

PROMEMORIOR FRÅN P/STM
NR 24

NONRESPONSE RATES IN 1970 - 1985 IN SURVEYS OF
INDIVIDUALS AND HOUSEHOLDS

AV HÅKAN L. LINDSTRÖM OCH PAT DEAN

INLEDNING

TILL

Promemorior från P/STM / Statistiska centralbyrån. – Stockholm : Statistiska centralbyrån, 1978-1986. – Nr 1-24.

Efterföljare:

Promemorior från U/STM / Statistiska centralbyrån. – Stockholm : Statistiska centralbyrån, 1986. – Nr 25-28.

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

R & D report : research, methods, development / Statistics Sweden. – Stockholm : Statistiska centralbyrån, 1988-2004. – Nr. 1988:1-2004:2.

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

Promemorior från P/STM 1986:24. Nonresponse rates in 1970 – 1985 in surveys of individuals and households / Håkan Lindström; Pat Dean.
Digitaliserad av Statistiska centralbyrån (SCB) 2016.

PROMEMORIOR FRÅN P/STM
NR 24

NONRESPONSE RATES IN 1970 - 1985 IN SURVEYS OF
INDIVIDUALS AND HOUSEHOLDS

AV HÅKAN L. LINDSTRÖM OCH PAT DEAN

NONRESPONSE RATES IN 1970 - 1985
IN SURVEYS OF
INDIVIDUALS AND HOUSEHOLDS

Håkan L. Lindström and Pat Dean

Abstract

Nonresponse is an important determinant of the quality of survey data. To evaluate the developments of nonresponse rates over time, we have collected nonresponse figures from Swedish surveys of households and individuals and illustrated these figures in time series. After having taken consideration to the changes in the design of each survey, we found trends that seem to stem from general causes. We discuss to what extent these trends are the result of a changing survey climate and to what extent they are the result of changes in Statistics Sweden's internal organization and data collection methods.

CONTENTS

- 1 Introduction
 - 1.1 Statement of purpose
 - 1.2 Background
 - 1.3 Principles of analysis

- 2 The regular surveys
 - 2.0 Contents
 - 2.1 The Income Distribution Survey
 - 2.2 The Labor Force Survey
 - 2.3 The Survey of Living Conditions
 - 2.4 The Consumer Buying Survey
 - 2.5 The Omnibus Survey
 - 2.6 The Party Preference Survey

- 3 A summary of the present state of nonresponse reporting

- 4 Findings
 - 4.0 Principles of presentation
 - 4.1 Total nonresponse
 - 4.2 Refusals
 - 4.3 Not-at-home
 - 4.4 Nonresponse - total and by reason among men and women
 - 4.5 Regions and reason for nonresponse
 - 4.6 Family type and nonresponse
 - 4.7 Summary of observations

- 5 The general factors that explain variation over time

- 6 Future studies of changes in the survey climate

- 7 References

1 INTRODUCTION

1.1 Statement of Purpose

Maintaining the confidence of potential respondents to reach an acceptable response rate is of utmost importance for every survey organization. We will use "survey climate" as a general term for the opinions and attitudes that determine the sampling population's propensity to respond to a survey.

When the rate and distribution of nonresponse changes the reason may be persistent changes in the survey climate. There may, however, also be other explanations, for example that the execution of the data collection has changed in either a planned or uncontrolled way. The survey climate can influence not only nonresponse errors, but also measurement errors, the amount of effort necessary to get an interview or a mail response and the cost that accompanies these efforts. However, since changes in nonresponse rates are the easiest to calculate there is good reason to keep an eye on them and try to explain their variation.

In the last decade there has been great variation in the nonresponse rates in surveys of individuals and households. Within this period several studies have been made at Statistics Sweden on the effects of nonresponse on survey results. There has also been a continuing process of surveying and improving the data collection methods. This report supplements earlier studies by thoroughly documenting nonresponse rates in 1970 - 1985 and thereby providing the means of studying issues related to both temporary and longlasting changes in nonresponse. The purpose of this study is :

- a. To describe the present state of nonresponse reporting and suggest improvements for the calculation and presentation of nonresponse.
- b. To collect data for times series and comment on their development and provide feasible explanations for these developments.
- c. To make recommendations on the need for a regular monitoring on nonresponse in the regular surveys of individuals and households. And if such a need is found, to recommend how this monitoring should be carried out.

Our primary interest lies in establishing to what degree the nonresponse rates of similar surveys move in unison and to identify the factors that contribute to such variations. For example, what part of the nonresponse is attributable to changes in the efficiency of our data collection agency and what part reflects changes in a general survey climate?

1.2 Background

Nonresponse rates are reported in most surveys at Statistics Sweden. The reporting functions:

- * as a measure of the quality of the survey
- * as a monitor on the efficiency of the data collection procedure
- * to facilitate the analysis of errors

To look at each survey individually is certainly inadequate for an analysis of changes in the survey climate, as there can be many reasons why one survey would have variations in its nonresponse rates. In order to get a general view of the changes in the survey climate, one must make a detailed and parallel study of the occurrence of nonresponse in several surveys of the same population. The surveys that we have chosen to study have the following in common:

1. The object of interest is the individual or the household
2. The sampled population is the Swedish population within some age limit, usually 16-74
3. Each survey is performed at least once a year
4. Participation in each survey is voluntary
5. All data collection has been performed by Statistics Sweden

Ideally the surveys' nonresponse reporting should be comparable and both calculated and presented according to the same principles. This was, however, not the case. There has been no coordination among the surveys, so the results are often not as comparable as one would wish. Observations about nonresponse trends in domains of study based on one survey could seldom be confirmed or disproved by another survey because few of the surveys reported comparable nonresponse rates. This leaves several questions open for speculation and makes several conclusions weaker than they

otherwise would have been. However, to make comparable calculations was not within the scope of this study.

1.3 Principles of Analysis

When we study the changes in the time series of a survey's nonresponse rate the explanations fall into three main categories:

- * changes in the survey climate
- * changes in the organization and/or efficiency of data collection
- * changes in the specific survey's design.

It is important that we do not mistake variations that have "technical" sources for true variations in nonresponse over the studied time period. A technical source of variation would be any changes in the planning stage or in the data collection method that could lead to changes in the response rate, for example: a redesign of the sampling plan, the amount of effort spent on reaching a specified response rate, the competition among surveys for data collection resources, and the time restrictions placed on the data collection agency.

We regard the nonresponse rates and especially the refusal rates as important indicators of the survey climate. We believe that increases in the nonresponse rates reflect an increasing resistance to surveys. Whereas changes in data collection procedures are mostly intended to reduce nonresponse rates or to counteract expected increases.

We do not assume that the nonresponse rates of the various surveys will react in exactly the same way to a changing survey climate. A survey's established level of nonresponse, its subject, burden of response, etc, all effect the way a given survey will respond to a change in the survey climate.

Explaining differences in the surveys' response levels and distributions lies outside the scope of this report. We will not devote much discussion to analyzing how differences in the surveys' choice of data collection method, duration of the data collection period, response burden, ability to motivate the interviewees to take part in the survey etc affect their nonreponse level.

2 THE REGULAR SURVEYS

2.0 Contents

The nonresponse rates presented in this study are drawn from:

- * The Labour Force Survey (LFS)
- * The Survey of Living Conditions (SLC)
- * The Survey of Consumers Buying Expectations (CBE)
- * The Party Preference Survey (PPS)
- * The Omnibus Survey
- * The Income Distribution Survey (IDS)

The classifications used to describe a nonrespondent are: refuser, not-at-home, sick at home, or under institutional care. The two last cases are of minor importance and not further considered in this report. The nonresponse rate is calculated from the net sample. The sampling frame is reliable and any over-coverage is small and easily identified. Since 1985, new routines have made a more detailed classification of nonrespondents accessible so future analysis of nonresponse can be more explicit.

The surveys were far from being equally detailed in their reporting of nonresponse in the domains of study. This lack of comparable data limited the scope of our analysis in many ways. This made interpreting the differences that do exist even more complicated.

To aid the reader in a critical examination of the time series, we summarize the main features of each survey, mention important redesigns and describe the standard method of reporting nonresponse. The data used in this report are compiled from both published and unpublished sources. In section 6 we discuss the problems and future options for a more detailed nonresponse analysis of both past and present surveys.

2.1 The Income Distribution Survey (IDS)

Until 1984 the IDS was a mail survey performed in November and December with a telephone follow-up in January among those who did not respond. In 1984, the survey went from using a mail survey with a telephone follow-up to using only telephone interviews. Interviewing now occurs in January and February after the income year.

Households are asked to report their family composition, type of housing and their occupations. Income data is obtained later from the public tax return registers and other registers. The total sample size is approximately 10 000 households. Half of these are newly sampled and the other half retained from the previous year. The sample is heavily stratified according to source of income, income size, age and family type. Some minor redesigns of the survey have been done. Stratification and allocation of the sample have varied a bit between the years. The principles for data collection have also changed. In 1980 and afterwards, households who refused in the first year were automatically classified as refusers also the second year if they did not respond by mail.

The IDS does not publish information on nonresponse. Unweighted nonresponse rates are calculated for the survey as a whole and can be obtained through personal contacts with the IDS staff. No consideration is given to the varying sampling probabilities. Nonresponse rates for each of the sampling strata can be obtained through extra effort. Since 1984, nonresponse reduction efforts have been directed towards strata with high nonresponse rates.

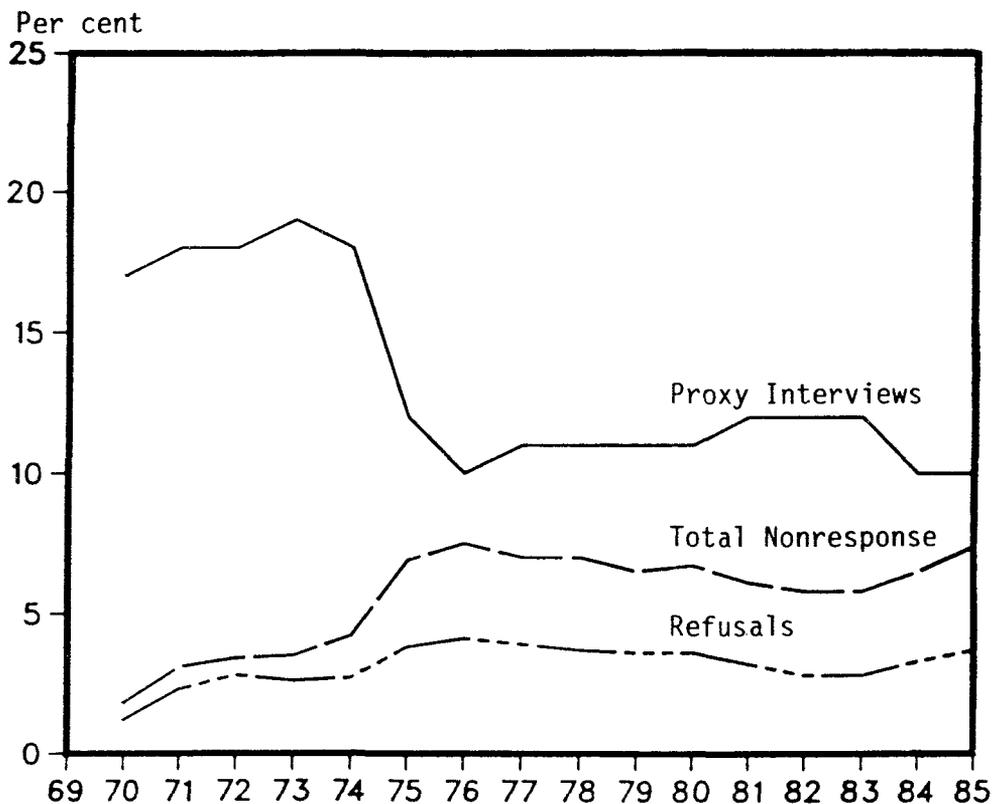
An extensive recalculation must be made if the nonresponse rates of the IDS are to be made comparable with probability weighted samples or with self-weighted samples. It is not known to what extent changes in the stratification have affected the presented nonresponse rates of IDS presented in figures 3, 4 and 5, which are unweighted.

2.2 The Labour Force Survey (LFS)

Until 1986 the Labour Force Survey was restricted to individuals in the ages 16-74 until 1986. It has a sample of about 20,000 individuals. The survey was performed each month, using a rotated sample with eight panels. For a sampled individual, the time lapse between two consecutive interviews is three months. Interviewing is done by telephone and takes about fifteen minutes. The data collection period is one week. Proxy interviewing is done with about one tenth of the net sample. Until 1975, all individuals had equal sampling probabilities, even though the sample was stratified. Since 1976, the sample has been stratified according to county, sex, and marital status. A disproportionate allocation for counties has been used. Individuals in counties with small populations are assigned higher sampling probabilities to ensure a minimum precision in each county.

Within each stratum, a systematic sample is drawn. The overall nonresponse rates for the period 1976-85 are somewhat lower than they would have been had the sampling design been the same as in the years before 1976 or after 1985. Since the sparsely populated counties were assigned higher sampling probabilities and have lower nonresponse rates than more populated areas, the use of unweighted nonresponse rates from 1976 to 1985 must have led to a lower overall nonresponse than otherwise would have been the case.

Fig.1 Nonresponse and Proxy Interview Rate in LFS.



However, it is not possible to isolate the effect of the 1976-85 design change since proxy interviewing was restricted in 1975 and this restriction is known to have countered the lower nonresponse that should have resulted from the new stratification. When the new proxy interview regulation came into effect, the nonresponse rates went up from 4.2 per cent in 1974 to 6.9 percent in 1975. Current regulations limit proxy interviews to members of the immediate family only. The interaction between proxy interviewing and nonresponse rate is shown in figure 1.

starting in 1986, LFS will exclude the age class 65 - 74 and go back to equal sampling probabilities for all individuals aged 16 - 64. The sample size has been reduced to 16 000.

The LFS reports nonresponse under the heading " The Sample" together with the results of each survey. Nonresponse is divided up according to reason, age, age and sex, and region (Stockholm, Göteborg, Malmö, and the rest of Sweden). The reporting has been consistent since at least 1970, except for a redefinition of regions in 1976.

2.3 The Survey of Living Conditions (SLC)

The SLC is a survey of individuals that started in 1974. The topic of the survey changes according to a rotating scheme. Data collection is done mainly by personal interviews lasting approximately one hour. There is also a low rate of proxy interviews. Proxy interviews are predominantly done among the elderly and those who have poor health. The length of the data collection period is at least three months. The ages included in the sample were 16-74 until 1980 when the upper age limit was raised to include 84 years old.

Major redesigns of the sampling plan were performed in 1977 and 1980. In 1975 and 1976 a simple random sample of persons (directly sampled) was drawn. The spouses (indirectly sampled) of those chosen for interviews were also interviewed. In 1977-79 the sample was stratified and unmarried parents with children younger than 16 were heavily over-represented in the sample. They then constituted about 1/24 of the age group under study. Because of the difficulties caused in both the field work and in technical matters, SLC chose to abandon spouse interviews in 1980, which surely has led to a decrease in nonresponse. Since 1980 there have been two age strata : 16 - 74 and 75 - 84.

The nonresponse rates presented here for the SLC represent individuals ages 16-74 who were directly sampled in a simple random sample. The sample size for those directly sampled has varied between 4800 and 8800 a year. They estimate the population values unbiasedly except in 1977-1979, when unmarried parents had to be excluded from the calculations for technical reasons. Still, design changes influence the data collection methods which in turn affect the final results. On two occasions, 1976 and 1983, experiments embedded in the survey design led to an increase in the nonresponse rates.

The technical reports of the SLC present nonresponse information only for the sample as a whole. More detailed information must be sought in unpublished internal reports. However, each of the unpublished reports was adequately detailed. Nonresponse was reported according to reason, demographic variables, income, and region. The available nonresponse data are sometimes incomparable from one year to the next because of changes in design, redefinition of variables, and inconsistent subgrouping. Nonresponse reporting on age, region, income, marital status, and nationality were not comparable because of inconsistent classification. For example, there was one age classification used from 1974-76, a second from 1977-79, and a third from 1980-85.

2.4 The Survey of Consumer Buying Expectations (CBE)

From 1973 to the present, the Survey of Consumer Buying Expectations has been performed on a quarterly basis. Interviewing is done in January, April, July, and October of each year. Data concerning households is collected via telephone interviews during a two week period each quarter. Each telephone interview takes less than ten minutes. The number of questions was reduced in July 1976 and has since been further reduced in July 1985.

CBE's sample consists of 6000 households divided into six income strata. Strata divisions are established according to income data from the Register of the Total Population (RTP). The strata limits are adjusted each year for inflation. The sampling fraction is greatest among households with very high and very low incomes. Originally a rotation system consisting of five equally large panels was used; at each survey, one panel was dropped and a new one replaced it. For financial reasons, the sample size has been reduced since July 1984 and the old panels abandoned. Since July 1985 the sample consists of 1500 households in January and July with a rotating system of three panels and of 4200 in April and October with an equal rotating system. In all other respects, the sampling design is unchanged. In July and October 1985 the entire sample consisted of new panels.

Some of the differences observed in the CBE before and after the redesign in 1984 have a purely technical explanation. Usually the longer a panel has been in a sample, the more its not-at-home and refusal rates move in opposite directions. The longer the panel members participate in the survey the more they will tire of the

the survey and refusals will increase, but not-at-homes will decrease because the interviewers will have more time to find and interview the difficult to locate respondents.

Unweighted nonresponse rates are reported by reason and panel in a uniform way. This survey is the only one that presents time series of its nonresponse rates. The information can be found in internal unpublished reports. The reports cover all of the quarterly surveys, but are restricted to the sample as a whole. How stratification and disproportionate allocation of the sample influence the unweighted nonresponse rate has not been evaluated. Nonresponse rates are also calculated for each: stratum, county, age group (age of the head of the household), region (greater Stockholm, greater Göteborg, greater Malmö, and the rest of the country), age and region, and stratum and region. These data are neither compiled nor reported but available on the original computer sheets in the archives. The reporting for stratum, county, age groups, etc, is consistent for the entire period. However, documentation is missing for a number of quarters and for this reason, time series on this material could not be presented here.

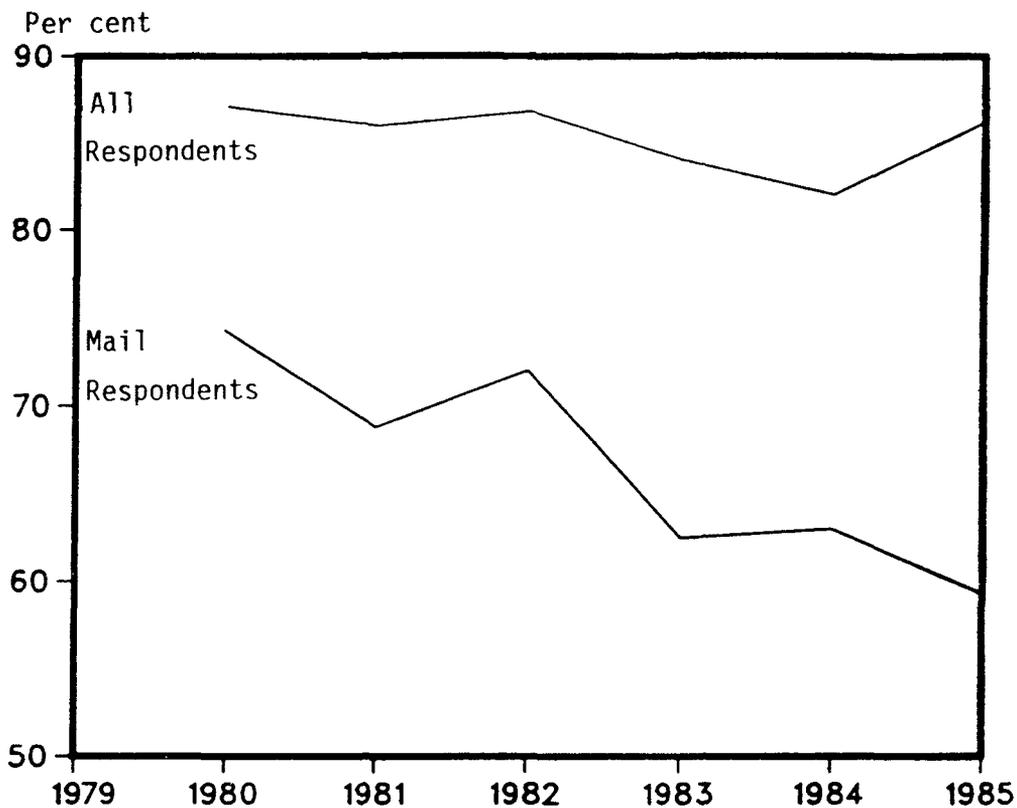
2.5 The Omnibus Survey

The Omnibus Survey is a multipurpose survey that has been performed three to five times a year since 1980. Since mid 1983, Omnibus has employed simple random sampling using a sample of about 1200 individuals in the ages 16-74. In the earlier samples, the sample size was approximately 800. Omnibus is a mail survey with a telephone follow-up within a subsample of those who do not answer by mail. For the Omnibus survey, nonresponse data are sparsely documented in the technical reports and in such a manner that the existing data cannot be further analyzed. The weighted total nonresponse rate is reported, but it is not easily connected with the table that presents the results of the data collection. A second table containing response and the sample's distribution on demographic groups cannot be recalculated to yield data on the nonresponse rates for these groups. The weighted results used in this report are summarized in figure 2.

The variations in total response may be explained partly by the fairly small sample size and the variety of topics dealt with. Although there has not been a great decrease in the total level of nonresponse, there has been a steady decrease in the number of

respondents who answer the mail survey directly without a follow-up telephone interview. In 1985 the mail response rate was 15 percent lower than in 1980. The reasons for this remarkably strong decrease should be further investigated. Either it is a result of hastening the data collection or a consequence of a worsening survey climate; costs increase when the number of follow-up telephone interviews must increase to maintain the quality of the survey.

Fig. 2 Response Rate for All Respondents and Mail Respondents in the Omnibus Survey



In this report the Omnibus rates will not be compared with those of the other surveys. The combination of Omnibus's variety of subjects, small sample size and the sparse nonresponse reporting proved an inadequate basis for comparisons.

2.6 The Party Preference Survey (PPS)

The Party Preference Survey started in 1972 and is performed twice a year - in May and November. A third survey was performed in February of the election years 1973, 1976, and 1979. After an intermission from May 1981 to May 1984, the survey has again been performed on a regular basis. The sampling units are all individuals who are eligible to vote, ie, all Swedish citizens over age 18. The sample size has consistently been approximately 9000 persons. Each sampled person is included in three consecutive surveys, and one third of the sample is newly sampled and the other two-thirds retained from the previous survey. Data is collected through telephone interviews during a two and a half week period. The sample is drawn by simple random sampling so no weighting problems occur when nonresponse rates are calculated. The absence of an upper age limit in the PPS probably contributes to an increase in nonresponse rates since very old people often have higher nonresponse rates than those in the more commonly sampled ages - 16-74.

Furthermore, the PPS performs a methodological study in every election year. The sample size is 3000 and no panel is used. The data collection period is two weeks long. The nonresponse rates of the methodological study are not included in the calculations of this report.

Nonresponse is extensively reported along with the results of the survey in Statistical Reports, series Be, and most recently, in Statistical Press Releases. The refusal rate and the not-at-home rate - but not the total nonreponse rate - are reported for each of the following variables: age, sex, age and number of children, region, new voters, marital status, income, occupational groups, counties, and type of living accomodation.

The nonresponse reporting of the PPS is outstanding. It is far more detailed and consistent than in any other survey. Even standard errors of the nonreponse rates are calculated. Equal response probabilities are assumed in each calculation. The results are easily accessible and easily interpreted.

3 A SUMMARY OF THE PRESENT STATE OF NONRESPONSE REPORTING

The importance put on analyzing the distribution of nonrespondents varies greatly from survey to survey. The Omnibus survey and the IDS calculate only the total nonresponse for the sample. The LFS, the SLC, the PPS and the CBE are more detailed. The LFS reports nonresponse rates by sex, four age groups, four regions and rotation groups. In each case, nonrespondents are categorized by not-at-homes, refusals and other nonrespondents. The SLC reports nonresponse rates by sex, marital status, seven age groups, six regions, income classes and nationality. Only the total nonresponse is reported by reason.

The most detailed nonresponse reporting is done in the PPS where the nonresponse is reported by refusals and nonresponse of other reasons for each class by age, sex, marital status, type of family, region and income class. The CBE is the only survey where they regularly compile time series of the nonresponse rates. Total nonresponse and nonresponse by reason and panel are calculated each time the survey is taken.

LFS and PPS are the only surveys that present detailed nonresponse in their reports. The other surveys report only their total nonresponse and possibly, the total nonresponse rate by reason. Anyone interested in more detailed information must seek it in unpublished technical reports, or even the personal working papers of the statistician responsible for that particular survey.

In addition to being difficult to obtain, those data that can be found are in two main aspects unsuited for comparisons with other surveys. First, almost all nonresponse rates are calculated from the unweighted sample data. If an equal selection probability method (epsem) sample is not used, comparisons of nonresponse rates are almost impossible to make because of the different effects introduced by the different sampling designs. When the sampling probabilities are not taken into consideration, it is impossible to distinguish a real jump in nonresponse from the effects of a change in sampling design. We looked at a few examples of design changes in chapter two. Second, it is an exception rather than a rule that the surveys present distributions of the nonresponse on the same classifications.

4 FINDINGS

4.0 Principles of Presentation

For those surveys that are performed several times a year, we have recalculated the seasonal or monthly data to yearly rates. Otherwise the main traits of the time series could easily have been dimmed by varying periodicity and a sometimes large seasonal variation. With yearly rates we mean the surveys' unweighted averages for each year.

In this report, the results are presented only in diagrams. Note that the scale of the vertical axis (per cent) is not the same in all figures. The scale has been chosen to make the times series of each figure readable. The original data will be given in a special technical report. The period we concentrate on is 1976 to 1985. By 1976, all the studied surveys were in production and had left their initial and, possibly, experimental stages. A reshaping of the interviewer staff was by that time on its way towards fewer interviewers and an average number of hours worked that is not much below that of today. By 1976, the various surveys were also feeling the full effect of the then new regulations set up by the data inspection act.

First, we present time series of the total nonresponse, refusal, and not-at-home for each each survey as a whole. When data are available, we present and discuss nonresponse in some important subgroups. The division of the samples into subgroups has two main purposes. The first is to find out if the rates in the subgroups behave like the population rates or if the rates of some groups are more affected by variations in the survey climate than others. The second is to identify subgroups where nonresponse is extremely high and prepare a discussion concerning if and how these rates can be reduced.

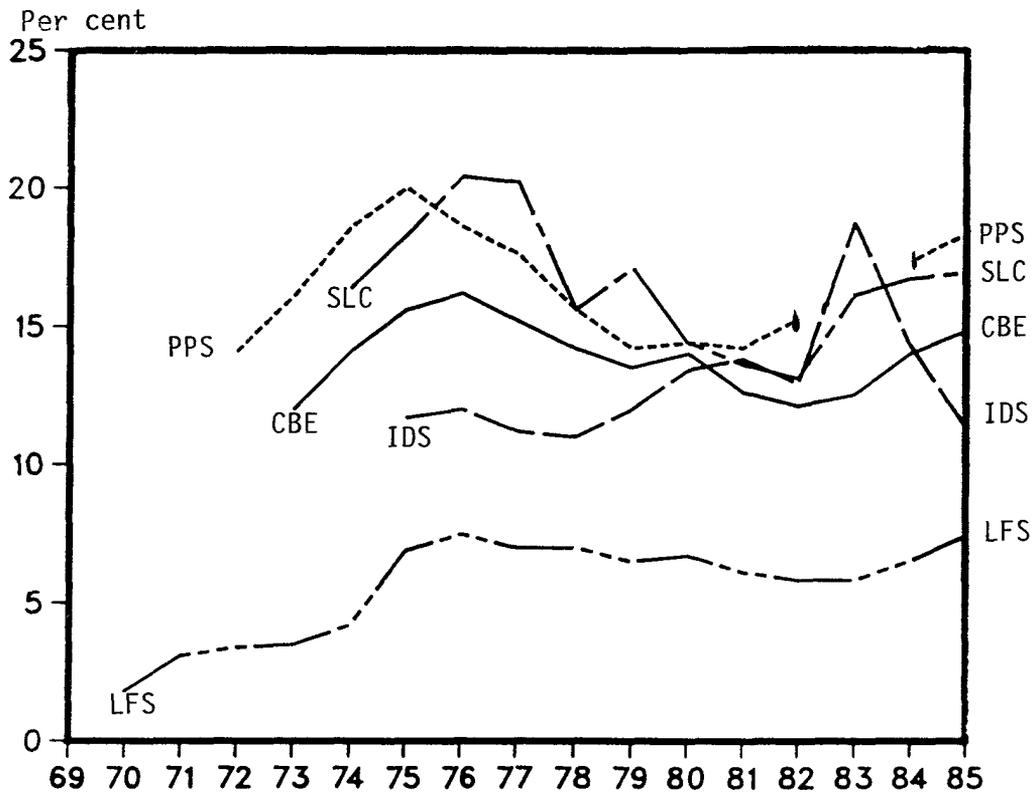
The only time series that has not been disturbed by any design changes whatsoever is the LFS during the years 1976-1985. Consequently they are more reliable than other surveys in reflecting true changes in the survey climate. In the other time series special attention should be given to the years subsequent to a redesign or intermission. The CBE has shortened its questionnaire and changed its rotation system. The PPS was not performed in 1982 and 1983. In the IDS, the data collection procedure was altered on several occasions to a greater or lesser extent and in the SLC there were major redesigns in 1977 and 1980.

The remaining sections of this chapter are mostly descriptive. The effects of design and organizational changes are discussed. However, the discussion of the influence of general factors is saved until chapter five. We do no significance testing since it would have required a considerable amount of calculation. We have restricted ourselves to comments on persistent trends and differences.

4.1 Total Nonresponse

Only the Labour Force Survey contains data collected previous to 1972. All the studied surveys started up in 1975. Although all comparisons must be made with caution, as emphasized before, we observe in figure 3 a pattern common to the nonresponse rates in four of the surveys for the period 1976-1985.

Fig 3. Total Nonresponse in Five Surveys



After an increase in nonresponse rates starting in the first year of each survey and lasting until 1976, there is a slow decrease that reaches its minimum in 1981- 82. It is notable that in the years 1980-1982 - with the exception of the LFS - the differences in nonresponse rates between the surveys are almost negligible.

The beginning of another increase is seen in the last few years. When the nonresponse rates increased after 1982 the SLC and the PPS rates went up a lot more than the others. However, they also decreased more than the others up to 1981-82. The present level is still below that of the mid-seventies.

IDS is an exception and has a comparatively constant rate in 1975-79. Giving up attempts to reach the refusers of the old panel in their second year of participation may have contributed to the higher level since 1980.

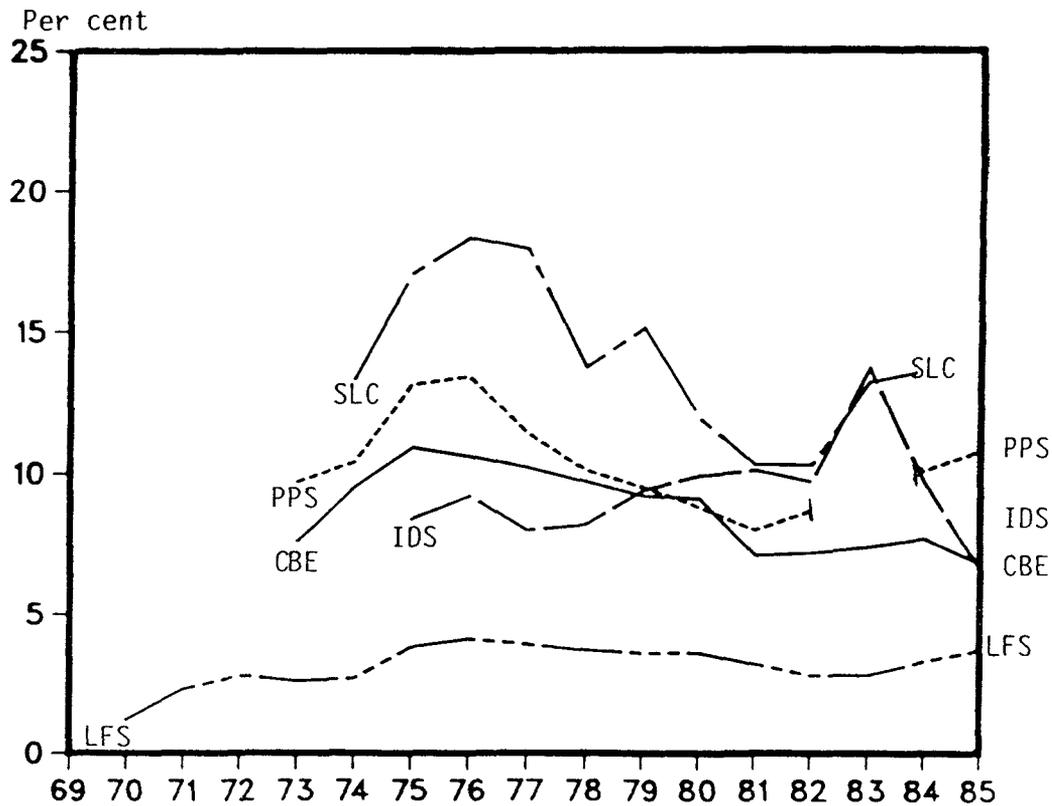
A lower propensity for nonresponse is indicated for the pre-1975 years. The early LFS has nonresponse rates that are about three percentage points lower than the rates after 1975. The early CBE and PPS had nonresponse rates equal to the low values of 1980-82.

4.2 Refusals

Refusal is generally the most important reason for nonresponse. The LFS is the only survey in which the refusal and the not-at-home rates are pretty much the same. Refusals are especially dominant in the SLC and the IDS. Refusal rates change over time in much the same way as total nonresponse rates as figure 4 illustrates. There was unquestionably an increase in refusals in the mid-seventies and a drop to very low levels in the early eighties that holds true for all the surveys except IDS. The great decrease in the SLC's refusal rate from 15.1 in 1979 and to 11.9 in 1980 occurred at a sample redesign and was probably mainly the result of the redesign. The largest change among the other surveys that year occurred in the PPS and was only 0.7 per cent. In the remaining four interview surveys, the refusal rates have still been lower in the last few years than they were in the mid-seventies, although they all show increased refusal rates after 1982.

IDS is different from the other surveys in that it has a comparatively constant nonresponse rate in 1975-79. Since 1979 there has been a steady upswing in nonresponse that was broken at the redesign of the survey in 1984. In 1983, IDS's refusal rate took a big jump. Those responsible for the survey hold the opinion that the increase in 1983 was brought about by the then current debate on confidentiality. In 1985 the level is back to 1975 level.

Fig. 4 Refusals in Five Surveys

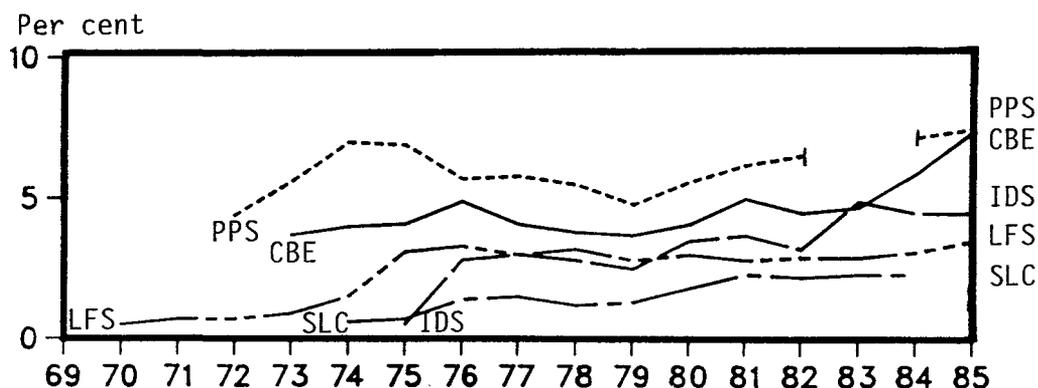


The CBE is an exception in that no notable increase in refusals occurs after 1982. This may be explained by the redesign of the survey that took place during this period. The relatively lower refusal rates in the 1985 CBE are not surprising when one considers that the rotation system was revised so that the entire sample was freshly sampled during this period. With the original design, the average refusal rates were one or two percent above those of the new panels.

4.3 Not-at-Home

Figure 5 shows a weak but pervasive increase in the not-at-home rates for all four surveys. One can also see that not-at-home rates vary less than the refusal rates. In the Labour Force Survey, there is an almost constant rate from 1975 to 1984. The SLC rates increase slowly, but this may be at least partially explained by the subsequent redesigns in 1976 and in 1980. No obvious change is seen in the CBE before 1984. The somewhat high 1985 value in CBE and the substantial increase in not-at-home rates of the PPS after the intermission from 1981 to 1984 is to be expected after the redesigns of the surveys and the introduction of new panels. If the CBE and PPS rates remain at the 1985 level in the coming years there will be reason to look for other explanations. An increase in the IDS is apparent but may be the result of successive changes in data collection methods. In 1979, all five surveys had very low not-at-home rates; the data collection routines used that year deserve special consideration.

Fig. 5 Not-at-Home in Five Surveys



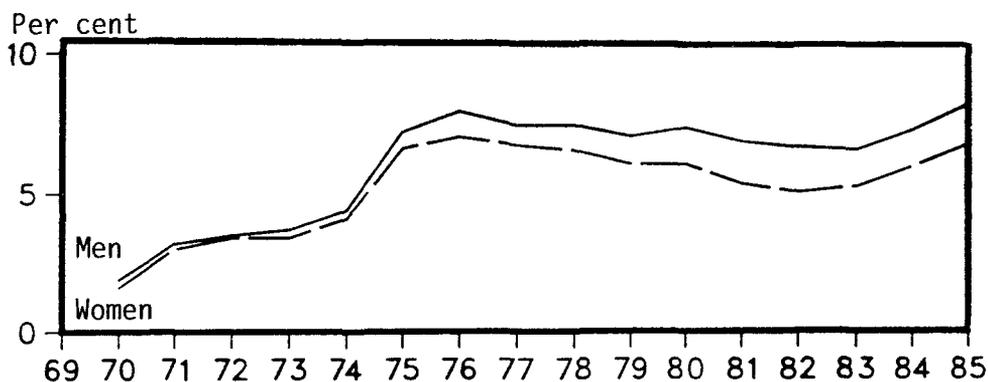
The different levels in the surveys' not-at-home rates have a number of different explanations. The main rule is that the longer the data collection period, the lower the not-at-home rates. This is the reason why the SLC, with a data collection period of three months or more has the lowest rate. The use of proxy interviews, as in the LFS, also reduces the not-at-home rate. With a rotating sample, there is a tendency for the not-at-home rate to be lower for a panel that has undergone a number of interviews. The LFS with eight panels (and proxy interviews) have around 1.5 percent lower not-at-home rate than the CBE with five panels up to July 1985. The PPS with three panels have four to five per cent higher not-at-home rates than the LFS. Evidently, the chances of reaching an interviewee increase with the number of times he or she is included in the survey. This improvement of the non-response rate is often counteracted by an increasing tendency among those who have participated a number of times to refuse more often. When the design is changed so that the degree of rotation is increased there will be more not-at-homes and less refusals. The opposite holds true if the degree of rotation is decreased.

Although the 1985 not-at-home rates are higher than those in 1976, there is still no conclusive evidence that they are increasing as a consequence of external factors like unlisted telephone numbers, locked entrances to apartment buildings, smaller families and simply that people spend less time in their homes. The extremely high number of interviewer worked hours in 1984/85 (1107 per interviewer) left less time to locate each sampled person or household. This may have contributed to the increase in not-at-home rates during those years.

4.4 Nonresponse, Total and by Reason Among Men and Women

Total nonresponse rates for each sex were readily available for the LFS, the SLC, and the PPS. Their time series are presented in figures 6, 7, and 8.

Fig. 6 Total Nonresponse Among Men and Women in LFS



Three important findings are apparent from the diagrams where the samples have been divided according to sex. First, there is no general distribution of nonresponse for men and women valid for all three surveys. In the LFS, men are more often nonrespondents than women. This is quite contradictory to the corresponding nonresponse rates for the PPS. In the SLC, there is no difference in nonresponse between the sexes. This lack of a consistent pattern is the consequence of the interviewee's being influenced by both the subject and the data collection method. For example, without the use of proxy interviews, there would have been a different nonresponse distribution in the LFS.

The second observation is that within one survey, the nonresponse distribution is fairly consistent over time. The differences in nonresponse rates between men and women in one survey vary within

narrow limits. In the LFS men always have higher nonresponse rates than women but women have consistently higher rates than men in the PPS.

Fig. 7 Total Nonresponse Among Men and Women in PPS

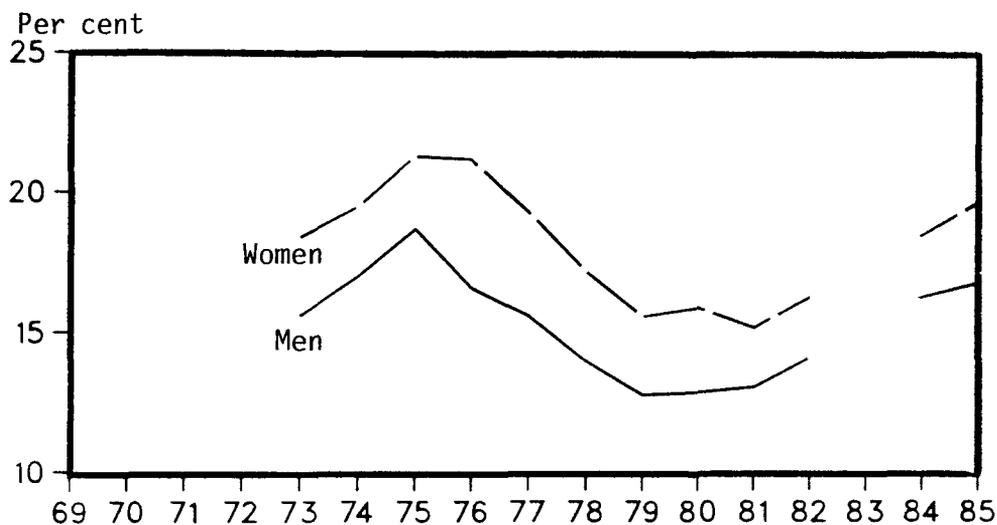
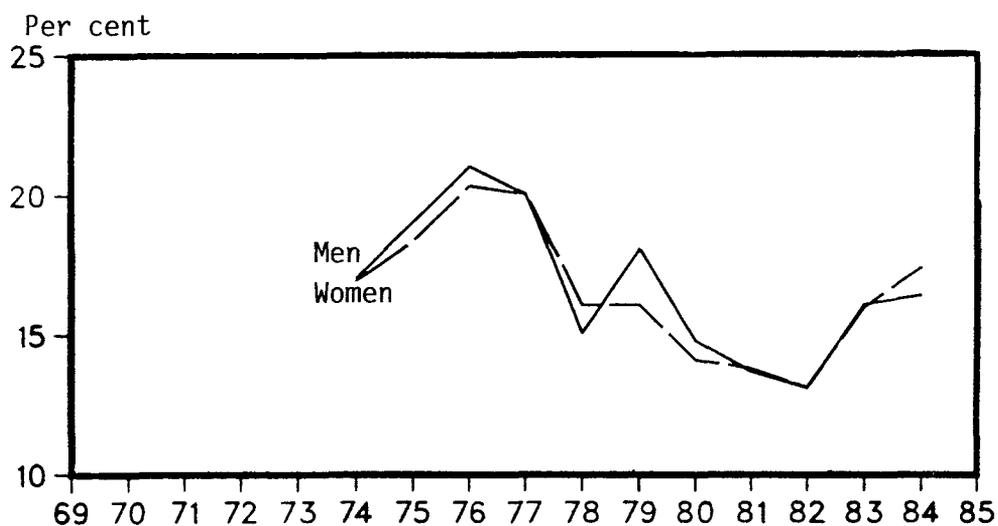


Fig. 8 Total Nonresponse Among Men and Women in SLC.



Third, the same changes that are seen in the time series of the nonresponse for the sample as a whole are mirrored in the nonresponse rates for each sex. Nonresponse rates both increase and decrease simultaneously for men and women.

These three observations will be demonstrated repeatedly in the four domains of study, but will not be emphasized every time they occur. They are also valid when the nonresponse among men and women is further divided into refusals and not-at-homes. The observation that there is no general distribution that holds for all three surveys is strengthened by the time series in figures 9 and 10.

Fig. 9 Refusal and Not-at-Home Rate Among Men and Women in LFS

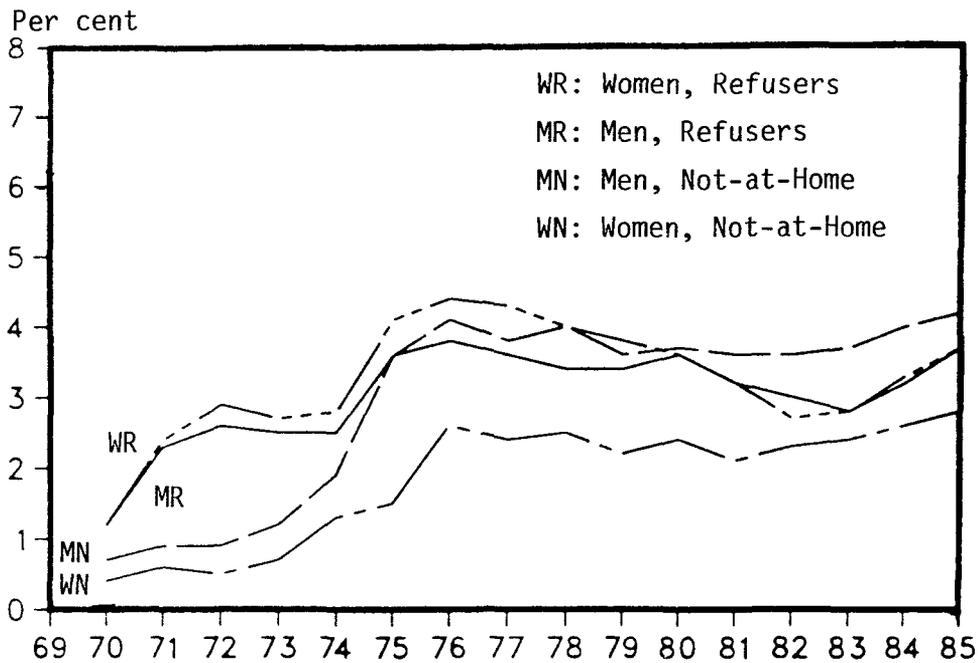
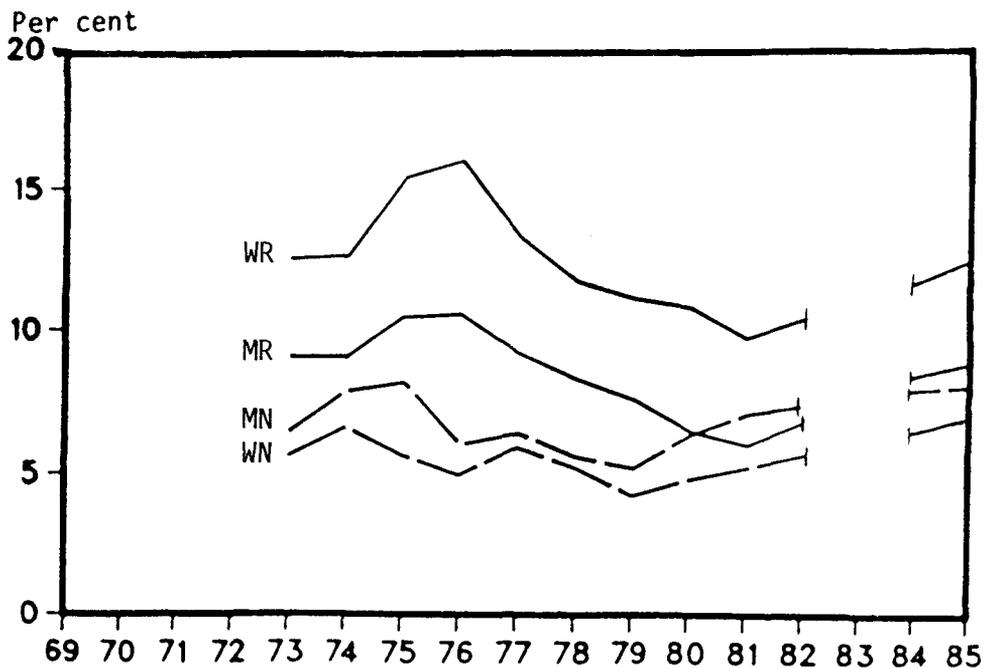


Fig. 10 Refusal and Not-at-Home Rate Among Men and Women in PPS

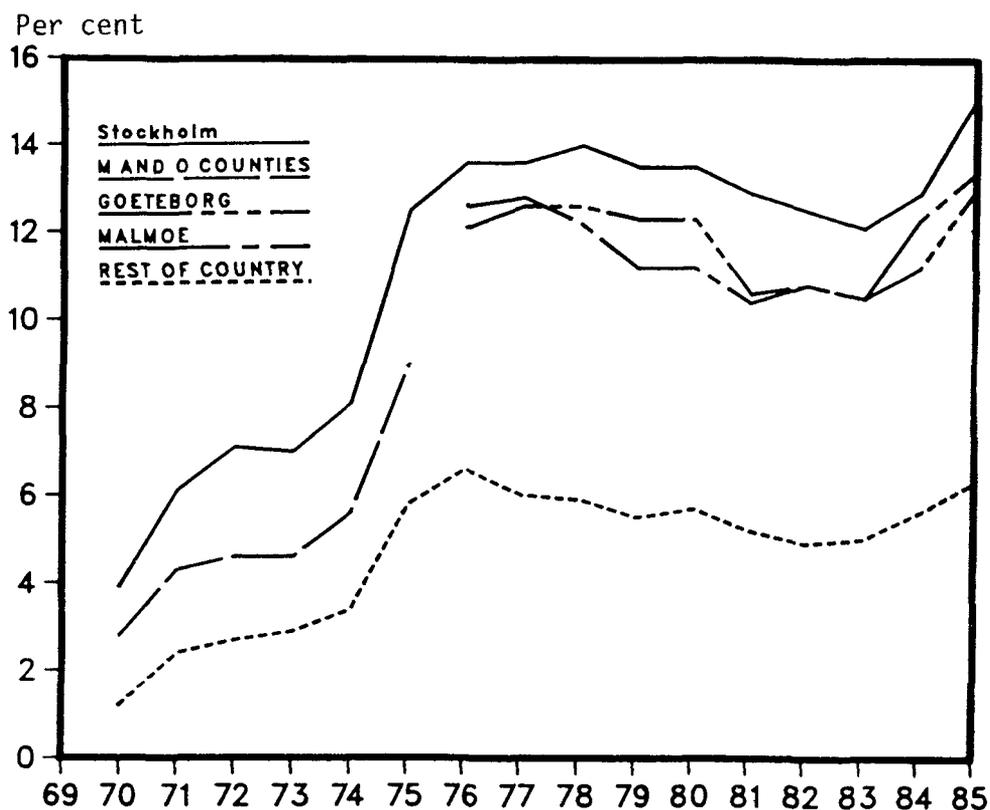


In the LFS, refusal is a slightly more frequent reason for nonresponse among women while not-at-home is the dominant reason for nonresponse among men. In the PPS, whose nonresponse rates are twice those of the LFS's, refusal is a more important source of nonresponse for men than not-at-home. While the not-at-home rates were only slightly lower among women than among men, the refusal rates are three to four percent higher among women in the PPS. Refusal and not-at-home rates are not calculated for any of the domains of study in the SLC.

4.5 Regions and Reason of Nonresponse

Regional domains of study are not similarly classified in the two surveys that calculate regional nonresponse. In the Labour Force Survey, nonresponse rates are calculated for the Stockholm, Göteborg and Malmö regions and the rest of the country (figure 11 12, 13). The rates before 1976 refer to counties. The counties of Göteborg och Malmö were then reported as one unit.

Fig. 11 Total Nonresponse in Regions in LFS



The highly populated urban areas have higher rates of total non-response than other parts of Sweden. Refusal rates decreased from 1976 to 1983 especially in the densely populated areas. In Stockholm the not-at-home rate seems to be constantly on the rise. The not-at-home rates for Malmö and Göteborg seem to vary randomly around a constant level. The same pattern holds for the rest of the country, but on a lower level.

Fig. 12 Refusals in Regions in LFS

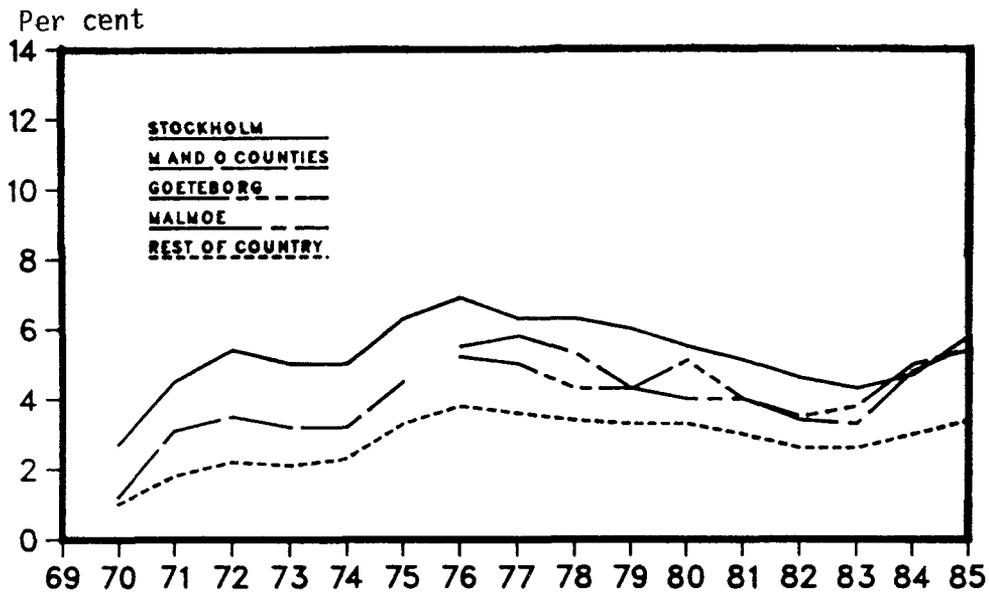
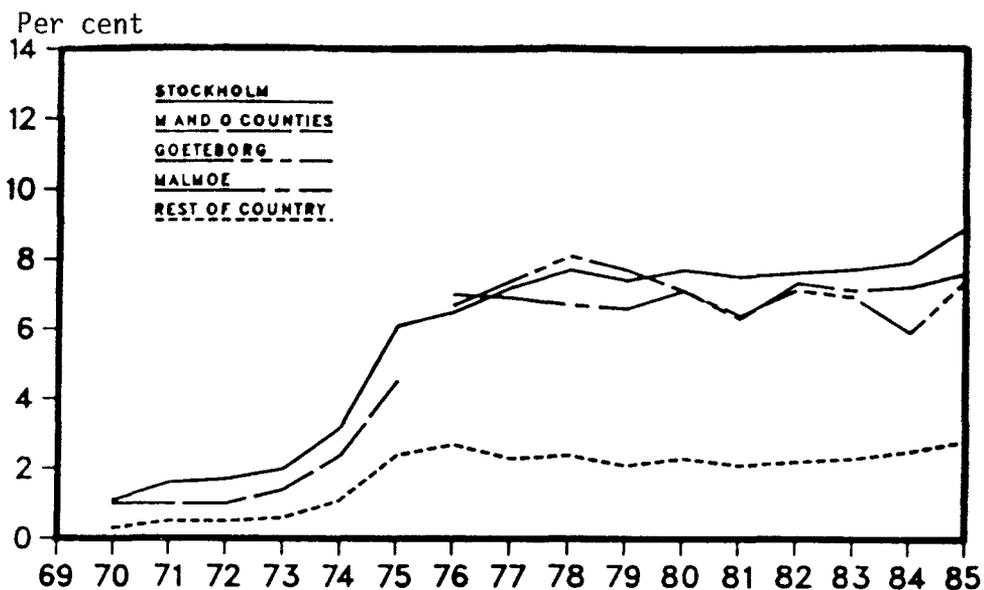


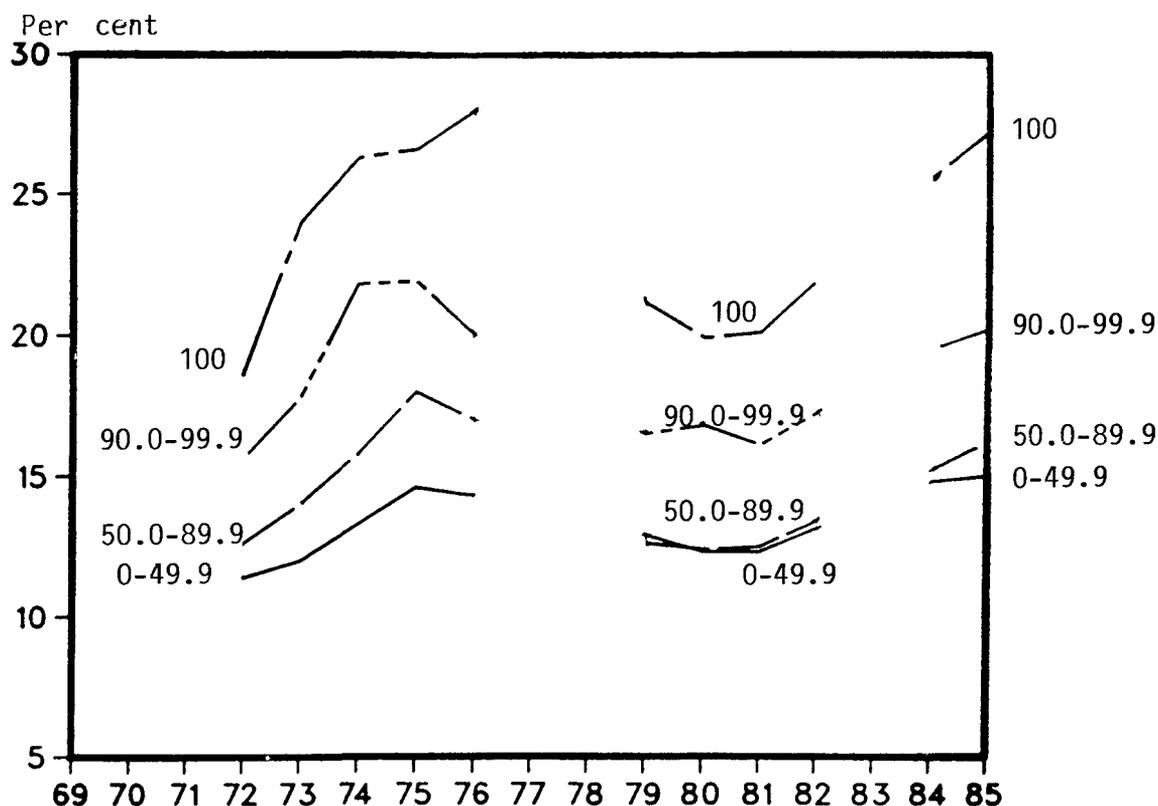
Fig. 13 Not-at-Home Rates in Regions in LFS



The PPS has chosen to present nonresponse rates in population density areas. The population density divisions of figure 14 support the findings based on the regional division. The more densely populated an area is, the higher the nonresponse rates. One lonely exception occurs in 1979 when the population density class 0-49.9 has 0.3 percent higher value than the next density class.

Since 1979 the nonresponse rates of the two lowest density classes have been close to each other. The differences between the two highest classes is substantial and amounts to around five and ten percent respectively.

Fig.14 Nonresponse by Population Density Area in PPS



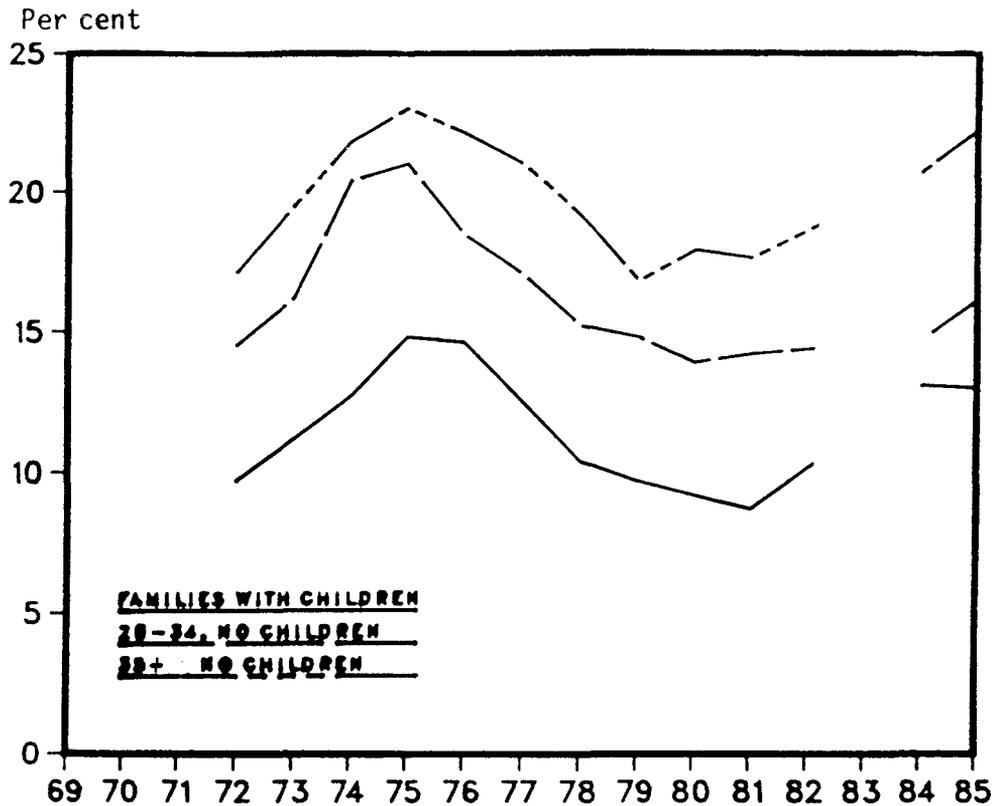
4.5 Family Type and Nonresponse

The variables on which nonresponse is reported are usually chosen to illustrate some specific aspect of the individual. In the PPS there is particularly interesting classification regarding the individuals family type according to the population register.

The classification is:

- * member of a family with children
- * 20-34 years of age and do not belong to a family with children
- * 35+

Fig. 15 Total Nonresponse by Family Type in PPS



The differences in the nonresponse among the family type categories are as great as the differences in nonresponse among the population density regions. The total nonresponse rate among persons belonging to families with children are, as is seen in figure 15, four to five per cent below 20-34 year olds with no children and ten per cent below those 35 and over with no children. Since 1975, the refusal rate - figure 16 - lies between those of the two classes of childless households. The not-at-home rate among families with children as presented in figure 17 is remarkably low - two to three per cent - and only a third or less of that in the other groups.

Fig. 16 Refusals by Family Type in PPS

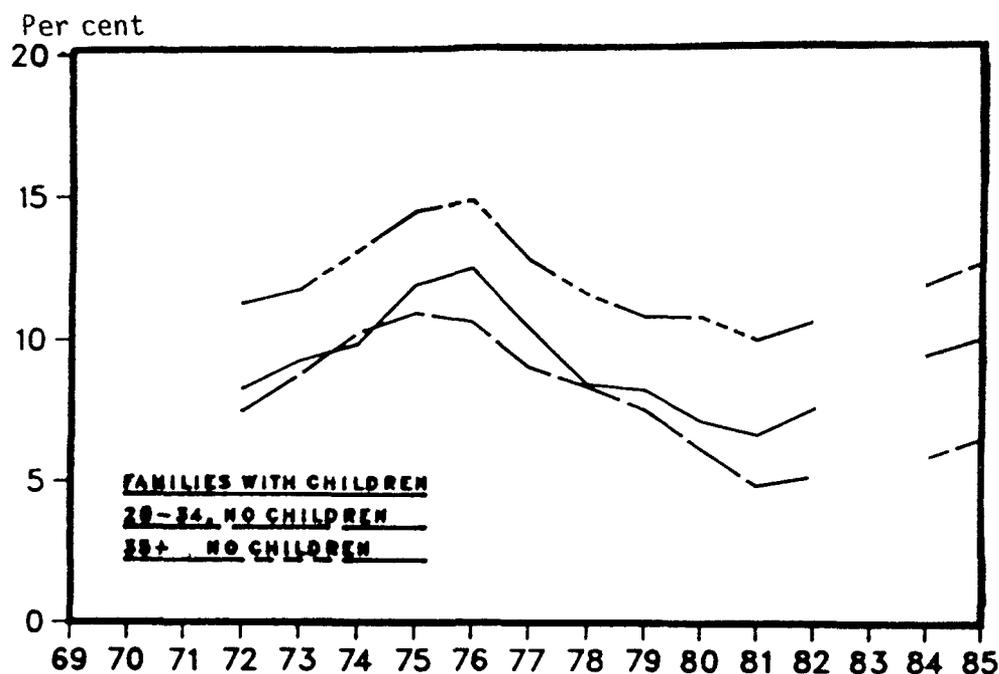
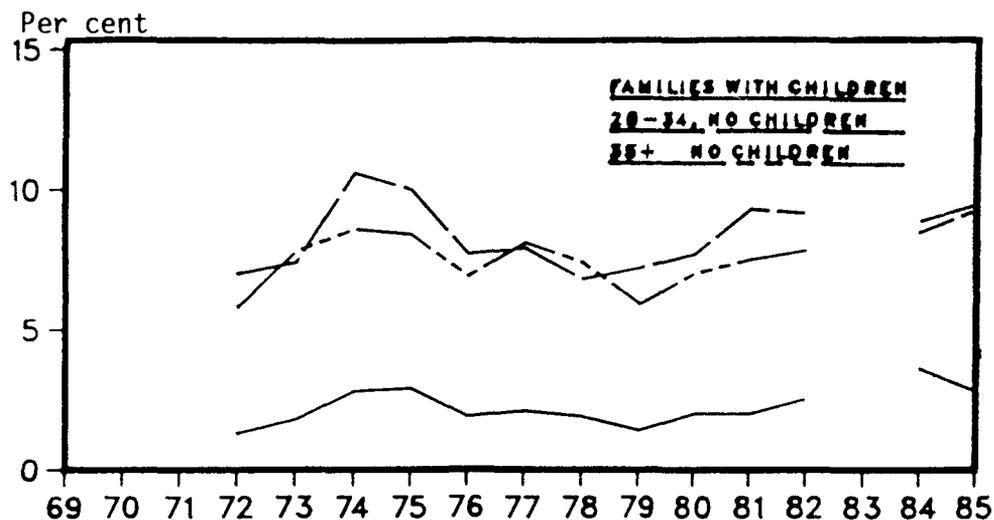


Fig. 17 Not-at-Home Rate by Family Type in PPS



4.6 Summary of the Observations

All of the time series of nonresponse in the interview surveys behaved in much the same way during the 1976 - 1985 period. From a high point in 1975 or 1976, there was a steady decrease and a remarkably low level was reached in 1980-1982. Nonresponse in the LFS changed less compared to the other surveys which is apparent from figure 3. In the last few years, there has been a new increase. However, this increase still has not reached the nonresponse level of the mid-seventies.

Refusals are responsible for a large part of the variation in nonresponse. In 1985 the not-at-home rates for all surveys except the LFS were higher than in 1976. It cannot be excluded that increases in not-at-home rates are explained by changes in design. Other explanations are also feasible but cannot be substantiated.

When the samples are divided according to domains of study, the time series have pretty much the same appearance. There is, however, no pattern common to all the surveys for the distribution of the nonrespondents on domains of study and reasons of nonresponse. Yet, the relationship between the level of the nonresponse rates in domains of study and by reason still hold true during the period, indicating that very persistent factors are at work in each survey.

5 THE GENERAL FACTORS THAT EXPLAIN VARIATIONS OVER TIME

After having taken the technical changes that could have led to changes in nonresponse rates into consideration we still can identify a general pattern in the time series of nonresponse rates in CBE, LFS, PPS and SLC the four interview surveys in the period 1976-85. This pattern is distinct and must be explained by variations either in the survey climate or in the efficiency of the interviewing itself.

Nonresponse, and especially refusals, reached an all time high in 1975-76 after and during an intensive debate led by the mass media on questions of confidentiality and the use of statistics. This debate was unleashed by the concern that 1975 census data could be misused and by other issues on privacy and confidentiality. The debate was partly prompted by the institution of the Data Inspection Board whose task is to supervise and regulate confidentiality aspects of data collection. Obtaining permission from the Data Inspection Board to perform the survey became an integral part of planning a statistical survey.

Surveys now became the topic of public debate and often while still in the planning stage. The unfamiliar situation arose that while trying to conduct interviews, the interviewer found the sampled persons more argumentative and questioning than before. Statistics Sweden's personnel were at that time neither adequately trained nor prepared to argue these issues. The consequence was an increase in nonresponse. Feeling a bit uneasy about the situation, the interviewers might have been less committed to their work than usual. In this way, the combination of the uneasiness and the lack of preparation for the new situation and a new set of attitudes in the sampled population might have contributed to the increase in nonresponse.

The public debate on the production of statistics and confidentiality related issues continued with varying intensity during the entire period that followed, but faded a bit after 1976. The public probably became less uneasy after hearing the same fears voiced repeatedly, without ever having heard of any actual incidences of breach of confidentiality. However, it does not follow that the decline in nonresponse up to 1980-82 was exclusively a consequence of an improvement in the survey climate up to that

period. To an unknown degree, the lower nonresponse rates are the result of increased efforts and efficiency in the data collection and of reduced response burden.

Statistics Sweden considered the increase in nonresponse to be a serious threat to the quality of the surveys results. Each survey formulated and distributed information to the respondents on the measures taken to protect confidentiality. There were also special efforts to educate the interviewers in questions of confidentiality and to train them so that they could better deal with argumentative or reluctant respondents. The agency made a concentrated effort to counteract the increase in nonresponse by appointing a committee with the then serving general director as chairman. This was the UBIS group (the Committee on Questions Concerning Respondents and Nonresponse in the Production of Statistics). The committee was appointed in 1976 and presented its report in 1981. The committee's central task was to improve information in interview surveys, make suggestions for continuing education for interviewers, and to develop measures that would lead to a reduction in the level and effects of nonresponse. Nonresponse fell strikingly and uninterruptedly during the period when UBIS conducted its work and when nonresponse was considered a priority issue.

During the same period, a reduction in the number of interviewers began that is still continuing. This probably led to the remaining interviewers having a more professional attitude towards their work. In 1970-71, the interviewer corps consisted of 533 people, but by 1984-85, the corps had shrunk to 215. At the same time the average number of hours worked increased from 378 to 1107 and Statistics Sweden expanded its training programs for the remaining interviewers. It has also become easier for the central staff corps to monitor the field work. Examples of such monitoring work would be reminding individual interviewers about deadlines and distributing the work more efficiently.

After 1982 there has been a new and continuous upswing in nonresponse rates that is still continuing while this is being written. The discussion surrounding the proposed register-based census (FOBALT) and the 1985 Population and Housing Census (FoB 85) have been heated and are likely to have tainted the survey climate again. However, there are other credible contributions to the rise in nonresponse. Since the UBIS project was successful, nonresponse issues lost their high priority and it is likely that efforts to keep nonresponse under control declined.

Furthermore, an important reason can be the interviewers' increasing work load which was heavier in 1984/85 than in any previous year. With a dwindling number of interviewers each of the remaining interviewers have to cover a larger region than before. During vacations and other absences the local interviewer must be replaced by a colleague who does not know that region equally well and has to spend more time on travelling to the region. As a result there is less time available for each individual or household that should be contacted. Primarily not-at-home but also refusal rates can be negatively affected by this. A continued reduction in the number of interviewers can have serious consequences. For this reason, an analysis of the number of interviewers and the efficiency of the data collection is important.

6 MONITORING THE SURVEY CLIMATE IN THE FUTURE

There is obviously a vested interest in having a numerical indication that shows when it is getting easier or more difficult to conduct interviews in a population. A close watch should be kept on changes in the survey climate in the future.

The use of a single nonresponse indicator, an index or some kind of summary rate, has been suggested. If such an indicator shall be useful, it must include information on several important surveys and in some formal way summarize their nonresponse rates. There is, however, no self-evident method for constructing such an indicator. One must first decide what weights should be given to each survey with regard to its :

- 1) sample size
- 2) respondent burden
- 3) periodicity
- 4) degree of sample rotation
- 5) the type of sampling unit (individual or household)
- 6) the method for data collection (telephone interviewing, face to face interviewing or mail surveys)

Constructing a summary rate would also demand a great deal of research and programming. However, the study of the time series of nonresponse from 1976 to 1985 clearly demonstrates that there is no immediate need for a sophisticated index. Changes in nonresponse rates have been fairly uniform, having increased or decreased simultaneously. Nor is there an obvious answer to the problem of how the index should be interpreted if the surveys show conflicting nonresponse trends. Another problem would be the dominating influence that the LFS would have on the index. The LFS has a sample size as large as all the other surveys put together.

Useful information can be reached in a far simpler way - by plotting time series of the yearly nonresponse rates for each of the regular and important surveys. These surveys would be the LFS, the CBE, the IDS, the SLC and the PPS and they should use common classifications on the register variables sex, age, region and family type. Both the rates of refusals and the rates of not-at-homes should be calculated and presented in time series as well as the over all nonresponse rates, since these have been seen to

have importance both in different surveys and in the different domains of study of the surveys. Efforts should also be made to recalculate nonresponse rates according to the same principles for the last five years.

When a simple system of nonresponse calculation and presentation is adopted, it will not be expensive to maintain. The responsibility for updating the system must be shared by the statistical methods unit (P/STM), the surveys concerned, and the interviewing agency, D/IE.

The detailed description and analysis of the nonresponse rates in domains of study and by reason rates should certainly be the responsibility of each survey.

8 REFERENCES:

Johansson, G. (1980) Svartsmönster i AKU - studie av nyttillkomna personer i arbetskraftsundersökningarnas (AKUs) urval under första kvartalet 1977 (Pattern of Response in LFS - A Study of Newly Sampled Persons in the Labor Force Survey's Sample in the First Quarter of 1977), MIHS No 17.

Lindström, H.L. (1983) Non-response errors in sample surveys. SCB, Urval No 16.

Ribbing, E. (1976) Bortfallets struktur i några SCB-undersökningar. (The Structure of Nonresponse in a Few of Statistics Sweden's Surveys.) SCB, Stockholm.

Waara-Grape, B. (1971) Om bortfallsproblemet i Utredningsinstitutets undersökningar. (On Nonresponse Problems in the Research Institute's Surveys.) SCB, Stockholm..

SCB, (1980) Räkna med bortfall - En handbok om statistiska metoder i samband med bortfall. (Count on Nonresponse - A Handbook about Statistical Methods to Deal with Nonresponse.) Stockholm.

SCB (1981) Rapport från UBIS-projektet. - Utredningen rörande uppgiftslämnarfrågor och bortfallsproblem i statistikproduktionen. (Report from the UBIS Project - An Examination of Response Burden and Nonresponse Problems in the Production of Statistics.) SCB, Stockholm. A summary of this report also exists.

Tidigare nummer av Promemorior från P/STM:

NR

- 1 Bayesianska idéer vid planeringen av sample surveys. Lars Lyberg (1978-11-01)
- 2 Litteraturförteckning över artiklar om kontingenstabeller. Anders Andersson (1978-11-07)
- 3 En presentation av Box-Jenkins metod för analys och prognos av tidsserier. Åke Holmén (1979-12-20)
- 4 Handledning i AID-analys. Anders Norberg (1980-10-22)
- 5 Utredning angående statistisk analysverksamhet vid SCB: Slutrapport. P/STM, Analysprojektet (1980-10-31)
- 6 Metoder för evalvering av noggrannheten i SCBs statistik. En översikt. Jörgen Dalén (1981-03-02)
- 7 Effektiva strategier för estimation av förändringar och nivåer vid föränderlig population. Gösta Forsman och Tomas Garås (1982-11-01)
- 8 How large must the sample size be? Nominal confidence levels versus actual coverage probabilities in simple random sampling. Jörgen Dalén (1983-02-14)
- 9 Regression analysis and ratio analysis for domains. A randomization theory approach. Eva Elvers, Carl Erik Särndal, Jan Wretman och Göran Örnberg (1983-06-20)
- 10 Current survey research at Statistics Sweden. Lars Lyberg, Bengt Swensson och Jan Håkan Wretman (1983-09-01)
- 11 Utjämningsmetoder vid nivåkorrigering av tidsserier med tillämpning på nationalräkenskapsdata. Lars-Otto Sjöberg (1984-01-11)
- 12 Regressionsanalys för f d statistikstuderande. Harry Lütjohann (1984-02-01)
- 13 Estimating Gini and Entropy inequality parameters. Fredrik Nygård och Arne Sandström (1985-01-09)
- 14 Income inequality measures based on sample surveys. Fredrik Nygård och Arne Sandström (1985-05-20)
- 15 Granskning och evalvering av surveymodeller, tiden före 1960. Gösta Forsman (1985-05-30)
- 16 Variance estimators of the Gini coefficient - simple random sampling. Arne Sandström, Jan Wretman och Bertil Waldén (Memo, Februari 1985)
- 17 Variance estimators of the Gini coefficient - probability sampling. Arne Sandström, Jan Wretman och Bertil Waldén (1985-07-05)
- 18 Reconciling tables and margins using least-squares. Harry Lütjohann (1985-08-01)

- 19 Ersättnings och uppgiftslämnarbördans betydelse för kvaliteten i undersökningarna om hushållens utgifter. Håkan L. Lindström (1985-11-29)
- 20 A general view of estimation for two phases of selection. Carl-Erik Särndal och Bengt Swensson (1985-12-05)
- 21 On the use of automated coding at Statistics Sweden. Lars Lyberg (1986-01-16)
- 22 Quality Control of Coding Operations at Statistics Sweden. Lars Lyberg (1986-03-20)
- 23 A General View of Nonresponse Bias in Some Sample Surveys of the Swedish Population. Håkan L Lindström (1986-05-16)

Kvarvarande exemplar av ovanstående promemorior kan rekvireras från
Elseliv Lindfors, P/STM, SCB, 115 81 Stockholm, eller per telefon
08 7834178