Errors in Surveys, Tucson, Arizona, November 11-14, 1990.

FIGURE 10
All expenditure
Relative expenditure for the 13 four-week periods relatively to the total means
Single living persons

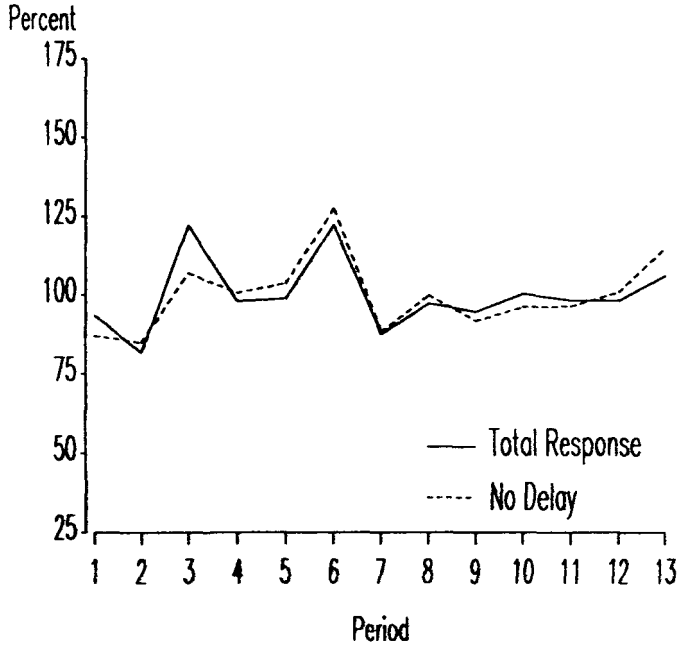


FIGURE 11
Food
Relative expenditure for the 13 four-week periods relatively to the total means
Single living persons

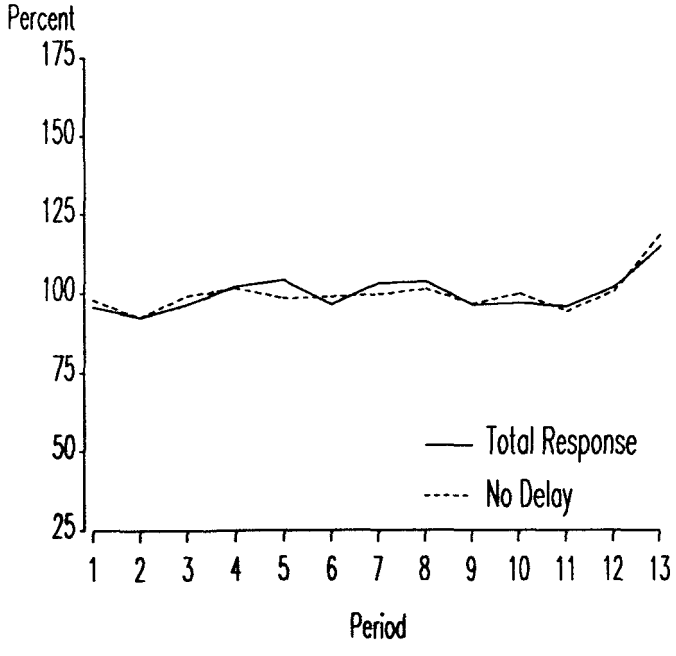


FIGURE 12
Eating out
Relative expenditure for the 13 four-week periods relatively to the total means
Single living persons

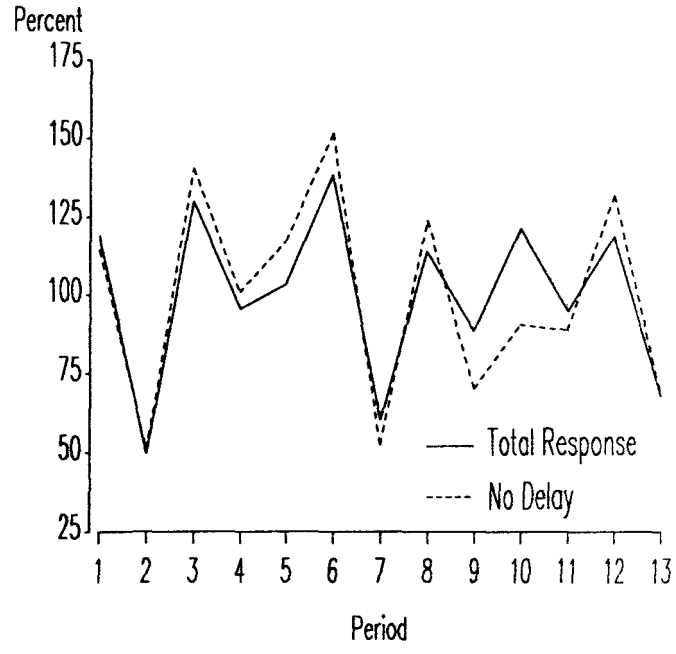


FIGURE 13
Entertainment
Relative expenditure for the 13 four-week periods relatively to the total means
Single living persons

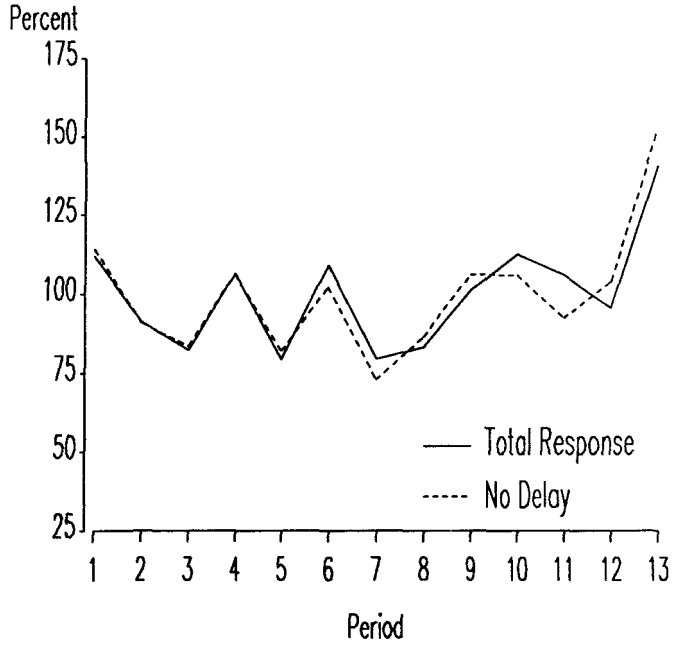


FIGURE 14
All expenditure
Relative expenditure for the 13 four-week periods relatively to the total means
Cohabitant households without children

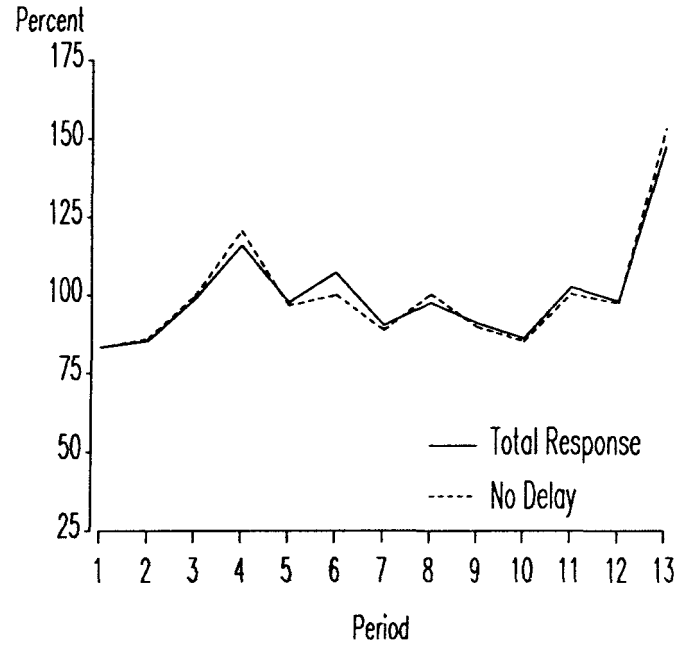


FIGURE 15
Food
Relative expenditure for the 13 four-week periods relatively to the total means
Cohabitant households without children

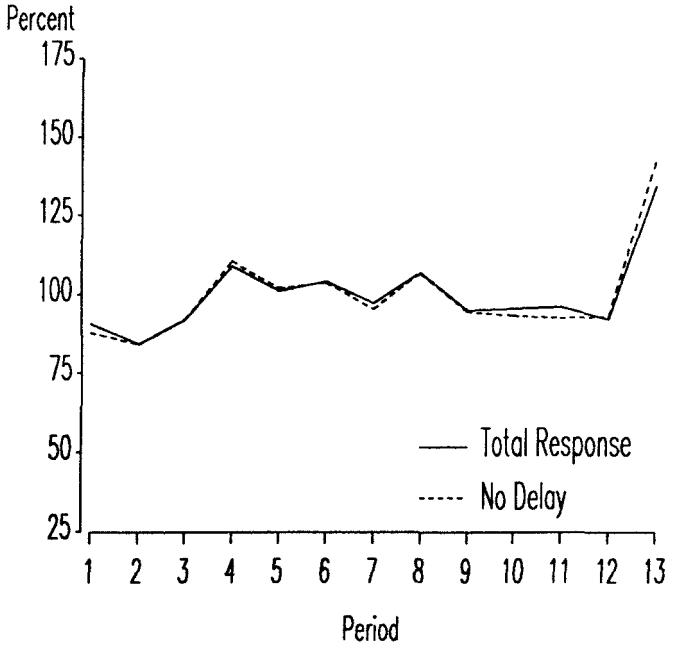


FIGURE 16
Eating out
Relative expenditure for the 13 four-week periods relatively to the total means
Cohabitant households without children

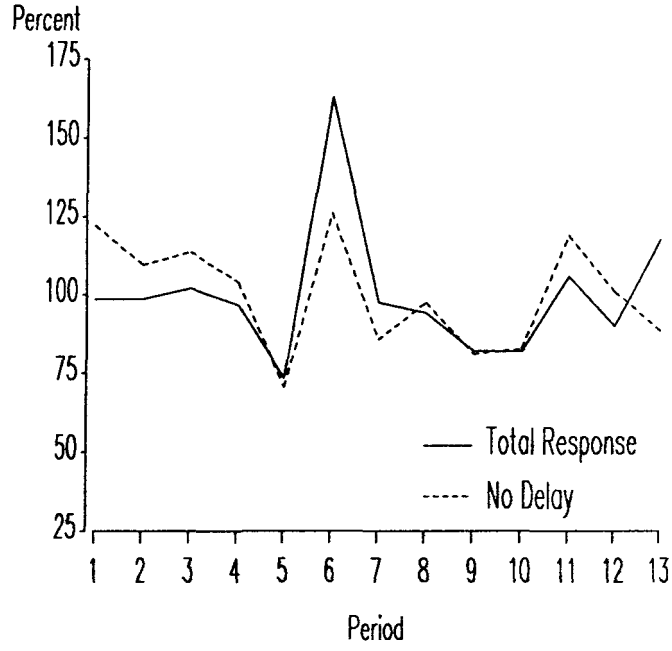


FIGURE 17
Entertainment
Relative expenditure for the 13 four-week periods relatively to the total means
Cohabitant households without children

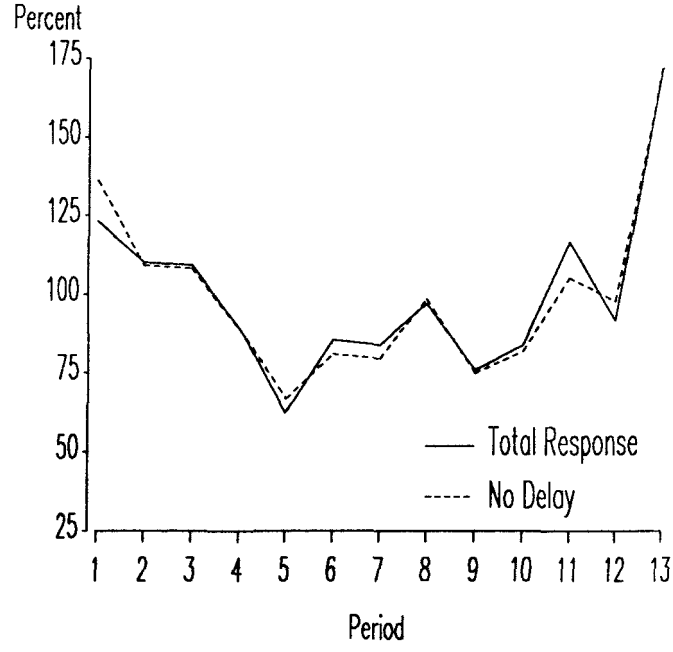


FIGURE 18
All expenditure
Relative expenditure for the 13 four-week periods relatively to the total means
Cohabitant households with two children

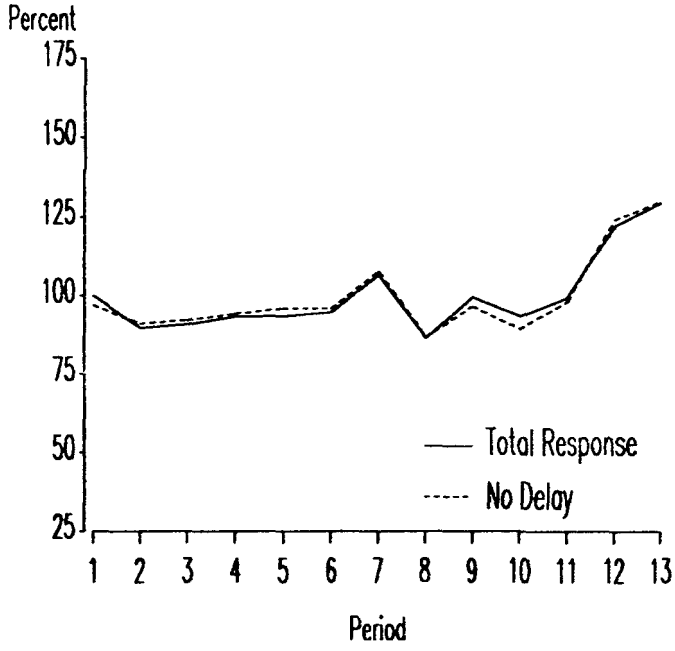


FIGURE 19
Food
Relative expenditure for the 13 four-week periods relatively to the total means
Cohabitant households with two children

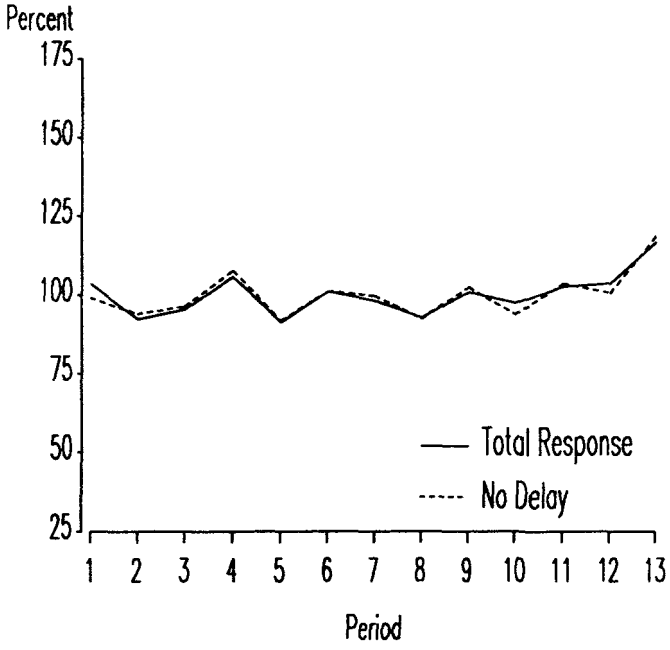


FIGURE 20
Eating out
Relative expenditure for the 13 four-week periods relatively to the total means
Cohabitant households with two children

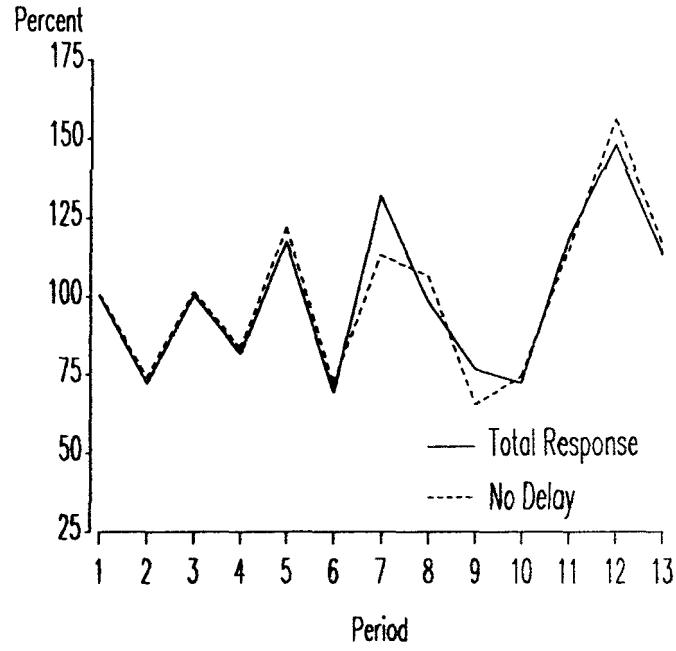
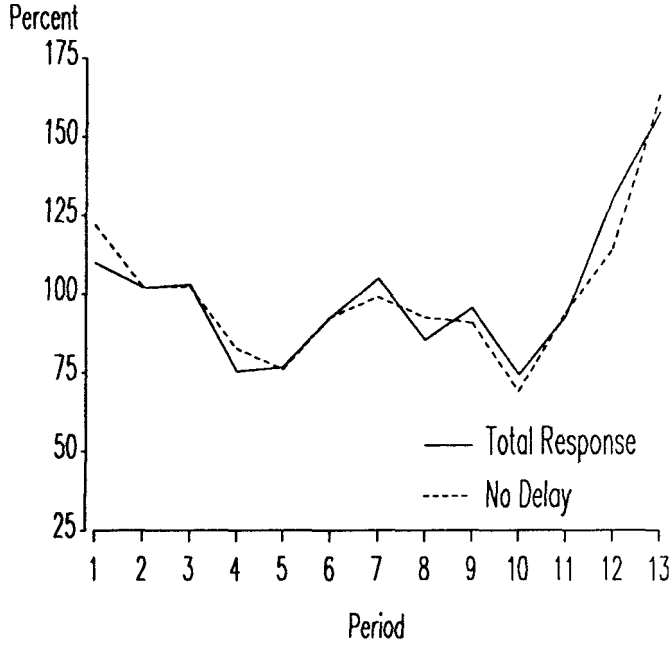


FIGURE 21
Entertainment
Relative expenditure for the 13 four-week periods relatively to the total means
Cohabitant households with two children



Reducing Nonresponse Rates in Family Expenditure Surveys by Forming Ad Hoc Task Forces

Lars Lyberg

1. Introduction

Most household expenditure surveys use a one-year data collection period to accommodate for seasonal variations in purchasing patterns. The typical design is one where the households are divided in subsamples that together cover the survey year and each household is requested to keep a diary of purchases for a number of weeks. This design imposes a great deal of respondent burden, greater than in other types of household surveys. This respondent burden usually results in high nonresponse rates. Nonresponse rates between 20 and 35% are typical and are a constant concern to the researcher.

Since nonresponse is such an important error source in expenditure surveys using diaries, the designer typically includes a number of preventive measures (for the data collection) so that the nonresponse rate can be kept on an acceptable level. Examples of such measures are advance letters, promotional material, interviewer-assistance in the respondent's record-keeping including delivery and pick-up of diaries, incentives, and repeated call-backs. These are used in many other types of surveys as well. What is perhaps special in expenditure surveys is that often the whole battery of measures is used.

The long data collection period allows for special efforts if it is discovered early that the nonresponse rate is unacceptable. In 1978 and 1988, Statistics Sweden formed special task forces to deal with the large nonresponse rates in its household expenditure surveys. The basic idea is that a special task force, with members from various specialties, for instance, subject matter specialist, interviewers, and methodologists, can come up with a highly efficient set of measures that can work both for the part of the sample that has been unsuccessfully approached and for the part of the sample that has yet to be contacted.

This paper describes the various measures taken by the ad hoc task forces and discusses why some measures actually planned for never were realized. No attepts with a few exceptions are made at estimating the measures' relative effect on reducing the nonresponse rates, simply because there was no room for experimental approaches. The main objective for the task forces was to do something that made sense to the members of the task force, based on their earlier experiences with reducing nonresonse rates. Even though it cannot be proved that some of the measures had an effect, due to a lack of controlled

experimentation, we strongly believe that final nonresponse rates would have been much higher if these actions had not been taken.

2. The 1978 Family Expenditure Survey

The 1978 Family Expenditure Survey (FEX) was preceded by pilot studies and experiments that predicted the nonresponse rate at about 24%. The survey started in January and preliminary calculations for the first two subsamples indicated a nonresponse rate of 17-20%, which was considered promising compared to the indications of the pilot study. Definite calculations for the first subsamples were not presented until the end of March and it then turned out that the preliminary calculations had been way off target.

Table 1 Nonresponse rates for subsamples 1 and 2

Respondent category	Number	Percent
Diary completed	358	60
Introductory interview completed and diary keeping probably in progress	34	6
Ill, might be approached later	6	1
Ill, will not be approached later	16	3
Noncontacts, might be approached later	7	1
Noncontacts, will not be approached later	2	0
Refusals, might be approached later	58	10
Refusals, will not be approached later	112	19
Other	5	0
Total	598	100

From Table 1, it is clear that the data collection had been protracted. Also, the final nonresponse rate would probably be at least 22%, possibly above 30%, but at the most, no higher than 40%.

2.1 The ad hoc task force and its plan of action

Given the discouraging result, an ad hoc task force was immediately formed. The task force consisted of subject matter specialists, supervisors and interviewers, and survey methodologists. It was decided that the plan would include a description and analysis of the nonresponse with the ultimate goal of reducing the final nonresponse rate. The task force agreed on the following goal: The nonresponse rate should not exceed 30% over the entire one-year survey period. After two months, the task force itself would be evaluated to determine whether it should continue or be terminated.

The task force produced the following plan of action:

1. All interviewers were informed about the seriousness of the nonresponse rates. It was hoped that the interviewers would, once having been made aware of the situation, increase their efforts to convert initial refusals. Another purpose was to seek advice from the interviewers on possible remedies or courses of action.

2. It was decided that tables of the nonresponse distribution by region, household size, age of head of the household, and interviewer should be produced.

3. A press conference was organized where the users of FEX were asked to confirm the importance of the survey.

4. Groups of interviewers were formed in large metropolitan areas. The notion was that creating networks for the interviewers might open avenues for them to help and support each other, thus resulting in lower nonresponse rates.

5. It is well known that incentives are an efficient way to bring down nonresponse rates. Nevertheless, to increase financial reimbursement to all households to decrease the nonresponse rate significantly was judged too costly. Therefore it was decided to investigate what the consequences would be of conducting some type of lottery where all participating household had a chance of being compensated, but only a few actually would be compensated. Prizes that were under consideration included a new car and government bonds.

6. A more intense follow-up of nonrespondents was considered essential. One way would be to centralize parts of the refusal conversion so that refusals would be contacted by people from the central interviewing staff but also by the FEX researchers themselves. Converting refusals could also be a main task for the interviewer networks.

7. At Statistics Sweden, all interviewers work with all continuing and intermittent surveys. This is at least basically true and the practice is a result of Statistics Sweden's interviewer policy. All interviewers are permanently employed and they should be able to work on all our surveys to maximize assignment flexibility. Such a policy has obvious drawbacks, though. For intermittent surveys, such as the FEX, assignment sizes per interviewer become very small. This is not very efficient from a training point of view and a widely held opinion is that if assignment sizes do not reach a critical minimum, some interviewers will not feel comfortable with the survey, which results in high nonresponse rates. There is also a psychological side. If an interviewer wins the cooperation of two out of three households assigned, the interviewer might consider this a good result while the formal nonresponse rate is 33%. Therefore, it was decided that the number of interviewers working with FEX should be decreased. A first step in that direction would be to terminate training of new interviewers and just use those who already had started working with FEX.

8. It was decided that interviewer debriefing groups should be organized so that interviewers' opinions could be properly heard and acted upon.

The plan of action described above was approved by the heads of the Survey Research Center and the Interviewing Staff.

2.2 Results of the plan of action

No plans survive confrontation with reality. Some suggestions had to be abandoned while most of the others had to be modified. Let us first point out though, that the general goal of the operation was achieved. The total nonresponse rate for the entire period decreased to less than 30%. The factors that contributed to this result are presented below.

The general information to interviewers and the organization of interviewer debriefing groups went according to plans. A large number of meetings were held with the interviewer and the ad hoc task force concluded that these intensive contacts must have alerted the interviewers and made them more interested in the FEX. As a matter of fact, the entire FEX benefitted from the various measures taken by the ad hoc task force. Most interviewers had a feeling of being involved in something important.

The analysis of the nonresponse distribution revealed that the geographic variability was considerable. The final analysis showed that there was a range in nonresponse rates for Sweden's 25 counties from 15.6% (Jämtland) to 36% (metropolitan Göteborg). Also one-person households had a nonresponse rate twice as large as households with four or more members. That is, of course, not surprising. It has always been true that expenditure surveys hold more interest for families with children than for smaller households. It is also generally true that households with very young heads (under age 24) and heads 55 year of age or older are nonrespondents more often than households with heads between 25 and 54.

During the entire survey period, nonresponse rates for individual interviewers were calculated. The stability of these figures is sometimes doubtful because of the small assignment sizes and the fact that so much reassignment was necessary that it led to calculations that sometimes were not fair to the interviewers. In the main, though, the range between interviewers was considerable. For instance, in one county, the 25% of the interviewers with the lowest rates had an average nonresponse rate of about 10% while the worst quartile had an average rate of about 60%. Parts of these differences are due to small assignment sizes and interviewing area, but they undoubtedly also reflect real differences in interviewer performance. Patterns like these, which were fairly stable throughout the survey period, suggested measures to be taken regarding respondents in metropolitan areas, small households, households with young and relatively old heads, and interviewers who needed help. This is all common knowledge.

Contact with the press never materialized as we had expected. The interest from the media simply was not there. Nevertheless, a few articles, mostly in consumer related magazines, were published, but the effects of that part of the public

relations work was marginal. Perhaps more important than receiving good press, we did not receive any bad press. We concluded that no or neutral press is better for respondent cooperation than bad press.

Networks of interviewers were formed in a number of places. Some of them worked very well while a few had to be abolished. The reasons that some networks did not function well were never clear. It is obvious, however, that if some interviewers were not interested in cooperative work, the network would eventually fall apart. Those networks that functioned well did a very good job of converting refusals.

On the other hand, the whole refusal converting process can have serious drawbacks. Usually, given that converting is successful the diary keeping has to be done during another time than the one initially assigned. This creates a kind of nonresponse in the time dimension, with fewer diaries kept during the beginning of the survey year and more towards the end. To our knowledge, the procedure of offering respondents another diary-keeping period in order to win cooperation is an unusual practice. As many as 17% of the participating households kept diaries in periods other than those initially assigned and the effects of this procedure are not clear since no evaluation study has been conducted. Nevertheless, such an option is a powerful tool in reducing nonresponse rates. But the procedure is effective only if the resulting decrease in nonresponse error outweighs the resulting error due to record-keeping in the wrong period.

The introduction of new incentives never worked and new incentives were not even tried. Organizing a lottery was considered dubious. First, there were problems in creating a situation where all households including those which already had participated in FEX had an equal probability of winning the prize. Second, and partly to our relief, the government was not enthusiastic about the idea. Even though the final decision was left to Statistics Sweden, the advice of the Department of Trade was followed and no lottery was conducted.

As mentioned earlier, the follow-up of refusals and reluctant respondents was one of the most important ingredients of the work of the ad hoc task force. Much of the work done during the second part of the survey period was on the follow-up of refusals and reluctant respondents. Refusals that otherwise would have been laid to rest were re-assigned to the interviewer networks and to individual interviewers. The centralized telephone group was also involved and even the ad hoc task force members made attempts, although not very successful ones.

The idea to create a special corps of interviewers for the FEX proved a failure. The Interviewing Staff did not like the idea. The task force was, however, successful at limiting the number of new interviewers trained for FEX, but no attempts were made at removing interviewers with high nonresponse rates from the survey. Nor was it permitted that those interviewers with especially high nonresponse rates be given additional training to correct for their inordinately high nonresponse rates. It is difficult to explain why measures directed at specific

interviewers were so controversial. At the time, there was a great deal of discussion and disparate opinions on the best way to run the interviewing unit and the singling out of individual interviewers was in conflict with the general goals of the Interviewing Staff.

The final result of FEX's more extensive type of diary-keeping is presented in Table 2.

Table 2 Field work results

Respondent category	Number of respondents	Percent respondents
Respondents	4,168	72.4
Year interview conducted and diary kept in correct period	3,141	54.5
Year interview conducted but postponed diary keeping	911	15.8
Only diary kept, but in correct period	62	1.1
Only postponed diary keeping	54	0.9
Nonrespondents		
Refusals	1,592	27.6
Illness	157	2.7
Language problems	33	0.6
Noncontacts	45	0.8
Total	5,760	100.0

As can be seen from the table, the final nonresponse rate was 27.6% which is well within the goal that had been set by the ad hoc task force.

3. The 1988 FEX

The 1985 FEX proved undramatic. The final nonresponse rate for the diary part of that survey came to 27% which was considered acceptable given the circumstances, that is, no special measures had been taken. In addition, most of the surveys conducted by Statistics Sweden between 1979 and 1984 had experienced slightly increasing nonresponse rates. One reason the nonresponse rate could be kept at the same level as in 1978 was that the respondent burden

had decreased somewhat. In 1985 fewer types of purchases had to be recorded by the respondents.

The 1988 FEX design was almost identical to the one used in 1985. However, after three months of survey work, it was evident that nonresponse would be much worse than it had been three years earlier. It was clear that something similar to the measures taken in 1978 had to be taken. In June, the proportion of households participating in the diary keeping was 45% compared to 61% at the same stage in the 1985 survey. Most of the increase in the nonresponse was due to an increase in the number of refusals and it was obvious that the goal nonresponse rate of 30% could not be reached unless special measures were taken.

3.1 A new ad hoc task force

In June of 1985, the Director General approved the establishment of a new ad hoc task force of essentially the same type as the one working in 1978. This time, however, it seemed as if the situation was worse than in 1978, mainly because the interviewers had a tremendous workload during the rest of the year with several intermittent surveys planned, including an Election Survey. The implications of this fact was that there was relatively little time available for repeated refusal conversion. Another important issue was the fact that the group was formed in the beginning of the summer, so the group had to fight not only the current situation but also the quite normal increase in nonresponse rates that always occur in Swedish surveys during the long summer vacation period.

3.2 Some preliminary analysis

Data from the period of January through the beginning of June showed that the proportion of refusals had increased for all types of households compared to the 1985 FEX for the same period. The same pattern emerged both for households by size and households by age of head. The variation of the average increase of eight percentage points in the refusal rate over the various breakdowns was moderate. Therefore it was difficult to identify types of households that contributed more than others to the increase.

It was also clear that nothing could be done about the basic design at this point. Statistics Sweden has a fairly stable corps of interviewers and central interviewing staff. A majority of these people had worked on both the 1985 and the 1988 FEX. Contacts with interviewers who have worked on both surveys revealed that they found it much more difficult to motivate respondents in 1988. The number of households stating "lack of time" as their main reason for not participating seemed to be much larger in 1988.

One reason for a decreased interest on the part of the sampled households might be the reduced value of the incentives. In 1985 the value of the compensation was

smaller than in 1978 without affecting the nonresponse rate directly. (However, with a compensation comparable to the one provided in 1978, the nonresponse rate might have decreased compared to 1978.) The 1988 compensation for participating households was about the same as in 1985, that is, a subscription to a consumer magazine, a lottery ticket, and a ruler with a calendar. It is possible that some households that might have participated in 1985 with that kind of incentive, by 1988 chose not to participate because inflation had outstripped the value of these incentives.

As mentioned above, the main difference between 1985 and 1988 was that interviewer workload was much larger in 1988. The number of interviewers had decreased from 220 in 1985 to 190 in 1988. Due to travelling costs, a larger portion of the corps had to work with the 1988 FEX, thus making all efforts on specialization more or less infeasible. Instead, a number of inexperienced interviewers had to work with the 1988 FEX and the average difference in nonresponse rates between inexperienced and experienced interviewers turned out to be five percentage points.

3.3 The plan of action

The following measures were taken by the 1988 ad hoc task force.

1. A revision of information to sampled households

Since so many nonrespondents had cited "lack of time" as their reason for not participating, a revision of the information for respondents was undertaken. The emphasis was shifted to how easy the task was and how little time it demanded. In fact, the words "diary" and "record keeping" were avoided; words like "note" and "write down" were used. Additional materials were designed for those households who were to participate during the Christmas shopping period. Also, the interviewers were equipped with more extensive material for use in converting refusals.

2. Increasing the compensation

It was decided to increase the incentive by sending sampled households a pocket calculator before the interviewer contacted the household. The effect of this gift on response rates was evaluated in an imbedded experiment. This was a rather unconventional move on the part of Statistics Sweden. First, those who already had participated did not receive the calculator. Second, the standard policy is usually to provide some sort of "reward" or compensation for promised or already secured participation. The analysis of the experiment showed that the gift had the desired effect. The response rate among those who received the calculator, in addition to the general incentives, was six percentage points higher than for those who received the general incentives only.

3. Changing the diary keeping period

Changing the diary keeping period was used extensively in 1978 in order to make it easier for households to participate. The 1985 FEX was much more restrictive on this point. Now it was considered that restriction was relaxed again.

4. Why are the successful interviewers successful?

It is a common dream among survey researchers to have interviewers with small nonresponse rates share their secrets with their less successful colleagues. This approach has been tried before without noticeable effect. So was the case this time as well. Some nebulous skill beyond the standard skills of interviewers is required for refusal conversion, attitudes or characteristics that depend on body language, voice, flexibility, charisma, and natural talent. Generally, these characteristics cannot be taught or learned, at least not to a noticeable degree. However, a number of not so successful interviewers were repeatedly contacted in order to restore some confidence in them. Some interviewers who expressed unwillingness to work with FEX were relieved from the survey.

5. Interviewer resources

In order to handle the heavy workload, four highly experienced retired interviewers were re-hired temporarily. It turned out that these interviewers were very successful. As a matter of fact, one of them had no nonresponse whatsoever.

Several hundred initially refusing households were contacted again and 18% of them agreed to participate.

3.4 Results

The final nonresponse rate for the entire survey was 37% which basically was the same as the rate in June. The measures taken were not sufficient to reach the stated goal of 30%. Had no special measures been taken, the final nonresponse rate would probably have been much higher given the summer and the upcoming large Election Survey. Scarce interviewer resources and the fact that the measures were taken late in the survey contributed to the relatively disappointing result.

4. Endnote

Experience from the work of these two ad hoc task forces can teach us a great deal. It is possible to get good results, but the process resembles the turning of an oil tanker; it takes time. Therefore, one should start as early as possible. If large nonresponse rates can be expected, a special catastrophe plan should be available from the start.

The very existence of an ad hoc group can create more interest and awareness among all parties involved. It is also important that the group be autonomous, at least to some extent. We have learned that it takes a number of different measures to handle a severe nonresponse situation. One of the most important ones, however, is to make sure that refusals can be recontacted in various ways. It is frustrating when we know that good refusal conversion rates can be obtained but not when too few experienced interviewers are available.

Details about the work of Statistics Sweden's ad hoc task forces are found in Lyberg (1980) and Lindström (1989).

5. References

Lindström, H.L. (1989): Krisgruppsarbetet och räknarexperimentet i HUT 88. Research and Development Reports, 14, Statistics Sweden (in Swedish).

Lyberg, L. (1980): Bortfallsarbetet under 1978 års hushålls-budgetundersökning. Metodproblem i individ och hushållsstatistik, 12, Statistics Sweden (in Swedish).

DRAFT, rev. 5
1991-07-17

A Study of Errors in Swedish Consumption Data

M. Ribe

1 Background

The present work was primarily called for by the appearance of some very large discrepancies between the Family Expenditure Surveys for 1985 and 1988 on the one hand, and the National Accounts on the other hand. Those inconsistencies rose questions about their probable causes, and about the adequacy of the Family Expenditure Surveys as a source for the National Accounts.

Family Expenditure Surveys for Sweden have been carried out intermittently; the three latest ones were in 1978, 1985 and 1988, and the next one is planned for 1992. For 1978 the data on consumption of the National Accounts of Sweden were adjusted, so as to agree with statistics from the Family Expenditure Survey for that year. During the following years the consumption data were updated by use of rates of change obtained from turnover statistics for retail trade. Then when a Family Expenditure Survey was again made in 1985, it turned out that for some goods, the results did not reasonably agree with the consumption data of the National Accounts. The tendencies remain for 1988. A few summary figures are given in Table 1.

There are several possible causes of the disagreement. They include sampling and nonsampling errors, both for the Family Expenditure Survey and for the retail turnover statistics, as well as differences in definitions and coverage between the two data sources. Several differences of the latter kind are known and may plausibly explain the disagreement of the figures for some goods, but for certain other goods there seems to be no such explanation, even though the disagreement is large.

So to some extent, it seems that the disagreement can only be explained by errors in data. Errors may be of importance both in the Family Expenditure Survey and in other data sources involved, such as the retail turnover statistics. Each of those various kinds of errors calls for examination.

2 Scope and idea of the present work

In the work reported here, however, we concentrate on measurement errors in the Family Expenditure Survey. Especially, an apparently crucial question is whether the character and size of the measurement errors in the Family Expenditure Survey may have changed between 1978 and 1985. In fact, the format of the questionnaire was rather substantially changed between 1978 and 1985. The question is how this change may have affected the comparability between those two years. That question also has independent interest, irrespectively of the use for the National Accounts.

To try to gain some clues on the character of the measurement errors, we may study the consistency of the data in some respects. Measurement errors in expenditure data, which are self-reported by consumers, have been rather extensively studied in previous work, in Sweden as well as in other countries. It is rather well known that there is an overall tendency to underreport expenditures, especially for certain kinds of goods. Furthermore, a recognized phenomenon is that the reporting of expenditures tends to decrease during the period of reporting for the household. Another important phenomenon is that many households choose to report for a later period rather than the prescribed period; this is also a likely source of bias.

By studying these phenomena in the response data, we may obtain some information on the structure of some major kinds of measurement errors. Of particular interest is whether this structure has changed notably between 1978 and 1985. The answer to that question may provide circumstantial evidence as to whether the measurement bias may have changed notably during the period, so as to be a likely cause of observed disagreements with the turnover statistics in rates of change in consumption of particular goods.

In order to study the mentioned error structure and its possible change over time, it should be useful to consider the dependence of the mentioned phenomena on background factors, such as age, region and composition of the household. Model-based analyses along these lines, for different kinds of goods, may reveal patterns which may be compared between the years.

3 The Swedish Family Expenditure Surveys

In the Swedish Family Expenditure Surveys, data are collected mainly by the use of diaries, where the purchases made in the respondent household are reported. The diaries are supplemented by interviews or mail questionnaires for data on some expenditures, such as purchases of more expensive items; here we will however deal only with the diaries. The use of diaries for collecting data on expenditures is very well-known and internationally widely used; cf. Kemsley (1979), Pearl (1968), Raj (1972, Ch. 15), and Sudman et al. (1971). An overview of the methodology of the Swedish Family Expenditure Surveys is given by

Näsholm et al. (1989), where particularly studies on issues of quality and response burden are reviewed.

As mentioned the format of the Swedish questionnaire was changed between 1978 and 1985. In the terminology of Sudman et al. (1971), the diary used in 1978 was a "journal diary". This means that the purchases were entered one after the other, in the order in which they were made, with all kinds of goods in one sequence.

The diaries used in 1985 and 1988, on the other hand, were "product diaries". Namely, the form was arranged by product group (kind of goods), such as food, other non-durable goods, wearing apparel (clothes and shoes) and home furnishings (furniture etc.).

In 1978 the diary had one couple of pages for each day in the period of reporting. Thus the day of each purchase is known, indicated by on which page the purchase was reported. The two- or four-week period of reporting always started on a Thursday. The product diary used in 1985 had one couple of pages for each week in the four-week period of reporting. That used in 1988 had more detailed product groups and had just one set of pages for the whole period, while the date of each purchase was to be entered by the respondent. In both 1985 and 1988, the respondent was allowed to choose the starting day of the reporting, given only that it was to start within a prescribed week and to last for four weeks.

Apart from these difference in the format of the diary, there were also other differences between the surveys of different years. In 1978 two samples were used, one with "complete" reporting, meaning that all purchases were to be reported, and one with "partial" reporting. In the latter only purchases exceeding 50 Swedish Kronor (SEK) were to be reported. (The amount of 50 SEK in 1978 is roughly 20 U.S. Dollars in 1990.)

From 1985, the reporting of food purchases is simplified. In 1978 all food items had to be specified by the respondent, like for other kinds of goods. In 1985 and 1988, however, food items were to be reported just as "food", without further specification. The specification is no longer needed, since there are now other surveys yielding statistics on the composition of the food consumption.

The sampling procedure is essentially simple random sampling on individuals. This means that the sampling probability for a household is proportional to the number of persons registered for the household. Systematic sampling is used in a population register sorted by geographical region. The possible effect of the sampling design on the analyses discussed in the present paper are presumably of minor importance and will not be dealt with further; cf. Nordberg (1989).

The main features of the Swedish Family Expenditure Surveys in different years are summarized in the following table.

Year	1978		1985	1988
	"Complete"	"Partial"		
Diary format	Journal	Journal	Product	Product
Purchases reported	All	> 50 SEK	All	All
Specification of food	Yes	Yes	No	No
Period of reporting	2 weeks	4 weeks	4 weeks	4 weeks
Starting day of reporting	Thursday	Thursday	Optional	Optional
Date/week of purchase	Date of purchase given by page of reporting	Date of purchase given by page of reporting	Week of purchase given by page of reporting	Date of purchase entered by respondent
Sample size	5700	7600	6000	6000
Response rate, per cent	72	80	73	63

4 Measurement Errors in Expenditure Surveys

Apparently, the response burden in an expenditure survey is rather heavy. Partly as a consequence of this, response errors may easily occur. It seems reasonable to assume that underreporting is the most prominent source of response error. It is easy to forget to report a purchase in the diary, as is both realized by common sense and sustained by psychology; cf. Braddeley (1979). In particular, some respondents may be reluctant to enter the purchases continuously into the diary, so that they later have to memorize the purchases when they finally fill in the diary. Such a practice is likely to further contribute to the rate of forgetting.

There may also be telescoping effects. Thus a purchase during the period of reporting may be non-reported if it was erroneously remembered to have been made before or after that period. Conversely, a purchase which was made before or after the period of reporting may be erroneously reported.

As demonstrated by Sudman et al. (1971), Näsholm et al. (1989), and Vacca et al. (1986), the diary format tends to affect the rate of reporting. Characteristics of the respondent may also influence the measurement errors; cf. Groves (1989), Sect. 9.7. Furthermore there may be an interaction between respondent characteristics and the measurement instrument, i.e., the diary, so that the record-keeping behaviour should be regarded as a product of those two factors. Tucker (1986) deals with modelling of this interaction.

Now, while underreporting and telescoping cannot be directly observed in the regular survey, there is another phenomenon which can. It is generally recognized that the rate of reporting, in terms of both number of purchases and total amount of purchase, has a tendency to decrease during the period of reporting; cf. Sudman et al. (1971), Raj (1972, Ch. 15), Kemsley (1979), Näsholm et al. (1989), and Harrison (1991). This decrement of the reporting rate seems to be usually interpreted as due to increased underreporting during the period.

The decrement in the rate of reporting is often studied in terms of the ratio between the number of purchases during the second week of reporting and that of the first. Let us call this ratio the *Reporting Rate Ratio*. Under unbiased reporting, this ratio would be expected to be equal to one. But in practice the ratio has indeed a systematic tendency to be smaller than one.

Now, the idea here is to study the behaviour of the reporting rate ratios. Since these ratios reflect measurement errors, it seems likely that substantial changes in the measurement errors could possibly cause notable changes in the observed behaviour of the reporting rate ratios.

5 Comparison of Reporting Rate Ratios

In Table 2 reporting rate ratios are given for a few aggregates of goods in 1978 and 1985. One may notice that the values for the Complete diary in 1978 are remarkably similar to those for 1985. Indeed, the similarity must be regarded as rather striking in view of the considerable difference in diary format. Even though this difference may very well entail a difference in reporting rate between the two years, the result still seems to indicate that there may not be a vast difference in performance between the diary for 1978 and that for 1985.

On the other hand, the reporting rate ratios for the Partial diary in 1978 are notably smaller, indicating a larger decrement in reporting. This tendency is also apparent in the tables of Lindkvist et al. (1980), where it is further shown that for some goods the reporting tends to be higher for the Partial diary. The interpretation of these findings does not seem quite clear. - For completeness, the numbers of purchases are given in Table 3.

6 Reporting Rate by Day

For 1978 it is possible to study the reporting rate not only by week but also by day. For purchases of everyday commodities in the Complete diary, Diagram 1 shows the distribution by day of the number of purchases. It is again seen that the first week tends to be on a higher level than the second, but that the tendency seems to be concentrated to the very first days. Apart from that, the curves are highly affected by the normal day-of-week variations.

For 1988 one may attempt to make a similar diagram. In that year's Survey, dates of purchase were entered by the respondents. There are two problems here. First, those dates were key-entered but not edited and not actually used in the normal processing, and hence their data quality is low, with a large proportion of invalid values. Second, the starting-day of reporting was optional, during a prescribed week, and hence it is somewhat uncertain what day to consider as the starting-day.

An attempt was made, however. Respondents with more than a reasonably small number of invalid date-values were excluded, and for the remaining ones, purchases with invalid dates were disregarded. The result is shown in Diagram 2. Here there surprisingly seems to be a low level of reporting for the first day or two, which must be interpreted as a consequence of the fact that the start of the reporting period is not well-defined. The data for 1988 thus seem to be difficult to analyze with respect to time of purchase, and hence data for 1988 are not considered in the rest of this paper.

7 Statistical Properties of Individual Reporting Rate Ratios

For everyday commodities, purchases are so frequent that it is meaningful to study the empirical statistical distribution of the individual reporting rate ratios k_2/k_1, where k_i is the number of reported purchases in week i ($i=1, 2$) for an individual household.

Diagram 3 shows the empirical distribution of individual reporting rate ratios in 1985. Note that the horizontal axis has a logarithmic scale. The smooth curve is the frequency curve of a lognormal distribution with the same median and interquartile range as the observed distribution; due to the logarithmic scale it looks like a normal curve.

It is seen that the observed distribution rather well follows the lognormal distribution, except for a spike at the value 1. Though spectacular in the diagram, the spike accounts for only a few per cent of the households. There is only a slight tendency to outliers. So, interestingly, we find no evidence that the decrement in reporting would be concentrated to some small group of respondents. Instead the well-behaved pattern of the distribution indicates that decrement in reporting is a normal phenomenon in the measurement procedure.

The arithmetic mean of the individual reporting rate ratios is 1.00, but this figure is apparently misleading in view of the skewness of the distribution. The median is 0.92, and the geometric mean (which is the natural location parameter for a lognormal distribution) is 0.90.

8 Modelling Reporting Rate Ratios

In a general vein related to those of Tucker (1986), and Groves (1989, Sect. 9.7), we will now try to model the dependence of reporting rate ratios on respondent characteristics. As a first approach, assume that the number k_i of purchases reported by a household in week i is Poisson distributed with parameter Θ_i. Let us then also assume that k_1 and k_2 are independent, both between households and between the two weeks.

Note however that the Poisson parameters for the number of reported purchases may vary between households. This being so, we now form a model by assuming that the ratio Θ_2/Θ_1 is a deterministic function of characteristics of the household. Namely, the dependence of response rate ratios on such characteristics is modelled by means of a logistic regression model (logit model)

$$(1) \qquad \ln \Theta_2/\Theta_1 = \Sigma \beta_i x_i ,$$

where the x_i are respondent characteristics or other factors; cf. Haberman (1974).

The assumptions just introduced may be unrealistic in at least two ways. On the one hand, there may be a stability in the purchasing behaviour between weeks, so that the number of purchases in a week would be more stable than a Poisson variable. On the other hand, there may be an opposite tendency, toward greater fluctuations in the number of purchases than in a Poisson process. Namely, the purchases need not be independent, for instance, if you buy particularly many clothes during certain periods.

To throw some light on the adequacy of the assumptions now introduced, we may consider the individual reporting rate ratios studied in the previous section. The observed variance of those ratios is 0.249. If the ratios were assumed to be ratios between Poisson distributed variables, then the variance would be expected to be approximately the squared mean ratio times the mean of $1/k_1 + 1/k_2$. That value is found to be 0.213, thus rather remarkably close to the observed variance. It is somewhat smaller, though, indicating that the assumptions might be violated by excessive instability rather than excessive stability. By common sense the latter conclusion would perhaps not be thought to be less valid for other kinds of goods than for everyday commodities.

To get a more realistic model we should have to allow for an underlying stochastic variability in the Poisson parameter (i.e., a doubly stochastic Poisson process), in such a way that Θ_2/Θ_1 need not be a deterministic function of household characteristics. To get a model of this kind we may modify model (1) into the form

$$(2) \qquad \ln \Theta_2/\Theta_1 = \Sigma \beta_i x_i + Z ,$$

where Z is a random variable, with a symmetric unimodal distribution of mean zero. The distribution assumptions on k_1 and k_2 are then to hold conditionally on Z. Further the realizations of Z are supposed to be independent between the considered kinds of goods, and independent and identically distributed between households. The maximum likelihood estimators of the parameters β_i for (1) are still valid for (2), but the variances of parameter estimators and test statistics are likely to be considerably larger under (2) than under (1).

In model (1) or (2), we are now interested in testing two hypotheses:

H1: $\beta_i = 0$

H2: $\beta_i^{(1978)} = \beta_i^{(1985)}$.

9 Application of logit model

The model (1) was estimated with the following regressors x_i:

Age (head of household)
- 16-24 (dummy variable)
- 25-44 (reference category)
- 45-64 (dummy variable)
- ≥65 "-

Type of household
- Single male (dummy variable)
- Single female (dummy variable)
- Couple (reference category)

Children ≤6 years of age (dummy variable)

Employment
- 3-level score: not employed, part-time, full time

Geographical region
- Metropolitan areas (dummy variable)
- Other major cities (reference category)
- Small towns, rural areas (dummy variable)

Socio-economic group
- 3-level score

The results are displayed in Table 4. For each of the mentioned factors (regressors), the estimated odds ratios are given, as obtained from the analyses

for each of the six goods aggregates (commodities). The odds ratios e^{β} here have, in principle, an unusually simple interpretation, namely as factors by which the reporting rate ratio is multiplied as a result of the presence of the characteristics in question. The value 1 for the odds ratio thus, as usual, means that the characteristic has no influence. The table also contains chi2 test statistics for testing the hypotheses H1 and H2.

It is seen that under the assumption of model (1), we would on the whole have to reject both H1 and H2 as generally valid. The grand total of the chi2 test variables is about twice the number of degrees of freedom. Considering model (2) as a more realistic one, we should however properly have to inflate the critical values for the tests. As is shown in Table 5, model (1) deviates rather distinctly from the saturated model, as measured by the likelihood ratios, whose values are often more than twice the number of degrees of freedom. The most important likely explanation of this deviation is seemingly the effect of the stochastic term Z in model (2), as will be further demonstrated in the next section.

Though we have not yet formally tested H1 and H2 under model (2), it thus heuristically seems likely that H1 and H2 may not be rejected under model (2). For further intuitive evidence, notice that for everyday commodities the estimated odds ratios are mostly close to 1. For other kinds of goods the estimated odds ratios fluctuate, seemingly without any sensible pattern. Simple sign statistics for the ß values confirm the impression. They are as follows:

1978	1985	
	Negative	Positive
Negative	14	17
Positive	13	16

For 1978 in Table 4, only values for the sample with "Complete" reporting are given. For "Partial" reporting, the random variability is greater and H1 is not significantly rejected, even under the assumption of model (1).

10 Test under model (2)

The reasoning at the end of the previous section may be formalized as follows. For each of the ten regressors (dummy variables and scores), consider the set of the estimated ß -values for the six goods aggregates, "standardized" by division by their estimated standard deviations (under model (1)). We may then test the hypothesis that for a given regressor i, H1 holds for each of the six goods aggregates. This can be done by a simple t-test on sample of the six standardized parameter estimates. Similarly H2 may be tested. - Note that there is still an approximation involved, in the reference to model (1) for the estimated standard devations used in the standardization of the estimated ß -values.

The results of these tests are given in Table 6. It is found that H2 is not significantly violated for any of the ten regressors. On the other hand H1 is, for one factor in 1978, and for two others in 1985. It is however particularly interesting to note that there is no significant change in the effect of any of the factors.

We can use a similar approach to test whether model (2) appears to fit the data reasonably. To that end, we compare the observed value of $k_2 / (k_1 + k_2)$ with that predicted by the model, for each "cell" of value combinations for all the regressors. Only those cells will be considered where at least five of the six kinds of goods are represented. For each such cell, form the residual for each of the five or six kinds of goods, standardize it by dividing it by its standard deviation, and then form a t-statistic by summing those five or six numbers and dividing the sum by the sample standard deviation.

It is found that for 1978 there are 96 such cells, with a variance between the t-values of 2.50, while for 1985 there are 124 cells with a variance of 1.75. Now, in a sample of about 100 variance values of independent t-distributions at four to five degrees of freedom, the expected mean is 1.6 and the expected standard deviation in 1.0 (see Cramér, 1946, Sect. 27.4), and thus the observed values are within the expected range. Even though the test is probably not too sharp it should indicate that there does not seem to be any really distinct interaction effects which would make randomness of Z a less plausible assumption.

11 Modified Jackknifing

An approach to estimate the standard deviations of the parameter estimates in model (2) is to use a form of the jackknife method; see Wolter (1985, Ch. 4). The parameters are re-estimated for replications of the sample, obtained by deletion of disjoint groups of (say) 100 households from the original sample. The variance of a parameter is then estimated by multiplication of the variance between the re-estimated parameter values by a certain factor.

Thus while the purchases are taken as observations in the parameter estimation, the households are taken as observations in the deletions during the jackknifing procedure. In this way the variability of Z is accounted for in the variance estimation and testing.

This approach has in the present context been tried only for everyday commodities in 1978 ("Complete" reporting). The results are given in Table 7. The chi2-values are here about a third or less compared to those of Table 4, and only one slightly significant violation of H1 is noticed among the ten factors. This reduction in the chi2-values is however probably paticularly effective in precisely the considered case, that is, everyday commodities in 1978. Namely, the large number of purchases there (cf. Table 3) would tend to stabilize model (1) and thus give relatively great importance to the variation of Z in model (2). It is hence not

possible to use the results of Table 7 for any conclusions by analogy for 1985 and for other commodities.

12 Effects of Delayed Reporting

Some respondent households participate at a later period than the prescribed one. Of all responding households, 24 per cent in 1978, and 14 per cent in 1985 reported for a later period. These delays may both affect the representativeness of the reporting, and possibly be linked to deviations in measurement quality. Lundquist (1991) has studied seasonal effects of the delays.

Here we take an approach to the general question of the possible impact of delays. A linear regression analysis was made on the whole sample, where the total expenditure amount for everyday commodities reported by the household was expressed in terms of the following variables:

Income
Number of persons in household
Children up to 6 years of age (dummy variable)
Number of rooms in dwelling
Geographical region.

For both 1978 and 1985, the R^2-value achieved was 0.41. Now, for the households with delayed reporting, one may compare the actually reported mean expenditure amount with that predicted by the model. For 1978, the observed value turned out to be 1.5 per cent lower than the predicted one, at a standard deviation of 1.2 percentage units. For 1985, on the other hand, the observed value was 1.9 per cent higher than the predicted one, at a standard deviation of 1.5.

This result does not exclude some possible disturbances in representability due to the delays, but on the other hand it does not signify any particular problems in representability or measurement quality because of them.

13 Conclusion

It may be concluded that the study has not revealed any particular misperformance of the main measurement instrument (the diary) used in the Family Expenditure Surveys. Methods of the kind used here, where no other external data are used for the evaluation, can of course only give circumstantial evidence on the accuracy of measurement. Still, if there were some really important deficiencies in the measurement instrument, and if they had changed over time so as to greatly change the bias, then this would not have been unlikely to show up in the analysis.

There may be reason to believe that if there were some really serious misperformance in the measurement instrument, then this would particularly have affected

certain groups of respondents, or respondents with certain characteristics. Such differences in turn would be likely to entail corresponding differences in reporting rate ratios. If some people had extremely great difficulties with the diary, they may very well have an extremely large decrement in their reporting rates.

However, such tendencies have hardly been noticed in the study. The decrement of reporting rates over the period of reporting appears to be a remarkably stable phenomenon. This phenomenon does not seem to be strongly linked to particular respondent groups or background factors, and especially no changes in the phenomenon over time seem to have such a linkage.

14 References

A.Baddeley (1979), The limitations of human memory: Implications for the design of retrospective surveys, (with Discussion by R.Thomas), The Recall Method in Social Surveys (Ed. by L.Moss and H.Goldstein), Studies in Education 9, University of London, pp. 13-30.

H.Cramér (1946), Mathematical Methods of Statistics, Princeton University Press.

R.M.Groves (1989), Survey Errors and Survey Costs, Wiley, New York.

R.B.Harrison (1991), Respondent burden and respondent fatigue, to appear at the Workshop on Methodological Problems in Household Expenditure Surveys and other Diary Surveys, Stockholm.

W.F.F.Kemsley (1979), Collecting data on economic flow variables using interviews and record-keeping, (with Discussion by B.Hedges), The Recall Method in Social Surveys (Ed. by L.Moss and H.Goldstein), Studies in Education 9, University of London, pp. 115-141.

S.J.Haberman (1974), The Analysis of Frequency Data, Chicago University Press.

H.Lindkvist and H.Lindström (1980), Jämförelse mellan fullständig (fub) och förenklad (feb) bokföring i HBU 1978, Metodproblem i Individ- och Hushållsstatistik No. 18, Arbetsmaterial (in Swedish).

P.Lundquist (1991), Seasonal variation in Swedish households' expenditure and response behaviour, to appear at the Workshop on Methodological Problems in Household Expenditure Surveys and other Diary Surveys, Stockholm.

H.Näsholm, H.Lindström, and H.Lindkvist (1989), Response Burden and Data Quality in the Swedish Family Expenditure Surveys, Proceedings of the Fifth Annual Research Conference, Arlington, U.S. Department of Commerce, Bureau of the Census, pp. 501-514. [Also as R&D Report, No. 1989:7, Statistics Sweden.]

L.Nordberg (1989), Generalized Linear Modeling of Sample Survey Data, Journal of Official Statistics, 5, 223-239.

R.B.Pearl (1968), Methodology of Consumer Expenditure Surveys, Working Paper No. 27, U.S. Department of Commerce, Bureau of the Census, Washington D.C.

D.Raj (1972), The Design of Sample Surveys, Wiley, New York.

S.Sudman and R.Ferber (1971), Experiments in obtaining consumer expenditures by diary methods, Journal of the American Statistical Association, 66, 725-735.

C.Tucker (1986), An analysis of the dynamics of the CE diary survey, Proceedings of the Social Statistics Session, Annual Meeting in Chicago, American Statistical Association, Washington D.C., pp.18-27.

E.A.Vacca and R.A.Killion (1986), Alternative diary formats for collecting expenditure information, Presented at the Business and Economic Statistics Session, Annual Meeting in Chicago, American Statistical Association.

K.M.Wolter (1985), Introduction to Variance Estimation, Springer-Verlag, New York.

TABLE 1

Private Consumption

According to Family Expenditure Survey vs. National Accounts

Billions of Swedish Kronor (GSEK), Current Prices

Commodity	Survey		National Accounts		Ratio Survey/Accounts Per Cent	
	1985	1988	1985	1988	1985	1988
Total consumption	396.6	488.6	437.8	570.2	91	86
Food (purchased)	78.7	89.4	81.9	99.8	96	90
Wearing apparel	29.6	35.5	32.1	40.5	92	88
Home furnishing goods	25.8	32.8	27.7	40.7	93	81
Leisure equipment	14.5	21.9	21.3	28.3	68	77
Pharmaceutical goods	17.4	20.9	24.5	31.6	71	66
Travel abroad	9.5	16.0	6.6	10.0	144	161

TABLE 2

Reporting Rate Ratios, Second vs. First Week

Per Cent

Commodity	Year	Type	Ratio
Everyday Commodities	78	Complete	92
	78	Partial	81
	85		90
Wearing Apparel	78	Complete	88
	78	Partial	80
	85		90
Home Furnishing Goods	78	Complete	92
	78	Partial	82
	85		88
Leisure Equipment, Jewellery, Watches, Clocks	78	Complete	94
	78	Partial	79
	85		87
Other Durable Goods	78	Complete	87
	78	Partial	74
	85		88
Pharmaceutical Goods	78	Complete	85
	78	Partial	77
	85		86

TABLE 3

Number of Purchases

During First Two Weeks

Commodity	Year	Type	Purchases
Everyday Commodities	78	Complete	404 620
	78	Partial	1 660
	85		119 010
Wearing Apparel	78	Complete	15 940
	78	Partial	11 060
	85		15 760
Home Furnishing Goods	78	Complete	8 940
	78	Partial	4 030
	85		7 290
Leisure Equipment, Jewellery, Watches, Clocks	78	Complete	15 120
	78	Partial	3 050
	85		10 110
Other Durable Goods	78	Complete	2 520
	78	Partial	850
	85		2 140
Pharmaceutical Goods	78	Complete	2 020
	78	Partial	120
	85		2 470

TABLE 4

Results of logistic regression analysis

(Note: The values for 1978 refer to the sample with "Complete" reporting.)

FACTOR Commodity	Odds ratio		Chi2		
	1978	1985	1978	1985	Change
Age 16-24					
Everyday Commodities	1.01	0.98	0.59	0.76	1.29
Wearing Apparel	1.06	0.87	0.48	2.18	2.48
Home Furnishing Goods	0.85	0.78	3.00	4.45	0.28
Leisure Equipment, Etc.	1.06	0.90	0.60	0.84	1.43
Other Durable Goods	1.10	1.02	0.27	0.01	0.07
Pharmaceutical Goods	0.73	1.36	1.63	1.43	3.04
Total (df.=6)			6.57	9.67	8.59
Age 45-64					
Everyday Commodities	1.00	0.95	0.00	11.01	8.48
Wearing Apparel	1.03	0.86	0.49	14.90	10.19
Home Furnishing Goods	0.90	0.87	3.36	5.14	0.18
Leisure Equipment, Etc.	0.95	0.94	1.63	1.50	0.02
Other Durable Goods	1.15	0.87	1.75	1.61	3.36
Pharmaceutical Goods	1.00	0.91	0.00	0.78	0.34
Total (df.=6)			7.23	34.94	22.56
Age ≥65					
Everyday Commodities	0.97	0.94	7.25	5.42	0.72
Wearing Apparel	0.82	0.68	6.86	24.16	2.89
Home Furnishing Goods	0.94	1.05	0.48	0.22	0.66
Leisure Equipment, Etc.	1.02	0.93	0.07	0.68	0.66
Other Durable Goods	1.19	0.89	0.59	0.28	0.85
Pharmaceutical Goods	1.19	0.89	1.13	0.52	1.61
Total (df.=6)			16.38	31.28	7.38
Single female					
Everyday Commodities	0.97	1.01	4.35	0.11	2.11
Wearing Apparel	1.01	0.98	0.05	0.15	0.17
Home Furnishing Goods	0.82	0.95	4.58	0.35	1.42
Leisure Equipment, Etc.	0.94	0.92	0.71	1.51	0.08
Other Durable Goods	1.12	1.03	0.47	0.04	0.15
Pharmaceutical Goods	0.78	1.18	2.58	1.74	4.32
Total (df.=6)			12.74	3.90	8.26
Single male					
Everyday Commodities	0.97	0.95	3.16	3.94	0.35
Wearing Apparel	1.36	1.04	7.71	0.16	3.04
Home Furnishing Goods	0.84	1.20	1.75	2.07	3.80
Leisure Equipment, Etc.	0.95	1.00	0.21	0.00	0.08
Other Durable Goods	1.10	1.25	0.18	0.91	0.15
Pharmaceutical Goods	1.13	1.03	0.17	0.02	0.07
Total (df.=6)			13.18	7.10	7.50

TABLE 4 (continued)

Results of logistic regression analysis

FACTOR Commodity	Odds ratio 1978	Odds ratio 1985	Chi2 1978	Chi2 1985	Chi2 Change
Children					
Everyday Commodities	1.05	0.97	9.36	0.94	5.26
Wearing Apparel	1.11	0.87	1.84	3.34	5.08
Home Furnishing Goods	1.00	1.10	0.00	0.65	0.36
Leisure Equipment, Etc.	1.06	1.02	0.64	0.04	0.12
Other Durable Goods	1.68	0.91	7.67	0.21	4.76
Pharmaceutical Goods	1.47	1.08	2.65	0.13	0.89
Total (df.=6)			22.16	5.31	16.47
Employment					
Everyday Commodities	1.01	1.00	0.42	0.06	0.01
Wearing Apparel	0.96	1.00	1.18	0.00	0.61
Home Furnishing Goods	1.00	1.02	0.00	0.10	0.05
Leisure Equipment, Etc.	0.95	1.13	1.71	5.12	6.66
Other Durable Goods	0.98	1.01	0.03	0.00	0.02
Pharmaceutical Goods	0.95	1.15	0.15	1.54	1.29
Total (df.=6)			3.49	6.82	8.65
Metropol area					
Everyday Commodities	0.99	0.98	2.94	2.21	0.23
Wearing Apparel	1.02	0.92	0.25	4.44	3.37
Home Furnishing Goods	0.89	0.97	4.38	0.23	1.14
Leisure Equipment, Etc.	0.99	0.92	0.03	2.52	1.23
Other Durable Goods	0.79	0.77	5.47	5.54	0.02
Pharmaceutical Goods	1.17	1.03	1.93	0.11	0.67
Total (df.=6)			15.00	15.05	6.66
Rural					
Everyday Commodities	0.97	0.99	15.93	0.43	1.67
Wearing Apparel	0.99	0.94	0.03	2.52	1.05
Home Furnishing Goods	0.96	1.06	0.72	1.13	1.84
Leisure Equipment, Etc.	1.04	0.97	1.12	0.35	1.26
Other Durable Goods	0.95	0.78	0.26	5.27	1.88
Pharmaceutical Goods	0.86	1.18	2.12	2.92	4.98
Total (df.=6)			20.18	12.62	12.67
Socioecon					
Everyday Commodities	1.02	1.01	2.14	0.97	0.10
Wearing Apparel	0.95	1.03	0.64	0.48	1.06
Home Furnishing Goods	1.05	1.00	0.25	0.01	0.14
Leisure Equipment, Etc.	1.26	0.99	11.18	0.09	8.73
Other Durable Goods	1.19	1.04	1.08	0.19	0.44
Pharmaceutical Goods	0.89	1.29	0.35	6.67	2.80
Total (df.=6)			15.64	8.41	13.29
(Grand total, df.=60)			(132.57)	(135.10)	(112.03)

TABLE 5

Likelihood ratio test of model (1) vs. saturated model

	Year	df.	Chi2
Everyday commodities	78	221	999.38
	85	304	420.08
Wearing apparel	78	201	539.10
	85	282	649.23
Home furnishing goods	78	191	454.17
	85	261	481.67
Leisure equipment etc.	78	206	338.98
	85	276	320.90
Other durable goods	78	165	241.54
	85	217	340.07
Pharmaceutical goods	78	149	191.18
	85	216	265.78

TABLE 6

Results of *t*-test of H1 and H2 for the commodities together, under model (2)

(Note: The values for 1978 refer to the sample with "Complete" reporting.)

Factor	*t*-values			*p*-values		
	1978	1985	Change	1978	1985	Change
Age 16-24	-0.09	-1.42	-1.01	0.931	0.216	0.357
Age 45-64	-0.29	-4.26	-2.78	0.782	0.008	0.039
Age ≥65	-0.96	-1.89	-2.26	0.381	0.117	0.073
Single female	-1.96	-0.17	1.34	0.107	0.875	0.237
Single male	0.01	0.29	0.01	0.990	0.780	0.994
Children	4.09	-0.79	-2.51	0.009	0.463	0.054
Employment	-1.21	1.75	1.98	0.281	0.140	0.105
Metropol area	-1.18	-3.07	-1.38	0.291	0.028	0.226
Rural	-1.43	-0.63	0.35	0.213	0.557	0.738
Socioecon	1.33	1.83	-0.41	0.241	0.127	0.698

TABLE 7

Results for jackknifing in model (2)

For everyday commodities in 1978, "Complete" reporting

Factor	Odds ratio	Odds ratio, 95 % confidence interval	Chi2	*p*
Age 16-24	1.01	0.95 -- 1.08	0.19	0.666
Age 45-64	1.00	0.97 -- 1.03	0.00	0.984
Age ≥65	0.97	0.91 -- 1.02	1.83	0.176
Single female	0.97	0.92 -- 1.02	1.56	0.212
Single male	0.97	0.90 -- 1.04	0.86	0.352
Children	1.05	0.99 -- 1.12	3.22	0.073
Employment	1.01	0.97 -- 1.05	0.10	0.754
Metropol area	0.99	0.95 -- 1.02	0.80	0.370
Rural	0.97	0.94 -- 1.00	4.51	0.034
Socioecon	1.02	0.97 -- 1.08	0.64	0.425

Diagram 1. Distribution of Purchases by Day
Everyday Commodities
Year 1978

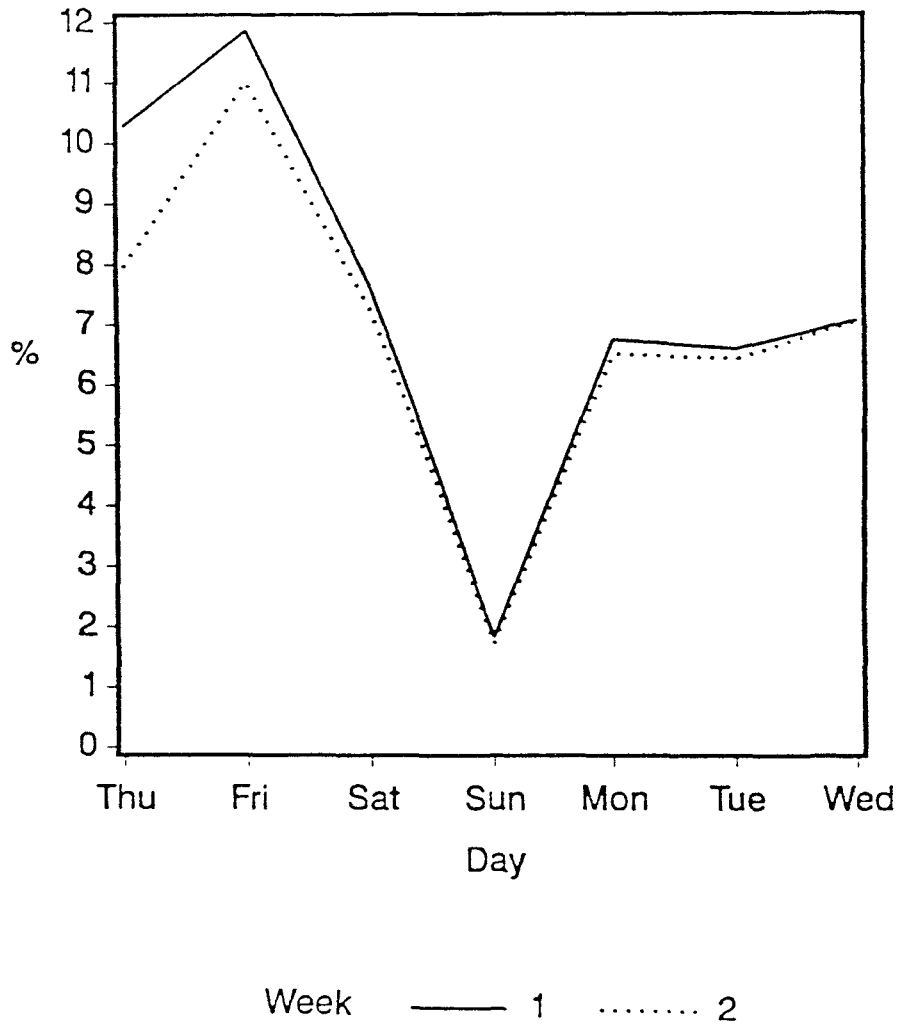
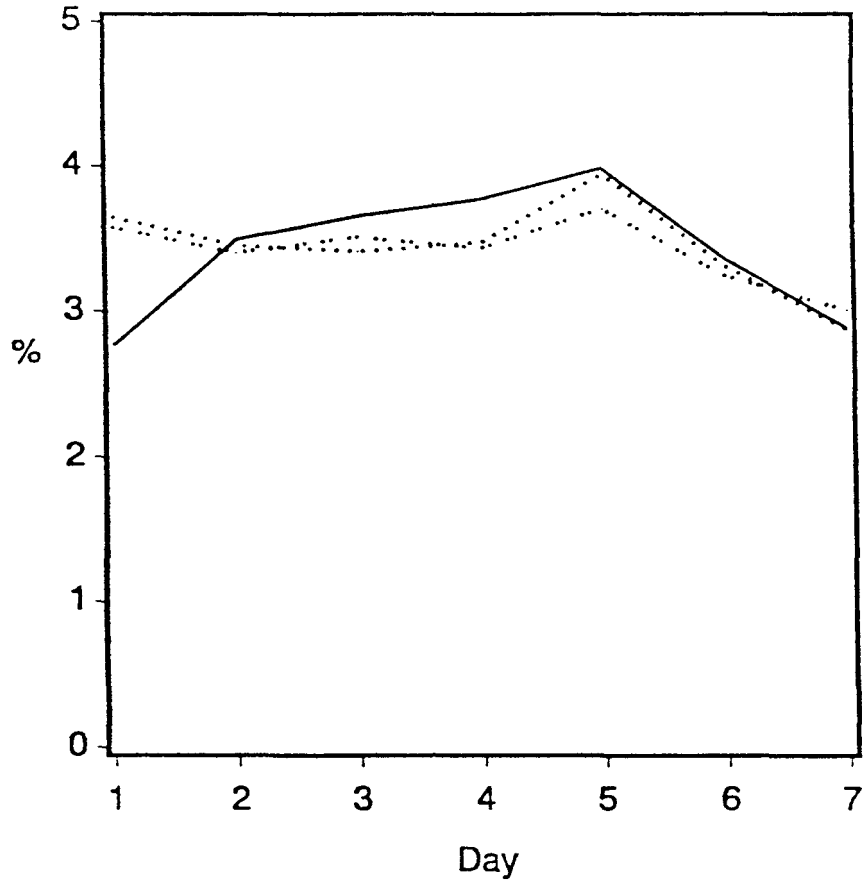
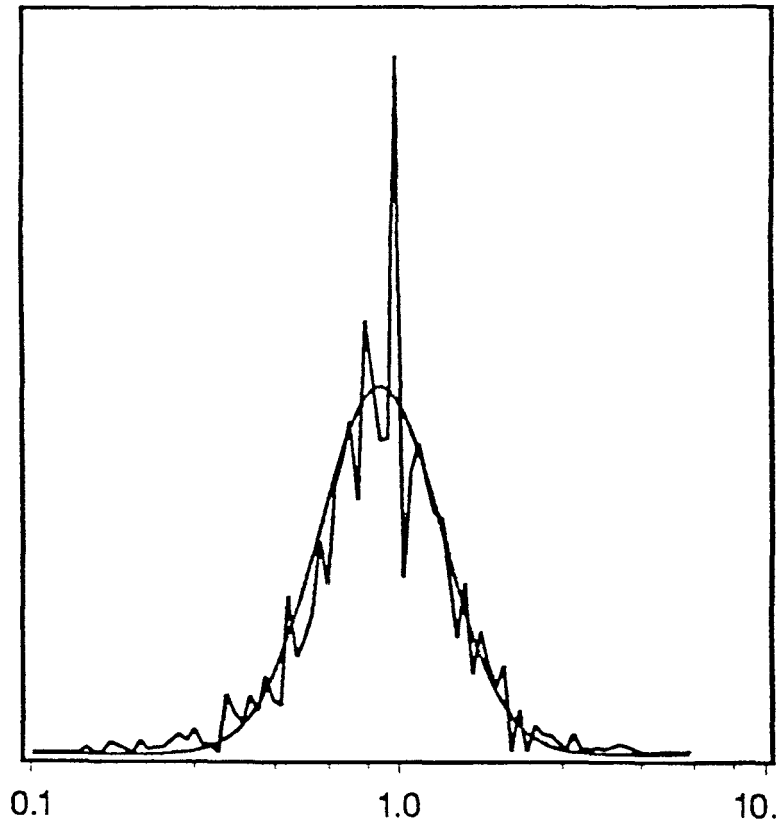


Diagram 2. Distribution of Purchases by Day
Everyday Commodities
Year 1988



Week —— 1 ····· 2 ·· ·· ·· ·· 3

Diagram 3. Distribution of Reporting Rate Ratios
For Individual Households.
Everyday Commodities



R & D Reports är en för U/ADB och U/STM gemensam publikationsserie, som fr o m 1988-01-01 ersätter de tidigare "gula" och "gröna" serierna. I serien ingår även **Abstracts** (sammanfattning av metodrapporter från SCB).

R & D Reports Statistics Sweden are published by the Department of Research & Development within Statistics Sweden. Reports dealing with statistical methods have green (grön) covers. Reports dealing with EDP methods have yellow (gul) covers. In addition, abstracts are published three times a year (light brown /beige/ covers).

Reports published during 1991:

1991:1 (grön)	Computing elementary aggregates in the Swedish consumer price index **(Jörgen Dalén)**
1991:2 (grön)	Översikt av estimatorer för stora och små redovisningsgrupper **(Sixten Lundström)**
1991:3 (grön)	Interacting nonresponse and response errors **(Håkan L Lindström)**
1991:4 (grön)	Effects of nonresponse on survey estimates in the analysis of competing exponential risks **(Ingrid Lyberg)**
1991:5 (grön)	Study of the estimation precision in the Chinese MIS surveys **(Bengt Rosén)**
1991:6 (beige)	Abstracts I - Sammanfattning av metodrapporter från SCB
1991:7 (grön)	Kvalitetsfonden 1990: Projekt - handläggning - utvärdering **(Roland Friberg och Håkan L Lindström)**
1991:8 (grön)	Tre återintervjustudier: Inkomstfördelningsundersökningen **(Jan Eriksson)**, Högskoleenkätuppföljningen **(Anders Karlsson)**, Årsarbetskraften **(Majvor Karlsson och Solveig Thudin)**
1991:9 (gul)	What metainformation should accompany statistical macrodata **(Bo Sundgren)**

Kvarvarande **beige** och **gröna** exemplar av ovanstående promemorior kan rekvireras från Inga-Lill Pettersson, U/LEDN, SCB, 115 81 STOCKHOLM, eller per telefon 08-783 49 56.

Kvarvarande **gula** exemplar kan rekvireras från Ingvar Andersson, U/ADB, SCB, 115 81 STOCKHOLM, eller per telefon 08-783 41 47.