Ib Thomsen, Øyvin Kleven, Jan Henrik Wang and Li-Chun Zhang

Coping with decreasing response rates in Statistics Norway

Recommended practice for reducing the effect of nonresponse

Statistisk sentralbyrå • Statistics Norway
Oslo–Kongsvinger

Rapporter

I denne serien publiseres statistiske analyser, metode- og modellbeskrivelser fra de enkelte forsknings- og statistikkområder. Også resultater av ulike enkeltundersøkelser publiseres her, oftest med utfyllende kommentarer og analyser.

Reports

This series contains statistical analyses and method and model descriptions from the various research and statistics areas. Results of various single surveys are also published here, usually with supplementary comments and analyses.

ISBN 82-537-7054-5 Trykt versjon
ISBN 82-537-7055-3 Elektronisk versjon
ISSN 0806-2056

Emnegruppe
00.90

Design: Enzo Finger Design
Trykk: Statistisk sentralbyrå/39

Standardtegn i tabeller	Symbols in tables	Symbol
Tall kan ikke forekomme	Category not applicable	.
Oppgave mangler	Data not available	..
Oppgave mangler foreløpig	Data not yet available	...
Tall kan ikke offentliggjøres	Not for publication	:
Null	Nil	-
Mindre enn 0,5 av den brukte enheten	Less than 0.5 of unit employed	0
Mindre enn 0,05 av den brukte enheten	Less than 0.05 of unit employed	0,0
Foreløpig tall	Provisional or preliminary figure	*
Brudd i den loddrette serien	Break in the homogeneity of a vertical series	—
Brudd i den vannrette serien	Break in the homogeneity of a horizontal series	I
Desimalskilletegn	Decimal punctuation mark	,(.)

Abstract

Ib Thomsen, Øyvin Kleven, Jan Henrik Wang and Li-Chun Zhang

Coping with decreasing response rates in Statistics Norway
Recommended practice for reducing the effect of nonresponse

Reports 2006/29 • Statistics Norway 2006

Over the last two decades, cooperation rates in most sample surveys appear to have declined in many countries, including Norway. As a consequence, the cost of conducting surveys has increased because repeated attempts must be made in order to seek information from reluctant sample members.
In this report, we provide a review of experiences in Statistics Norway in the field of survey nonresponse. The report presents a recommended practice manual, RPM, based on these experiences. As nonresponse is part of quality, we treat the response process as part of the production process in a systematic quality approach, and recommend methods and techniques applied to reduce unit nonresponse and its effects on the final official statistics.
The report is divided into five chapters:

Chapter 1 provides an introduction of some basic concepts within the systematic quality approach and some recommendations.

Chapter 2 presents some response rates and nonresponse biases and their development over time and some international comparisons.

Chapter 3 is concerned with household surveys. Key factors that have important impact on nonresponse are identified. For many of these factors, key process variables are measured and applied to find a good balance between cost and quality.

Chapter 4 gives a similar presentation concerning business surveys.

Chapter 5 presents techniques used to investigate the nature of the effects on nonresponse on the population estimates together with the most commonly used weighting methods.

Acknowledgement: Jan Bjørnstad, Director of Research of the Division for Statistical Methods and Standards, has given thorough and comprehensive comments to previous drafts, which has improved the readability of the report substantially.

Contents

1. Introduction 8
1.1. Introduction 8
1.2. Nonresponse as a process 9
1.3 Conclusions and recommendations 9

2. Development of some key process variables 11
2.1. Introduction 11
2.2. Some key concepts 11
2.3. Trends in nonresponse rates 11
2.4. Trends in field efforts to increase response rates 12
2.5. Relationship between nonresponse rate and nonresponse bias 12
2.6. Some international comparisons 13

3. Interviewer-administered household surveys 14
3.1. Introduction 14
3.2. Infrastructure of household interviewing 14
3.3. Processes of interviewer surveys 17

4. Business surveys 27
4.1. Introduction 27
4.2. Organisation of business surveys 27
4.3. Data collection process 29
4.4. Identifying and measuring process variables related to nonresponse 32
4.5. The use of key process variables to evaluate product quality 37

5. Weighting for nonresponse bias 40
5.1. Introduction 40
5.2. Use of administrative data to study the effects of nonresponse 40
5.3. Use of information from the interviewers 42
5.4. Weighting methods 42
5.5. Estimation of variances in the presence of nonresponse 46

Appendix
Interviewer-administered surveys in Statistics Norvay 47

References 48

Previously issued on the subject 51

Recent publications in the series Reports 52

List of Figures

1. Introduction
1.1. Cause-effect diagram 1. Reduce nonresponse errors 9
1.2. Process efficiency graph. Response rate and nonresponse bias of the variable "mean number of live births per woman", by number of call-backs 10

2. Development of some key process variables
2.1. Response rates in some Norwegian household surveys 1967-2004 11
2.2. Response rates in some Norwegian business surveys 2000-2004 11
2.3. Nonresponse rate and days of fieldwork. Norwegian Survey of Level of Living, 1995-2004 12
2.4. Nonresponse rate and bias on voter turnout. Norwegian general election survey 1989 - 2005 13
2.5. Response rates. The European Social Survey 2002/2003 and 2004 13

3. Interviewer-administered household surveys
3.1. Organisation of sample surveys in Statistics Norway 15
3.2. Basic training of interviewers 15
3.3. Response rate among interviewers. Survey on Level of living 2005. 136 interviewers 8500 cases 16
3.4. A macro flow chart of the interviewer process in Statistics Norway 17
3.5. Response rate in different surveys 2004, in percentages 18
3.6. A typical response rate development. ESS 2004. 18
3.7. Voter turnout among respondents and types of nonrespondents. Election Survey 1997-2005 21
3.8. Non-contact rate in different surveys 2004 21
3.9. A conceptual model for contacting the interviewer object 22
3.10 A conceptual framework for survey cooperation 24
3.11. A typical relationship between response rate and costs. ESS 2004 25
3.12. Cumulative electoral turnout in the net sample by days of fieldwork. General election survey 1997 26

4. Business surveys
4.1. Organisation of the data collection process in business surveys in Statistics Norway 28
4.2. Data collection process in business surveys 31
4.3. Weighted and unweighted nonresponse in the Business Tendency Survey. 4th quarter 2004 33
4.4. Development in nonresponse at critical dates in the data collection process for the QIS 34
4.5. Control chart for nonresponse in QIS for 1st reminder 35
4.6. Response profile for the QIS 4th quarter 2004 for web and paper questionnaire 36
4.7. Process efficiency graph: nonresponse and product quality in the BTS 4th quarter 2004 38

5. Weighting for nonresponse bias
5.1. Employed persons estimated in the Labour force survey (2001-2004) using three different weighting methods. Percent 44

List of Tables

2. Development of some key process variables
2.1. Response rate in two surveys from the Nordic countries 13

3. Interviewer-administered household surveys
3.1. A short history of the development of interviewer- assisted interviewing in Statistics Norway 14
3.2. Local interviewers, by gender, age and years in the job, in percentages 15
3.3. Response rate among local interviewers, by gender, age and years in the job, in percentages. Ordinary assigned cases. ESS 2004 N=2750 16
3.4. Types of incentives used in Statistics Norway 19
3.5. Experiments with prepaid incentives in Statistics Norway 20
3.6. Types of tailoring incentives in Statistics Norway 20
3.7. Telephone number present in the sample of ESS 2 004, in percentages 20
3.8. Different modes of contact in the first interaction attempt in ESS 2004, in percentages 23
3.9. Result of the first interaction in ESS 2004, in percentages 23
3.10. "Refusers" in the first round of Norwegian election survey 1997-2001 24

4. Business surveys
4.1. Factors used in weighting from stratum to industry level in the BTS 2004 for NACE 22.1 39

5. Weighting for nonresponse bias
5.1. Effects of nonresponse in the Norwegian Election Survey 2001 40
5.2. Response rate among voters and non voters 41
5.3. P-values for the variables 42
5.4. Turnout in last elections among respondents, refusals and not at homes/prevented 42
5.5. Turnout and ease of response in Election surveys 1997- 2003 42
5.6. Effect of weighting for 5 variables 43
5.7. Age distribution estimated by using the reciprocal of the inclusion probability as weight and calibrated weights 45
5.8. Income distributions estimated by using the reciprocal of the inclusion probability as weight and calibrated weights 45

1. Introduction

1.1. Introduction

In this document, we present a recommended practice manual, RPM, for including the response process as a part of the production process in the systematic quality approach and a recommended practice for how to reduce the rate of unit nonresponse and its effect on the final official statistics. Item nonresponse is not considered in this manual. Using experiences from surveys in Statistics Norway, SN, we divide the response process into three sub processes:

- Identification of influential factors and key process variable, which affect unit response rate.
- Give examples of how to obtain information about nonrespondents in order to study the nature of nonresponse bias.
- Present some methods used in order to reduce the nonresponse bias and recommend some "best methods" depending on the type of survey and auxiliary information available.

In the late 1960s, a division responsible for household surveys was set up at SN. At that time, the theory and practice for sampling errors was well established and a number of high quality textbooks were available, e.g., Cochran (1977), Kish (1965). Theory and practice concerning nonresponse was, however, more scarce and the need to cope with these problems was urgent. In 1980, the U.S. National Research Council organized a meeting on nonresponse. Researchers from several countries and institutions were invited to present and discuss their experiences. The results of these activities are published in Madow et al (1983). Statistics Norway by Thomsen and Siring (1983), contributed to this meeting. For SN and probably for many other statistical institutions, this meeting represented a turning point in the development of efficient methods to cope with nonresponse and its effects on quality. Since then, there has been a remarkable development in techniques and methods, which can be used to handle nonresponse. An excellent presentation of the development can be found in Groves et al (2002), which was stimulated by the International Conference in Survey Nonresponse, held in Portland, Oregon, USA, October 1999.).

In this RPM, we shall present some of the measures taken by SN in order to meet the challenges created by increasing nonresponse rates in many surveys. These measures include improving collection methods, weighting procedures and application of the Statistical Act. The paper is divided into 5 chapters. In chapter 2 we present some response rates and nonresponse biases, their development over time and some international comparisons. The chapter is meant as a background for the following chapters, and no attempts are made to explain the observed development or the variation between countries. Chapter 3 is concerned with household surveys. We regard the control of response rates as a process, which means that we identify key factors, which have important impact on response rates. For a majority of these factors, we define some key process variables and apply the principles of continuous quality improvement to find a good balance between response rate/bias and costs. In chapter 4 a similar presentation is given concerning business surveys. In chapter 5, attention is given to methods used to investigate the nature of the effects of nonresponse on the population estimates, and we present the most common weighting methods used to reduce these effects. In the literature one often finds the expression reweighting when special weighing is made to reduce the effects of nonreponse. This is done in order to distinguish it from weighting due to the sample design. As no sample design issues are central in this paper, we shall use both weighting and reweighting to mean weighting for nonresponse, unless otherwise explained.

As the causes and effects of nonresponse vary from one survey to another, and because the Norwegian statistical system covers a large number of surveys, it is not possible to cover all aspects of all surveys in detail. We present and discuss basic principles and methods, which can be applied to all surveys. By using examples from a large number of different surveys, we also demonstrate how these methods are used in connection with specific surveys.

1.2. Nonresponse as a process

Nonresponse is part of quality. In 2002, the Leadership Expert Group (LEG) on quality presented a report, Eurostat (2002), recommending to consider product quality as a result of a process. Within this framework, the product quality is generated by an underlying process or sequence of processes. To follow up this work, Eurostat recently issued a handbook, Eurostat (2005), on improving quality by analysis of process variables.. In the present paper we shall use recommendations from this handbook when we discuss measures to cope with nonresponse. The reason for choosing this approach is that we believe that it will become of vital importance and serve as an important tool in future efforts to further improve our ability to cope with nonresponse issues.

To control this process, it is necessary to identify the underlying, influential factors and some measurable key process variables for each factor. Any improvements of the process is based on measuring these key variables, adjusting the process based on these measurements, and finally study the effects on nonresponse rates and nonresponse bias. Two relevant references are Deming (1991) and Sæbø, Byfuglien and Johannessen (2003).

The comprehensive Cause-effect diagram given below is presented in Eurostat (2005). It lists factors, which influence nonresponse, and is divided into four processes. For many of the factors listed in the diagram, it is very difficult to identify and measure possible key process variables. In the present report, attention is therefore limited to some important key variables, quality results variables and the relation between them.

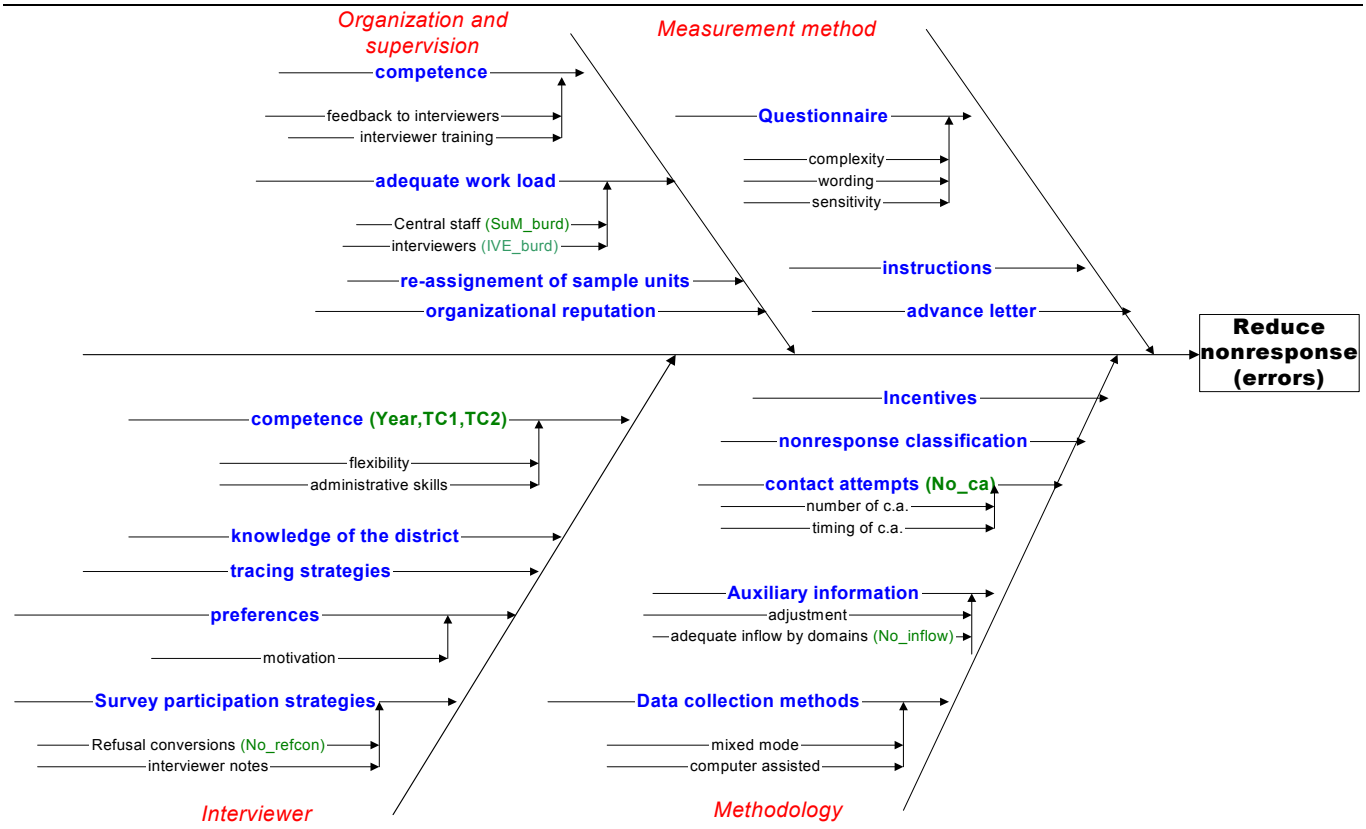
1.3 Conclusions and recommendations

Efforts to control nonresponse and its effects on quality should follow the principles and methods for improving quality by analysis of process variables. This means to identify key process variables and their effects on nonresponse. Following this principle, the conclusions and recommendations will be presented separately for some key variables:

Response rates

Statistics Norway has managed to maintain high response rates in business surveys by making it mandatory to respond and by applying user-friendly collection modes. Concerning most household surveys, response rates have been decreasing or constant in spite of increased efforts to maintain a high response. One important exception is the Labour Force Surveys for which response rates are relatively high. The reason is that it is mandatory to respond and proxy interviews are used. see Solheim, Håland and Lagerstrøm (2001).

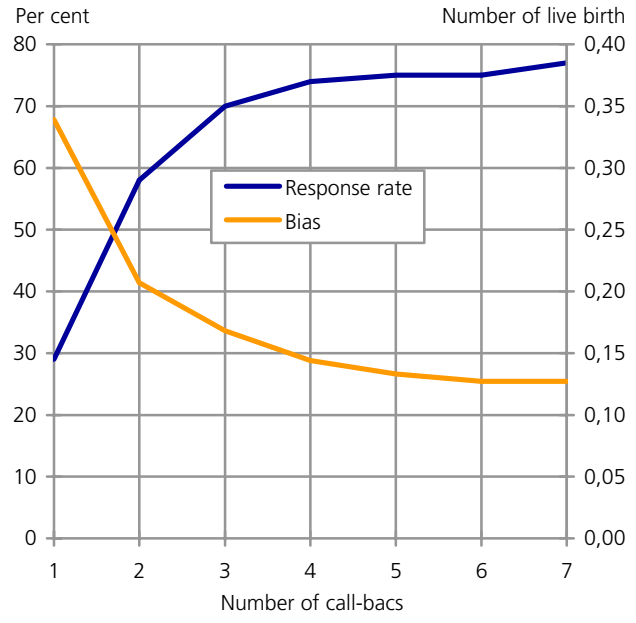
Figure 1.1. Cause-effect diagram 1. Reduce nonresponse errors



Nonresponse bias.
For most survey organizations, a high response rate is in itself considered part of good quality. It is, however, clear that the real quality concern is nonresponse bias. The relationship between the two concepts is very complex. In a recent published article, Groves (2005), the nonresponse rates and nonresponse bias is presented for 301 variables from 30 different methodological studies. One of the conclusions is that nonresponse rate alone is not a good predictor of the magnitude of the bias. In SN, very little work has been done in this area. As the availability of data from administrative registers offers many opportunities to study the relationship between nonresponse rate and its induced bias, studies should be undertaken to throw further light on this question.

Process efficiency graph
As mentioned above, there has been a substantial increase in time and resources spent in order to maintain high response rates. It is therefore important to monitor whether this increase has had any impact on the quality of the results. When analyzing this complicated relationship, we have found it useful to use graphs showing a product quality variable as a function of the time (resources) used to get a higher response rate. We have denoted this graph, The Process Efficiency Graph (PEG). This graph can take several forms, as an example, we shall use data from the Norwegian Fertility Study (FS). Other examples are given in Chapters 3 and 4.

Figure 1.2. Process efficiency graph. Response rate and nonresponse bias of the variable "mean number of live births per woman", by number of call-backs



The sample in the FS consists of 5047 women in the age group 18-44 years. The graph shows how the nonresponse rate decreases, and more importantly, how the nonresponse bias for the variable "Mean Number of Live Births" is reduced as a result of call-backs. We shall not discuss the final choice of the number of call backs here, but point out the usefulness of the PEG when making this choice.

Weighting for nonresponse bias
In many cases, the nonresponse bias on a number of register variables is routinely calculated, and used as a process variable in the decision concerning further weighting. When weighting for nonresponse, one usually starts with poststratification along some important auxiliary variables from administrative registers. If the number of auxiliary variables is high, calibration is usually used after a careful evaluation of various weighting techniques. In this manner, the marginals of some important auxiliary variables are kept constant between various surveys and equal to the marginals in the register. In addition, weighting often reduces sampling variance because high quality, administrative registers are used to find the weights. Weighting is therefore often chosen, even when the effects on the nonresponse bias are negligible. In cases where the weighted population estimates differ from the unweighted ones it is important to justify the choice of estimate, which essentially means to choose a nonresponse model. During this sometimes complex endeavour, it is useful to use a sensitivity analysis presented in chapter 5.

Harmonization of weighting methods
Up until now, weighting methods have varied from one household survey to another. The main reason is that information concerning the distribution of households by size and type has varied over time. Once the new household register, based on administrative files is ready, it is time to harmonize weighting methods between various household surveys, using this register as a basic calibration/poststratifying variable. During this work, increased attention should be given to the consequences of weighting for the variance. See paragraph 5.5.

2. Development of some key process variables

2.1. Introduction

There has been a substantial technological and organizational development in collection methodology during the last two decades. In chapter 3 and 4, a comprehensive and up-to-date presentation is given of various processes involved in dealing with nonresponse. In this chapter, we shall present some graphs showing how some key variables developed in various important surveys during the last 40 years, and make some comparisons with other countries. Finally, we will have a short look at the relationship between response rates and response bias.

2.2. Some key concepts

For many years, there has been a need to standardize the definitions of key concepts in connection with nonresponse. In order to meet these needs, The American Association for Public Opinion Research has published a comprehensive report covering most surveys of persons and households. For a comprehensive discussion of the development of nonresponse standards, see Smith (2002). As far as we know, a similar detailed report concerning business surveys does not exist. There is a need for such a report. Some definitions are suggested in chapter 4.

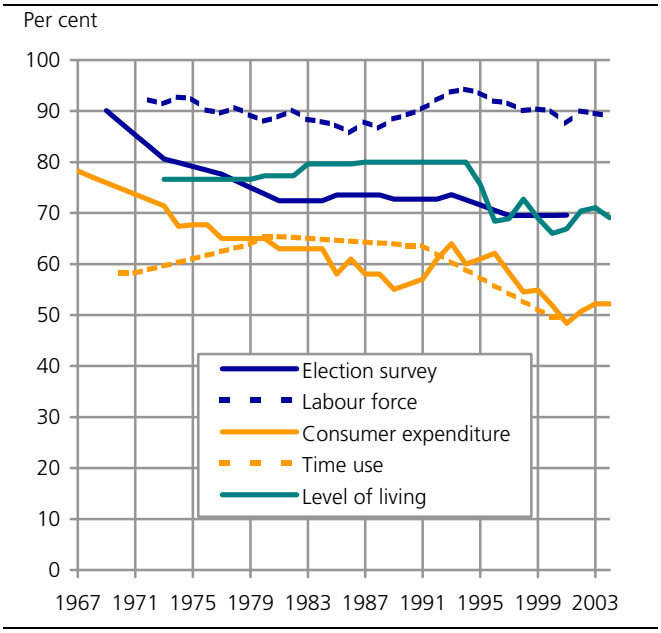
In this "historical" overview, only very few process variables are available, and we shall concentrate on the following four variables:

- **Response rate**: The number of completed interviews with reporting units divided by the number of eligible units in the sample.
- **Refusal rates**: The number of units for which the respondent refuses to be interviewed divided by the number of eligible units in the sample.
- **Non-contact rates**: The number of units for which no respondent was reached divided by the number of eligible units in the sample.
- **Number of days used in the field for data collection:** This key variable is used as an indicator of resources spent to increase response rates. Other indicators are used in chapter 3 and 4.

2.3. Trends in nonresponse rates

Nonresponse has occurred in household as well as business surveys as long as such surveys have been undertaken. As a background to the present analysis, we shall present some graphs describing the development of response rates in some important surveys from Statistics Norway. A detailed analysis of these graphs is not possible, because very few process variables are available for many of the surveys. However, based on visual inspection of the graphs and some general insight, we have drawn some conclusions.

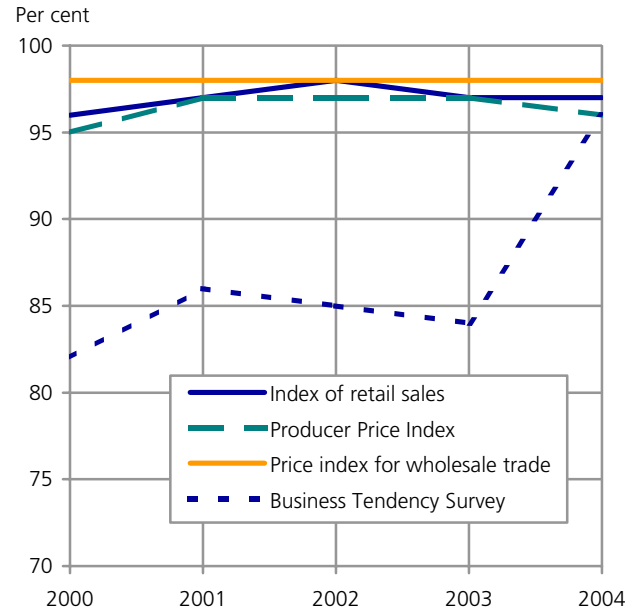
Figure 2.1. Response rates in some Norwegian household surveys 1967-2004



In Figure 2.1 is seen that there is a clear negative trend in the development of response rates for most household surveys. This trend has been observed in a number of countries. It is, however, important to notice that it has been possible to maintain high response rates in the Labour Force Surveys. Surveys like Expenditure Surveys and Time Use Surveys have

the lowest response rates throughout the period. This clearly illustrates that response burden seriously affects the nonresponse rate.

Figure 2.2. Response rates in some Norwegian business surveys 2000-2004

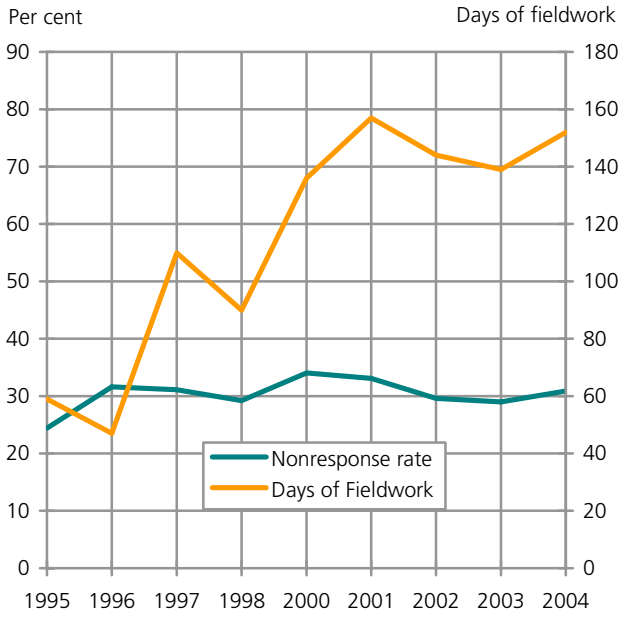


Concerning business surveys, Figure 2.2 shows that business surveys have in recent years had very high response rates. A more comprehensive explanation is given in chapter 4, but it is believed that making the surveys mandatory combined with user-friendly collection modes, have contributed to this positive development.

2.4. Trends in field efforts to increase response rates

The findings shown in figure 2.3 are typical for most household surveys. During the period, there has been a substantial development in collection methodology. Unfortunately, there are few clear indications of the effects on response rates. It is generally believed that without this development in collection techniques, the response rates would have been smaller during the later years. Some of these questions are further studied in the following two chapters. In figure 2.3 we have introduced the key process variable "Number of days in the field" as an indicator of the amount of resources allocation to follow-up efforts to reduce nonresponse.

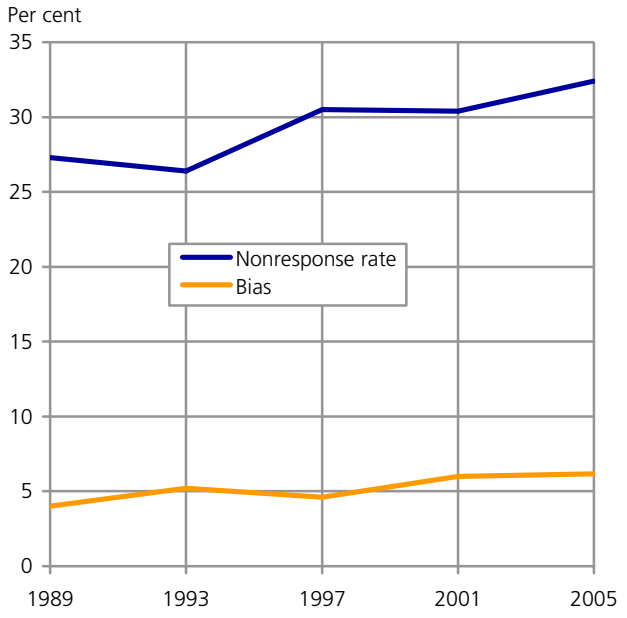
Figure 2.3. Nonresponse rate and days of fieldwork. Norwegian Survey of Level of Living, 1995-2004



2.5. Relationship between nonresponse rate and nonresponse bias

Attention is often concentrated on response rates, which is usually simple to measure. It is, however clear that it is the introduced bias, which raises quality concern. The relationship between nonresponse rate and nonresponse bias is complicated and often impossible to determine, simply because the bias is usually very difficult to estimate for most variables in a survey. A large number of variables, which are highly correlated with the target variables of a survey, are available from administrative registers. For such variables, it is possible explore the relationship between nonresponse bias and rate. Figure 2.4 shows the development of response rates and nonresponse bias of electoral turnout in the General Election Surveys 1989-2005. Similar graphs can be made for a number of other variables.

Figure 2.4. Nonresponse rate and bias on voter turnout. Norwegian general election survey 1989 - 2005



2.6. Some international comparisons

In recent years there has been a growing interest for making studies of international response trends for household surveys. In the early 1990's, The International Workshop on Household Survey Nonresponse initiated the establishment of a database with nonresponse data from 16 countries over a number of years, see De Heer (1999) and de Leeuw and de Heer (2002). One of the conclusions from an analysis of these data is:

"There is ample empirical evidence that response rates are declining internationally. Nonresponse is indeed an increasing problem in the developed world."

To study the variation between the countries, the nonresponse rate was split into two components, the noncontact rate and the refusal rate. It was found that noncontact rate was associated with average household size and certain aspects of the design of the fieldwork, like strictness of the supervision of the interviewers. Refusal rate was associated with economic indicators of the country, supporting a hypothesis that lower cooperation with government surveys is associated with better economic conditions (see Harris-Kojetin and Tucker ,1999). Norwegian data are not included in the database. In table 2.1 response rates for two important household surveys are presented together with response rates from other Nordic countries.

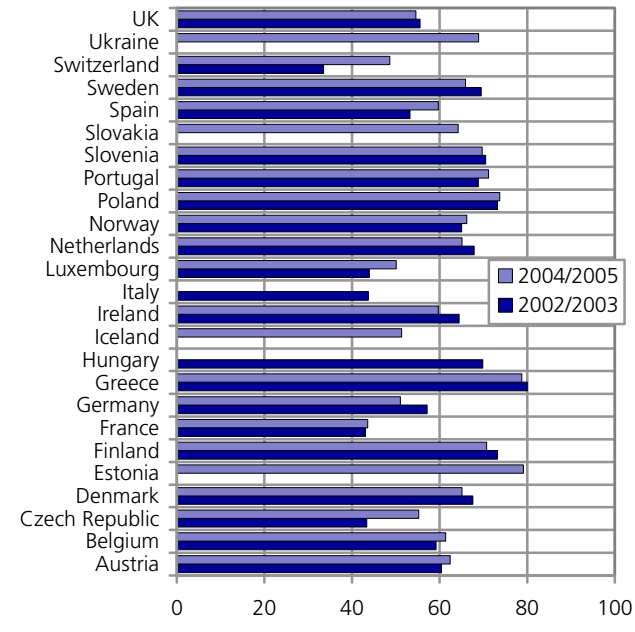
Table 2.1. Response rate in two surveys from the Nordic countries

Country	Denmark	Finland	Sweden	Norway
Labour force survey				
1992	82	92	86	92
1993	82	93	88	94
1994	76	92	87	94
1995	74	93	87	94
1996	75	92	87	92
1997	-	91	86	92
Expenditure survey				
1994	67,4	62,9		60
1995	71,3	67	64	61
1996	67,8	65,4	54	62
1997	65,1	-		58

2.6.1. The European Social Survey

The European Social Survey offers a special opportunity to compare nonresponse information between countries. The aim of the survey is to collect and analyse data on values, attitudes and beliefs among Europeans. Two rounds of the survey have been undertaken, in 2002/2003 and 2004 in about 20 countries. Several factors, which create differences in response rates, were controlled for by using standardized data collection procedures and registration of process data from the field (see Stoop , 2005). In figure 2.5 the response rates in the participating countries are shown. Despite all efforts to standardize the procedures, the response rates vary substantially between the countries in both rounds. In Koch and Stoop (2005) refusal rates are correlated with economic indicators, confirming the relationship found in Harris-Kojetin and Tucker (1999). It is also found that statistical agencies all perform better than when commercial organisations are responsible for the data collection.

Figure 2.5. Response rates. The European Social Survey 2002/2003 and 2004



Source: www.ess.nsd.uib.no

3. Interviewer-administered household surveys

3.1. Introduction

Statistics Norway was established in 1876, and was one of the pioneers in survey sampling. Leslie Kish wrote, *"The 1895 paper of A.N. Kiair can well serve for an official birth date for survey sampling"* (Kish 1995:813). The institution was not equally pioneering with regard to interviewer-administered surveys. While a Norwegian section of the Gallup organisation was established in the late 1940s, it was not until 1966 that a regular division and permanent interviewer corps was established in Statistics Norway. Collection methodology has seen a substantial development, especially since the electronic era started. In the beginning of the 1990s, laptops and computer-assisted personal interviewing were introduced. Since 1995, the field interviewers have carried out both face-to-face interviews and telephone interviews from their home. Only ad-hoc surveys are carried out using pen and paper. Computer Assisted Interviewing provides rich data on the interviewing process, including the development of response rate, and the opportunity to analyse process data and data from the survey immediately after the survey has started.

Table 3.1. A short history of the development of interviewer-assisted interviewing in Statistics Norway

1876	Statistics Norway is established.
1891	Director General Kiær carries out one of the first sample surveys in the world.
1897	Kiær publishes: Den repræsentative undersøgelsesmethode
1958	Survey of Consumer Expenditure carried out. Gallup Norway does the fieldwork.
1966	Division for sample surveys established. 100 interviewers
1971	Labour force survey starts. Number of interviewers increased to 300
1992	Computer-Assisted Telephone Interviewing starts. Blaise is introduced.
1995	Computer-Assisted Interviewing only. Interviewers regrouped and decreased to 150.
2000	Windows platform with a fully integrated computer exchanging system. Computer-assisted survey management

Today, field interviewers located throughout the country and CATI interviewers operating from Oslo carry out about 160 00 interviews per year. Statistics Norway conducts about 10 interviewer-administered surveys per year, most of them simultaneously (see Appendix 1 for details).

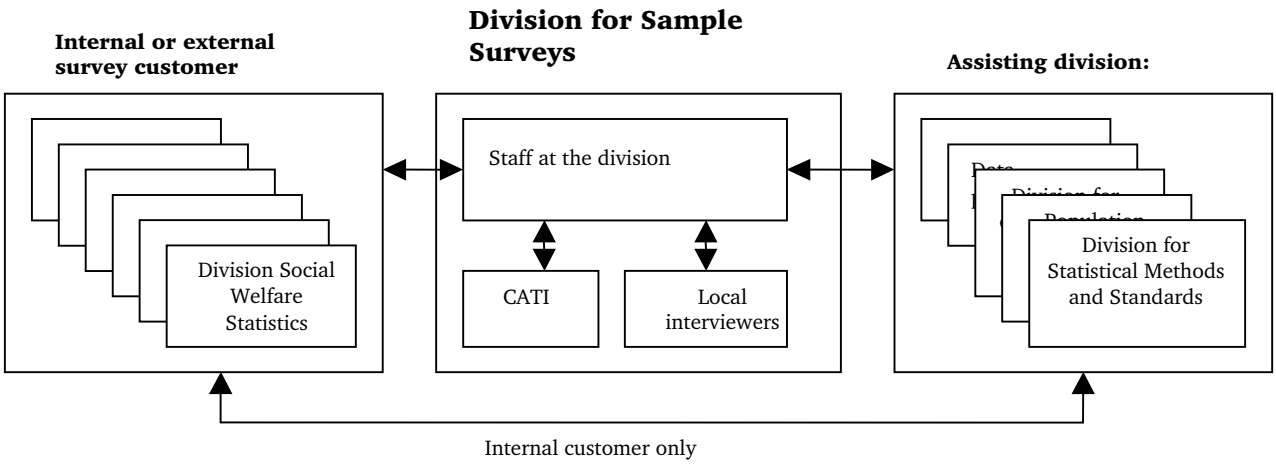
3.2. Infrastructure of household interviewing

In this section we describe the infrastructure of interviewer-administered household surveys in Statistics Norway. By infrastructure we mean the conditions surrounding the work process (environment), the equipment and tools used in the process (technology), the people doing the work, the materials being processed and the way it is all executed (method). There will of course be differences between surveys, but there are also many similarities between them.

Internal organisation

With a staff of around 30, the Division for Sample Surveys is responsible for all interviewer-based data collection. This involves the planning, programming and daily monitoring and supervision of the data collection. Topics, questions, and samples are proposed and decided by the division in charge of the statistics in collaboration with the Division for Sample Surveys. Only the survey division staffs have communication with the interviewers. The Division for Statistical Methods and Standards is responsible for the standardised sampling frames and assists on methodological issues in sampling design and estimation. The Division for Data Collection Methods assists in matters of questionnaire design. Statistics Norway also undertakes surveys for outside clients on commission.

Figure 3.1. Organisation of sample surveys in Statistics Norway



Interviewers

There are 150 decentralised interviewers and 35 centralised interviewers. The 150 decentralised interviewers are located within the 109 primary sampling units of the standardised sampling frame. The 35 centralised interviewers are working in the CATI facility in Oslo. The interviewers are regularly employed, and there is a tariff agreement regulating the working conditions. Interviewers can choose between 500 and 700-hour contracts per year. Senior interviewers have a contract of 1200 hours per year. All interviewers work on several surveys at the same time, normally 3-5 per week. About half of the local interviewers employed are males, and ages vary. Over 50% have been working more than 5 years, and 25 % have been in the job for over 10 years.

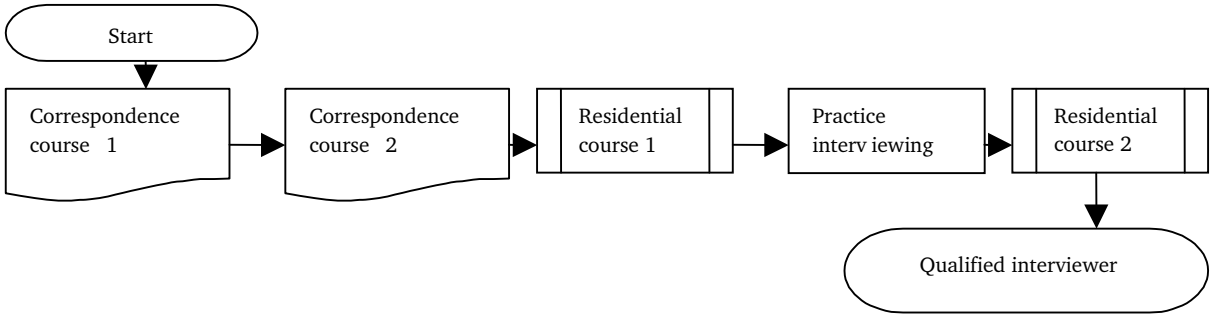
Table 3.2. Local interviewers, by gender, age and years in the job, in percentages

	Total	Men	Women
Gender	100	54	46
Age			
Under 40	12	6	18
40-49	15	13	17
50-59	45	40	52
Over 60	28	41	13
Years			
1 year or less	23	27	18
2-3	22	24	20
4-9	28	23	35
Over 10 years	26	26	27
N	130	70	60

Basic training of interviewers

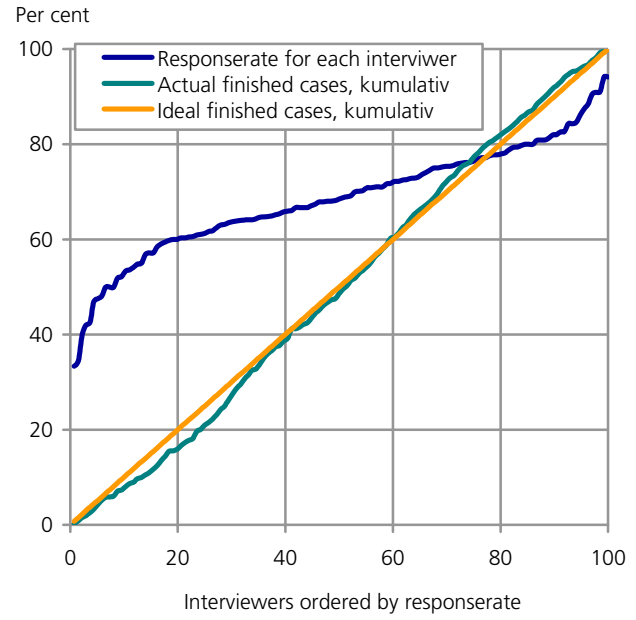
The basic training of interviewers follows several stages. First they undertake a correspondence course, and then they participate in a residential course of several days with intensive training. Normally there is one course before they start and another one or two months after their first work experience. In this basic training the candidates are role-playing refusal conversion, learning to track the "hard to get". Software has also been developed for training purposes where the interviewers shall give a response to a programmed interviewer object unwilling to participate.

Figure 3.2. Basic training of interviewers



One member of staff is responsible for following up new interviewers, and to advise them on situations they perceive to be difficult. New interviewers are also participants in telephone conferences with other more experienced interviewers. How to deal with refusals and tracing strategies are discussed in these conferences.

Figure 3.3. Response rate among interviewers. Survey on Level of living 2005. 136 interviewers 8500 cases



Continuing training with focus on nonresponse

As is shown in figure 3.3, there are major differences in the response rates among the interviewers. We believe it is crucial to continue the training on nonresponse issues. To keep the interviewers motivated, and to try to let the least successful learn from the more successful, national or regional interviewer conferences with continuing training are arranged. These conferences normally take place over two days. Since the late 1990s, they have been organised every other year. In 2004, the main topic was communication between the interviewer and interviewer object, with focus on how to get the interview, and how interviewers could exchange refusal-avoiding strategies. There was also a strong emphasis on good tracing strategies.
The response rate among the interviewers differs. Figure 3.3 shows the response rate among local interviewers in the Survey of level of living 2005. The response rate varies from 33% to 94%, and the median is 68%. Less than 10 percent of the interviewers have a response rate lower than 50%, and about 15 % have a response rate over 80%.

Response rates also differ moderately between different categories of interviewers (displayed in table 3.3).

Table 3.3. Response rate among local interviewers, by gender, age and years in the job, in percentages. Ordinary assigned cases. ESS 2004 N=2750

	Total
Gender	
Men	63
Women	62
Age	
Under 40	65
40-49	54
50-59	64
Over 60	64
Years	
1 year or less	58
Between 1 and 2 years	67
Between 2 and 4 years	61
Between 4 and 6 years	57
Over 6 years	65
N	130

Statistics Act

One important environment feature is the Statistics Act. Some household surveys such as The Labour Force survey are made mandatory under the provisions of the Statistics Act. The Act stipulates that Statistics Norway is the central body for the preparation and dissemination of official statistics in Norway. The institution is subject to supervisory guidelines and financial frameworks set for its business at any time by the Government and the Storting (the Norwegian Parliament), but the Statistics Act stipulates that Statistics Norway is an independent institution in its field that includes a comprehensive research activity. This means that Statistics Norway is responsible for the total statistical product within the guidelines and budgets set by superior bodies. Statistics Norway also determines the statistical methods that are to form the basis for preparation of given statistics, and is responsible for how and when statistics are published.

Technology

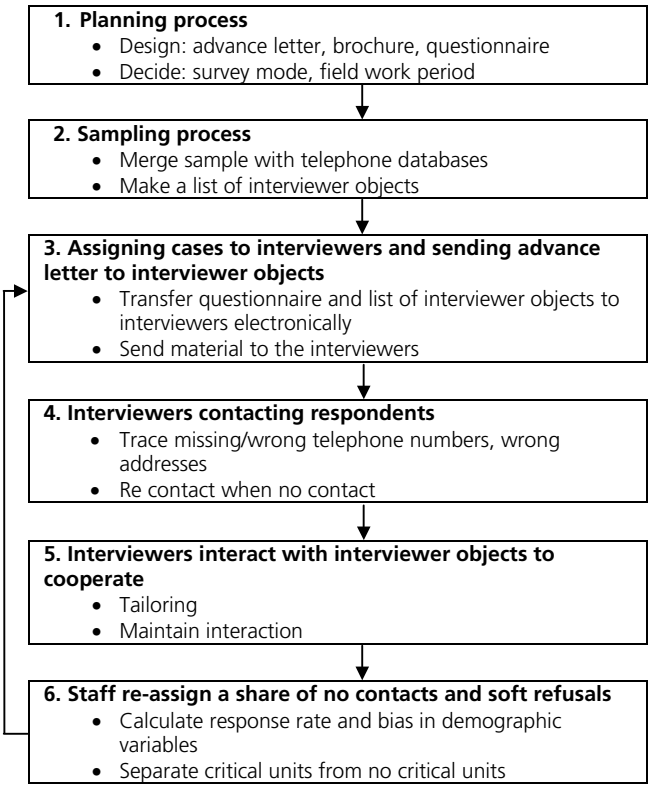
The interviewers in the field are equipped with laptop computers, and there is a fully integrated computer-assisted survey management system. Most communication between the survey division and the interviewers is done electronically. The CATI interviewers work from regular PCs in Statistics Norway's offices and not from laptops. The software used for programming the questioners is Blaise. The field interviewers download questioners and the list of interviewer objects. Completed interviews, non-eligible sample units and nonrespondents are returned electronically to the office via the Internet. This gives the survey managers a daily update of how the interviewers are performing in the field. Every interviewer has a telephone, paid and owned by Statistics Norway, installed in their home.

Registers and databases
Statistics Norway make extensive use of administrative registers. There are about 60 administrative registers. All sampling in surveys is based on registers. A sample unit is normally a person. The sampling frame, which is a working copy of the National Population Register from 1 January 1999, is updated monthly. Other registers, including phone registers, are updated regularly before they are used as sampling frames.

3.3. Processes of interviewer surveys

In this section we describe how we normally carry out a survey. The emphasis is on the process and key variables to monitor during the process. Several examples from recent surveys are included. Before we go into more detail on each process it is useful to give a brief overview of the processes. Figure 3.4 shows a macro flow chart of the processes.

Figure 3.4. A macro flow chart of the interviewer process in Statistics Norway



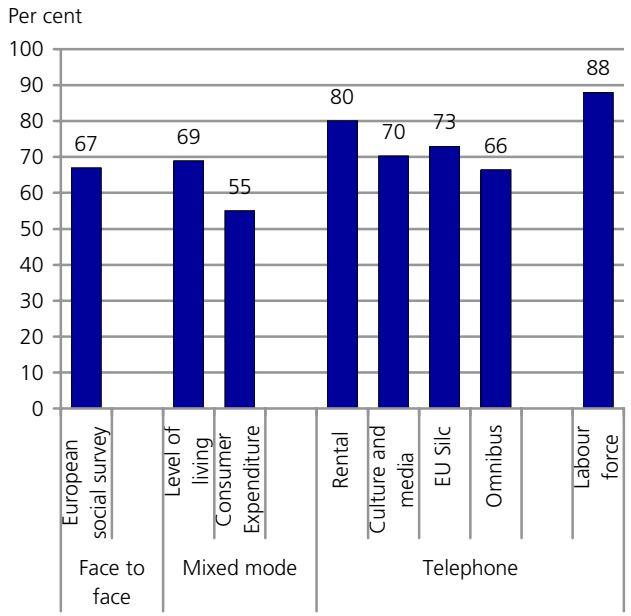
In figure 3.4 we have divided the interviewer process into six steps. A survey always starts with the planning process (step 1). Here the questions in the questionnaire are decided. Which questions to ask and how to word them? What kind of mode should the survey be administered by? Which interviewer objects should be sampled and how long should the fieldwork period be? In step 2, the sample is drawn from the population register, and merged with databases that consist of each sampled unit's registered address, age, gender, household structure and listed telephone number. We always merge the sample with all telephone listings that are available, also in face-to-face interviewing, because it is more cost effective for the interviewers to first try to interact with the respondents by phone. In step 3, cases are assigned to the interviewers. Advance letters are sent to the interviewer objects. We can say that after the advance letters are sent, the fieldwork starts. In step 4, the interviewers try to contact the interviewer object. In this step the interviewers often have to go through several tracing strategies before they come into contact with the selected interviewer object. In step 5, the interviewer's task is to succeed in getting the interviewer objects to cooperate. Interviewers will normally have several weeks to try to establish contact with the interviewer object and succeed in getting a cooperation. In step 6, the staff decides which of the interviewer objects are eligible and regarded as critical, and hence necessary to re-assign.

3.3.1. Planning process

3.3.1.1 Designing the questionnaire
In interviewer-assisted interviewing, the effect on the response rate by the questionnaire design is uncertain. In mail design, the length of the questionnaire has an effect on the response rate. There is no perfect correlation between long questionnaires and lower response rates, but very short interviewer-assisted surveys, like The Rental Survey of 1 minute, get significant higher response rates than surveys taking more than an hour, like The Consumer Expenditure Survey. Keeping the questionnaire as short as possible is probably a good idea, however it cannot be argued that adding a question or two necessarily will lower the response rate. An important feature is the indirect effect a long questionnaire has on the interviewer's willingness to persuade a interviewer object to cooperate. The same can be said if several interviewers think the wording and flow of a questionnaire is poor.

3.3.1.2 Data collection mode
Data collection mode can have an impact on the response rate. Usually, the topic and resources decides which mode to choose. However, nonresponse should also be taken into consideration when deciding this. Mail mode generally gives a lower response rate than telephone and face-to-face. In international literature, it is often stated that face-to-face interviewing gives the highest response rate. However, this does not correspond with the experience in Statistics Norway (figure 3.5). Telephone surveys can produce higher response rates than face-to face, probably because it is easier for some respondents to do the interview over the telephone. Typically, a mixed mode between face-to-face and telephone is recommended, but for a number of topics it will cause serious mode effects. For other topics mode effects are considered less important than nonresponse errors. Mixed mode is common for many surveys conducted by Statistics Norway.

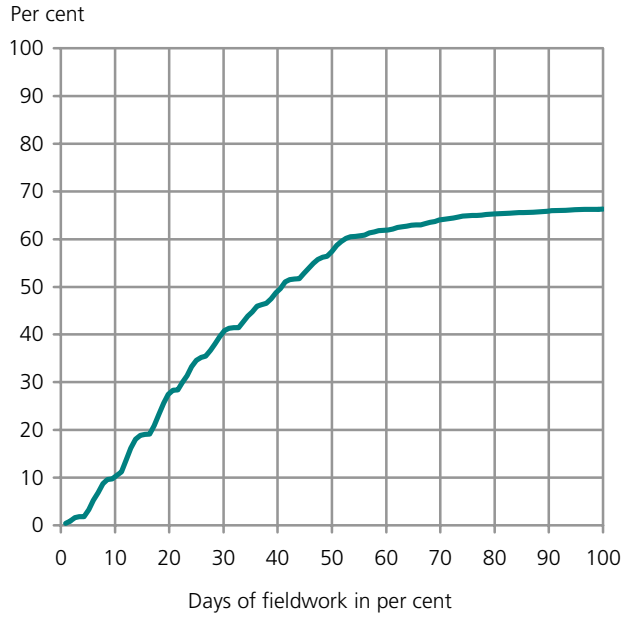
Figure 3.5. Response rate in different surveys 2004, in percentages



3.3.1.3 Fieldwork period

The length of the fieldwork period has a major impact on the response rate, and more importantly on the response bias. A fieldwork period that is too short will result in a high non-contact rate. The shape of a fieldwork period has normally a concave shape, see figure 3.6. It normally takes a short period of time to reach a response rate of about 40–50 percent. It then takes the same period of time to reach the next 20–30 percent.

Figure 3.6. A typical response rate development. ESS 2004.



Despite the fact that the improvement of the response rate is relatively low at the end of the fieldwork period, a relatively long fieldwork period is recommended where possible. Some people are not available or not willing to cooperate in the beginning of the fieldwork period. Hence the fieldwork period has to be sufficiently long so that there is a practical chance of coming into contact with persons when they may cooperate. Another factor is that it is wise to first seek out the "easy to get" and then target the resources on the "hard to get". Since the potential responses are diminishing throughout the fieldwork period, it is reasonable that there is a lower return.

There is no easy way to decide on an ideal fieldwork period, and clearly some of the fieldwork periods in some surveys are too long when they are evaluated afterwards. Analyses should be carried out throughout the fieldwork period. If good data on the number of contacts and the timing of contacts are missing, data on when the interview was done can be used as a dependent variable in analyses on bias and cost efficiency. This will be elaborated later in the chapter.

3.3.1.4 Designing the advance letter and brochure

Research shows that sending an advance letter has a positive effect on the response rate (de Leeuw et al. 2005). The Statistics Act and data protection regulations instruct Statistics Norway to obtain active consent to participate in a voluntary survey and inform the interviewer object of the survey. A rational way to do so is to send a letter in advance. Advance letters work as a quality check on the address register, whereby the envelopes that are returned from the post office indicate that the address in the database is wrong. Actions on those respondents can quickly be taken.

Martin Luppes (1995:462) gives three reasons why it is wise to send an advance letter:

- to announce the visit of an interviewer
- to give information about the survey
- to stimulate the willingness of the interviewer objects to participate in the survey

According to Groves et al. (1992) past research shows that 70-80 percent of respondents who receive an advance letter remember getting it, and of these the same percentage remembers reading the letter. Thus, an advance letter not only eliminates the element of surprise, it also provides tangible evidence that the interview is legitimate and that the call is neither a sales gimmick, nor a practical joke (Dillmann 1978:243).

Advance letters are the starting point in persuading reluctant interviewer objects to answer the survey. Some do not spend a great deal of time or cognitive energy on deciding whether to participate or not. The heuristic decision whether or not to participate is made according to what is called compliance principles (see box).

Some of these principles should serve as a checklist when advance letters are designed. Advance letters in Statistics Norway are highly standardised. The Statistics Norway brand is always on the letter, and it is signed by the general director - this is an example of using the authority principle. A great deal of consideration is given to what the survey should be called. We always put in a sentence in bold letters: that *"we cannot replace you with another person*", the scarcity principle. We often use sentences meant to appeal to their sense of helping others. Qualitative research in information material done by Statistics Norway also concludes that the respondents don't spend much time reading the letter. Hence it should be kept short; a maximum of one page, and appealing, and written in a direct native language, whilst retaining its authority principles. It should not be mistaken for a commercial advertising letter. To compensate for the "one page only" rule, and in order to retain the authority principle, we also design a brochure to give a more tabloid presentation in bright colours. The letter is addressed to a person (and not: "to the household"), which gives it a hint of personification and authority at the same time. The main thing is that the interviewer object remembers the advance letter when the interviewer calls.

The **compliance** principles are principles that individuals use when deciding whether it is appropriate and adaptive to comply in a social activity. Groves (1989) and Groves, Caildini and Couper (1992) have identified six compliance principles of interest in nonresponse research. Luppes, based on Cialdini (1990) introduces a seventh:

- *Reciprocation*, the tendency to participate if a reward is given
- *Consistency*, the tendency to behave in a similar way in situations that resemble one another
- *Social validation*, the tendency to behave according to norms and values of the social group to which one belongs (or thinks one belongs)
- *Authority*, the tendency to be more willing to comply if the request comes from an authority
- *Scarcity*, the tendency to comply because the interviewer object gets the feeling of being in a unique position
- *Liking*, the tendency to comply with requests from attractive requestors
- *Helping*, the tendency to help other people who are in need and who are dependent upon them for aid.

Designing the brochure

The idea behind the brochure is that some people buy the authority principle and some have to be persuaded by other means. The main thing is that the interviewer object remembers the advance letter when the interviewer calls, and a good brochure can aid in this. In the brochure, more details are put in together with graphs and pictures. In one case, we included a picture of two scientists who are well known from television to help interviewer objects figure out what it was about. The information material serves as an important conversation piece the interviewer can use when she/he comes into contact with the interviewer object.

3.3.1.5 The use of incentives

Incentives have proven to have a positive effect on response rates. The use of incentives has a long history in mail surveys, and has also become common in interview-assisted interviewing. A meta analysis of the experimental literature on the effects of incentives (Church 1993, Singer et al. 1999 and Singer 2002) classifies incentives along two dimensions: whether the incentive is a monetary or non-monetary reward, and whether it is prepaid or promised. A prepaid incentive is offered unconditionally with the initial advance letter or questionnaire mailing, while a promised incentive is made contingent on the interviewer object answering the survey. It is useful to also separate the monetary incentives in actual money and a lottery. The most common incentive in Statistics Norway is to arrange a lottery among the ones who agree to participate (Table 3.4), a form of promised incentive. Although this has shown to have little direct effect on the response rate in the literature (see Singer 2002), it is considered to be popular among interviewers when converting reluctant respondents. In the Survey of Consumer Expenditure, the respondents receive a gift card of 300 NOK if they participate. Prepaid incentives are used regularly in the European Social Survey.

Table 3.4. Types of incentives used in Statistics Norway

	Prepaid	Promised
Money		Consumer expenditure
Lottery	ESS, Media use	Level of living, Income and living conditions, Election survey
Gift	Consumer expenditure LAG	Physical activities in the age group 55-75 years

Prepaid incentives have been tested out in several surveys and have given a higher overall response rate; about 4-7 percentage points on average (Table 3.5). The effect differs among different sub-groups, but incentives have a negative effect for very few and small sub-groups. Money as a prepaid incentive has not been tested in Statistics Norway.

Table 3.5. Experiments with prepaid incentives in Statistics Norway

Survey	Incentive	Response rate incentives	Response rate no incentives	Diff
Consumer expenditure 2004	Gift, calculator	56.7 (1 100)	51.0 (1 100)	+5.7***
Media use 2002	Lottery ticket	75.7 (325)	71.7 (325)	+ 4.0
Survey among nurses and nurses assistants 2002	Lottery ticket	31.1 (499)	24.6 (676)	+ 6.5***

Another factor is whether the incentive is a tailoring strategy or not. Tailoring means that some groups of respondents receive an incentive while others don't or different groups receive different incentives. In mail surveys, the ones who have not responded by a given time receive an incentive in the follow-up mailing. In panel surveys, an incentive can be offered to those who refused in previous waves. In the ESS, the number of incentives is increased when refusals are re-assigned, and those who don't return their self-completion questionnaires receive two lottery tickets in the follow-up mailing. Most respondents receive one lottery ticket while some receive five.

Table 3.6. Types of tailoring incentives in Statistics Norway

	Prepaid	Promised
Conditional	Election survey.Survey among nurses and nurses assistants	ALL
Conditional and increased	ESS	

The literature on the effects of incentives (Church 1993, Singer et al. 1999 and Singer 2002) concludes that prepaid incentives yield significantly higher response rates whereas contingent (promised) incentives do not. Prepaid monetary incentives yield higher response rates than gifts offered with the initial mailing, and response rates increase with increasing amounts of money. The effect in telephone and face-to-face surveys is lower than in mail surveys. Another very important factor is that incentives seem to speed up the response, and that using incentives can actually lower the total cost of the survey.

When using incentives, consideration must be given to whether this could affect the respondent's answers. There is no significant empirical evidence that incentives make people change their answers, but there is evidence that it affects the item non-response.

3.3.2 Sampling process

In Statistics Norway, all sampling in surveys is based on registers. A sample unit is a person or family. Address sampling is never used. For most household surveys, the 2-stage standardised sampling frame is used. The Labour Force survey has its own sampling frame. In pure telephone and mail household surveys, a simple random sample in one stage is used.

Merging the sample with the telephone databases

In both telephone and face-to-face surveys, the first contact attempt by the interviewer is by telephone. The sample is matched against public telephone registers before the list is sent to the interviewers. When a person is selected, the whole household is selected at the same time. A match with a telephone number is not, therefore, necessarily with the actual sampled person, but can be with another family member. In face-to-face surveys, there are always 5-10 percent that are unmatched with a telephone number. As an example, we present in table 3.7 the percentage of present telephone numbers in the ESS 2004 survey. Here we see that there is a match of 96.6 %, which is high. The percentage differs a little between different sub-groups. Note that there is a lower match for the 1-person households, non-Norwegians and people living in Oslo. This is also the case for people over 80.

Table 3.7. Telephone number present in the sample of ESS 2 004, in percentages

	Per cent	n
All	96.6	2 750
Gender		
Male	96.3	1 358
Female	96.8	1 392
Age		
15-19	97.5	158
20-29	95.4	413
30-49	96.2	1 007
50-66	98.3	708
67-79	98.3	287
80+	91.0	177
Household		
1-person household	91.8	845
2-person household	98.2	722
3 or more person household	99.0	1 183
Citizenship		
Norwegian	96.6	2 619
Non-Norwegian	88.6	123
Geography		
Oslo	90.7	322
Other part of country	97.4	1 917

A key process variable is the number of unmatched cases with telephone numbers and respondents.

3.3.3. Assigning cases to interviewers and sending advance letter to the interviewer objects

The workload has a major impact on how interviewers perform their tasks. Overworked interviewers produce more nonresponses than interviewers with manageable workloads. The assignment of persons in the sample to interviewers is of vital importance. The workload of each interviewer must be considered. The staff at the Division for Survey sample uses a standardised system that is updated weekly in order to know which

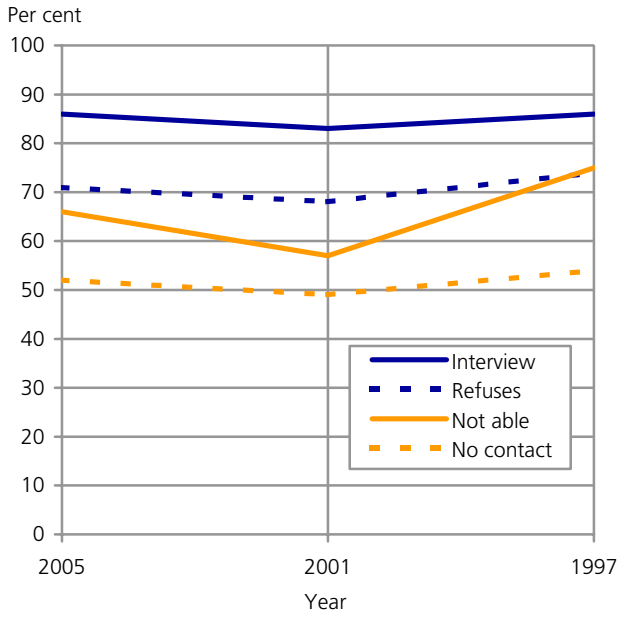
interviewers can be assigned. One of the main things to avoid is assigning new tasks to interviewers that are clearly not capable of getting the job done in time. This standardised system is the basis of weekly staff meetings, where a decision is made on which surveys to increase or decrease interviewer resources.

In Statistics Norway, it is most common that the advance letters are sent from the office to all the interviewer objects at the same time. In a long fieldwork period, as in the Time Use Survey, the interviewer is responsible for sending the advance letter. The effect of the advance letter will decline over the fieldwork period. In long fieldwork periods, the advance letters can only be sent to the respondents that are due to be contacted within the next 2-3 weeks. This can be handled both by the office and the interviewers. There are advantages and disadvantages in both alternatives. There is a gain if the interviewers can tailor the mailing very near the first contact attempt. The down side is that some interviewers do not do this tailoring, and that it is time-consuming for the interviewers.

3.3.4. Contact strategies

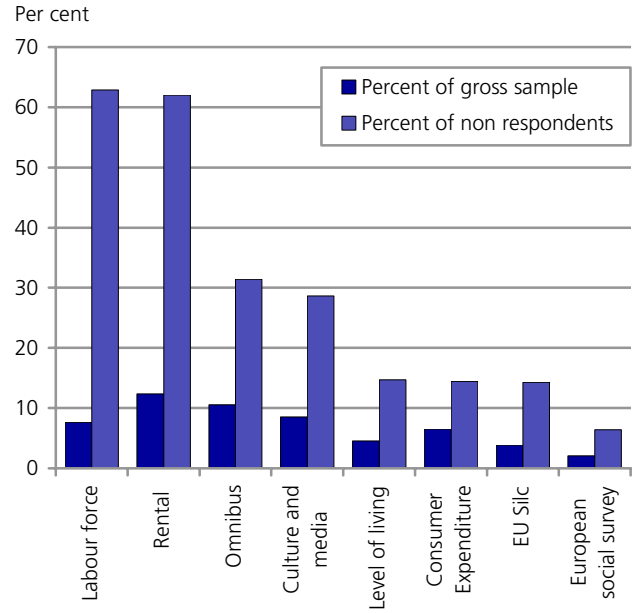
When dealing with nonresponses, we are normally more concerned with those we don't get in contact with than other types of nonrespondents. Much emphasis is therefore put on minimising the non-contact rate. Research shows that the interviewer objects that we don't get in contact with for various reasons are those in the nonresponse category that differs the most from the respondents we succeed in interviewing. As an example, figure 3.7 shows the voter turnout (from register) by different non-respondent categories. Here we see that the non-contact category has a mean 20 percentage points less than the respondents.

Figure 3.7. Voter turnout among respondents and types of nonrespondents. Election Survey 1997-2005



In most surveys, non-contacts are a small share of the total nonresponse as shown in figure 3.8. In surveys with a fixed and short fieldwork period, like the labour force survey and the rental survey, non-contact is a more frequent nonresponse cause. Great efforts are put into the process of minimising the non-contact rate. The following describes the theory and practise that lead to a low non-contact rate.

Figure 3.8. Non-contact rate in different surveys 2004

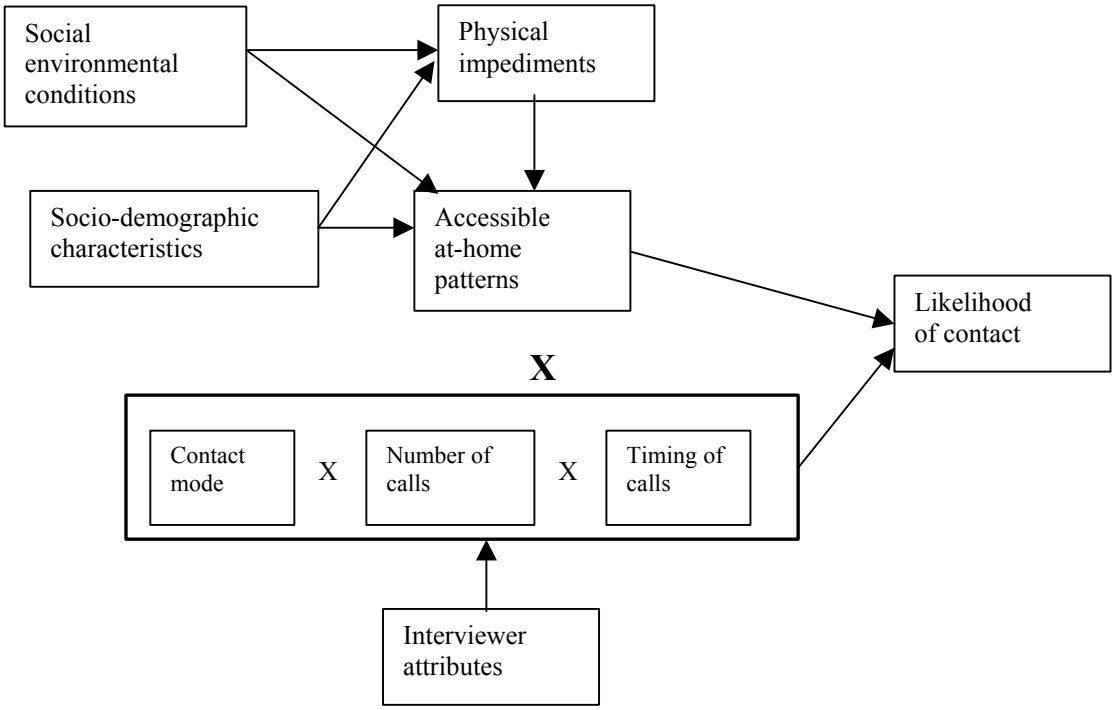


The process of contacting an interviewer object is, in theory, rather straightforward. In the words of Groves and Couper: The success is a function of the times at which at least one member of the household is at home, the times at which interviewers call, and any impediments the interviewers encounter in gaining access to the household (Groves and Couper, 1998:26). Figure 3.9 shows a conceptual model for contacting sample households, originally presented by Groves and Couper (1998). In this model we have also included contact mode, since we believe this to have an important impact on the success of contact attempts. Contact mode, number of calls and timing of calls are part of what we call contact strategies. These strategies have to vary depending on the surrounding attributes.

There are always two or three different contact modes used in the surveys in Statistics Norway. Before the interviewers contact interviewer objects, the office sends an advance letter to the household. Many interviewer objects will be reached within a couple of days. Several of the more difficult to contact interviewer objects cannot be reached just by increasing the number of calls without considering the timing of the calls. Hence the interviewers call at different hours. From the sample list they know the age, gender and household structure of the interviewer object. Some interviewers use this information frequently while others don't. This information

gives more details on when not to call than on when to call. If the interviewers do not get in contact with somebody in the household, either the interviewer object directly or somebody else who can explain when the interviewer object will be at home, other contact strategies have to be used. One simple method is to send a card with the interviewer's telephone number, explaining that the interviewer has tried to reach the interviewer object. The interviewer object is asked to return the card with information on when to call. Relatively few do, so it is more effective to pay a personal visit at the address.

Figure 3.9. A conceptual model for contacting the interviewer object



Source: Adapted from Groves and Couper 1998:26.

In face-to-face surveys there are always 5-10 percent that are unmatched with a telephone number. This can be interpreted as physical impediments. In many surveys those lacking a telephone number will routinely receive a card asking if they have a telephone number on which they can be reached. The interviewers are instructed to try to find a telephone number before they visit the address. If an interviewer visits an address and doesn't find anyone at home, leaving a card is recommended.

It is almost impossible for the interviewers to keep track of every contact attempt they make over the telephone. Hence we have separated this into different modes of interaction, where a telephone call is only registered if someone answers the call. A visit to the address should always count as a contact attempt. Because the office sends the advance letter this must always count as a form of interaction between the survey organisation and the interviewer object or its household. In a survey where the interviewers are instructed to call at different times and to go to the address if someone has not answered within 2 weeks, the non-contact rate is only 7% (Table 3.8). The interviewers come into contact with the interviewer object in about 65% of the cases, 61% by telephone. If they don't reach the interviewer object, they frequently reach somebody else in the household. This means that even in a face-to-face survey, telephone is the main contact interaction channel, with more than 80% of the interaction being done by telephone. Note that the number of personal visits at the address is three times higher, and about 15 percentage points of the total, than the number missing in the match with the telephone registers. In cases where the interviewer doesn't come into contact with the interviewer object, he/she must try to gather information on when the interviewer object might be home. In 25% of the cases this is possible. Where the interviewer doesn't come into contact with anyone, the interviewer just has to call again.

Table 3.8. Different modes of contact in the first interaction attempt in ESS 2004, in percentages

	Interaction with respondent	Interaction with other than respondent	Non-contact	Not eligible
Personal visit at the address	4,6	2,6	7,0	0,2
Telephone	61,2	22,1	.	0,6
Information through survey organisation	0,5	0,9	.	0,3

Table 3.9. Result of the first interaction in ESS 2004, in percentages

	Interview	Appointment	Refusal	Incapacity	Not available	Other
Interaction with interviewer object	2.6	40	14.3	1.3	1.2	8.7
Interaction with a family member	.	9.2	1.5	1.8	13.7	5.9

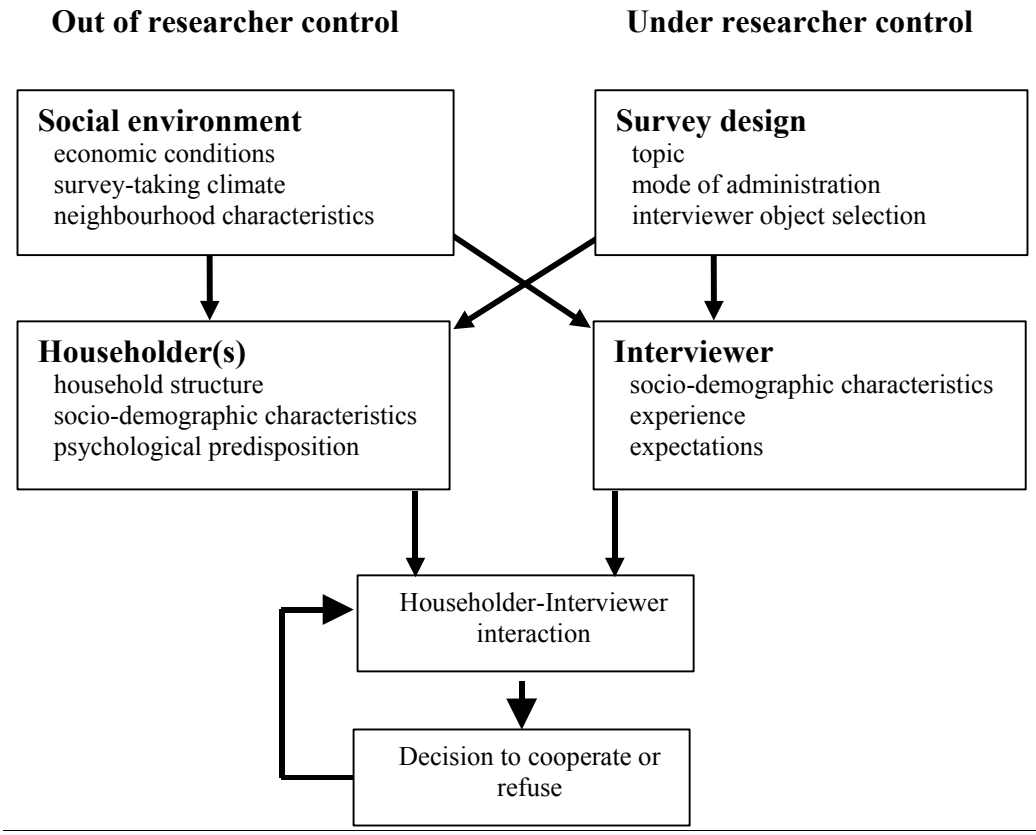
It is important to monitor the non-contact rate for each survey and each interviewer. If the non-contact rate increases in a survey it could increase the bias in the estimates.

3.3.5 Interaction strategies

When the interviewer comes into contact with someone, the job takes another form. Generally there is a small chance of getting the interview right away. This will vary depending on the mode and length of the survey. In a short telephone mode survey, like the LFS, contact with interviewer object usually means that the interview can take place right away. In a long face-to-face interview, normally an appointment has to be made. Table 3.9 shows the result of the first interaction, counting only the cases that resulted in contact with somebody. Only about half of these interactions resulted in interviews or appointments. Fifteen percent resulted in a direct refusal, and another 15% gave information that the interviewer object was not available at the time, and the rest had to be coded as other results. For many of the interviewer objects, it was not hard at all to get a yes.

In figure 3.10, a conceptual framework for survey cooperation in household interview surveys is presented as described by Groves and Couper (1998:30,31). The figure lists influence of the social environment, householder, survey design features, interviewer attributes and behaviour, and the contact-level interaction of interviewers and householders. On the left side are features of the population under study, which is out of the control of the researcher. This is a feature of the social environment and the household. The influences on the right side are the result of design choices made by the researcher, affecting the nature of the survey requests and the attributes of the interviewers who deliver them. At the foot of the figure, the interaction between the interviewer and the householder, where all these influences come into play, is presented. Cited from (Groves and Couper 1998:31): “Which of these various influences are made most salient during that interaction determines the decision outcome of the householder”. We shall not go into detail on the different influences. The important thing to bear in mind is that there are some things that are under the control of the researcher, survey organisation and interviewer, and other things that are not. We need to focus on the things that are under our control.

Figure 3.10 A conceptual framework for survey cooperation



Source: Groves and Couper 1998:30

3.3.5.1 Tailoring and maintaining interaction
The first interaction between the interviewer and the interviewer object, or another member of the household, is a brief conversation. It begins with the self-identification of the interviewer. Normally the interviewer reminds the householder of the advance letter received prior to the household call and adds some more descriptive matters about the reason for the call.

The goal of the interviewer in the first interaction is to avoid getting a no, preferably with the acceptance to call back later. The strategy to use is to maintain interaction. The interviewer must not pose the question of whether the interviewer object should cooperate too early in the conversation.

Interviewers should use the strategies of tailoring and maintaining interaction because people who refuse often refuse because of conditional factors. There is little empirical evidence to suggest that those who refuse are people who never do surveys out of principle. Panel surveys show that people who refuse in one wave often participate in another wave of the survey (see Table 3.10). In the Norwegian election survey, a representative part of the sample (1321 respondents) was part of the survey in three waves; the first after the election in 1997, the second before the election in 2001 and the third after the election in 2001. In the first round, the cooperation rate was 80% (interviewers succeeded in 80% of the cases they came into contact with), while 216 refused, not counting the later non-eligible or non-contacts. Of this group of original "refusers" 35% cooperated in both successive waves, 12% cooperated only in the second, 5% in the last and 48% refused in all three waves. This means that only about 8% of the total eligible sample refused all three times. There is reason to believe that the amount of total refusals will be smaller if the number of waves increases. However, a group of people will probably never cooperate no matter what they are offered. Our point is that this is not a large group and that the refusal rate in a cross-sectional survey can be smaller if the interaction face improves.

Table 3.10 "Refusers" in the first round of Norwegian election survey 1997-2001

	Per cent	N
Cooperate in 2nd and 3rd wave	35	76
Cooperate in 2nd wave, refuses in the other	12	26
Cooperate in 3rd wave, refuses in the other	5	11
Refuses in all three waves	48	103
Total	100	216

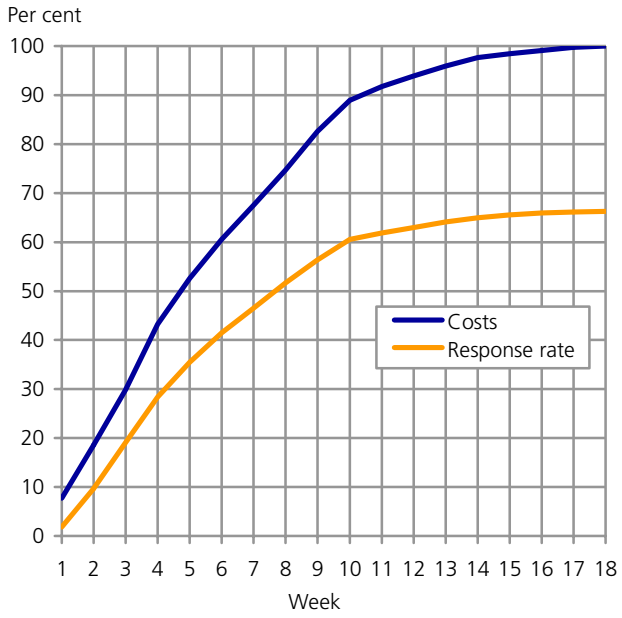
3.3.6 Re-assigning cases

Re-assigning cases means that cases that are sent back from interviewers with a nonresponse status are sent back into the field, often to another interviewer. If it is a refusal, it is normal to send a new advance letter

explaining more of the survey and why it is important to participate. A more personalised style is often used in this letter. The name of the interviewer is often included in the letter. Re-assigning is distinguished from the more routine call-backs that the original interviewer does. These things are, however, connected, with the number of cases that is re-assigned often being a function of how well the original interviewer has performed. A high non-contact rate will give a high proportion of re-assigning. A high refusal rate can indicate that the interviewers have not performed well in the tailoring-maintaining interaction face with the interviewer object; hence this will also give a high proportion of re-assigning.

In a survey organisation, there is sometimes a trade-off between letting the interviewers work on their cases for a relatively long period of time and instructing them to return their cases after a relatively short time. The first will give the interviewers time to do a lot of call-backs and the latter will give the staff at the organisation more flexibility to target the re-assignments.

Figure 3.11. A typical relationship between response rate and costs. ESS 2004



3.3.7. Monitoring the development in response rate and response bias

Monitoring the data collection process, and continuously evaluating the product quality during the fieldwork period is strongly recommended. In nonresponse terms, this means making decisions on whether one should extend the fieldwork period and/or re-assign more cases. Questions to be answered are:

- How much should the variance in the estimates be reduced?
- How much should the bias in the estimates be reduced?
- How much resources should be spent on this?
- Will this influence other surveys?
- Will this seriously delay the publication of the statistic?

The cost of re-assigning cases and/or extending fieldwork periods is relatively high. The late respondents are more expensive than the earlier respondents. In figure 3.11, we see how the curve of the cumulative costs increases more than the return in response rate. This is because the interviewers have to use relatively more time to track the hard to reach. This means more call-backs; visits to the address and so on. One other important factor is that most cases that are re-assigned do not result in an interview. It is vital to take this information into consideration during the planning process. A fair amount of the resources should be allocated for re-assignments in the original budget. Extending the fieldwork period will have a negative influence on other surveys if it exceeds the original fieldwork period. If this is the case, publication of statistics from the survey as well as statistics from other surveys may be delayed.

The trickiest question is the one about bias. To measure bias of a target variable we need to know the population total for this variable. In most cases we don't, and that is why we did the survey in the first place. For demographic variables like gender, age, education, income, we can match the sample against the register. The traditional way is to do this after the survey is finished. In recent years it has become more common to do this during the fieldwork period and to take actions based on this information (see Cobben et al., 2006). In this section we shall present some graphical tools developed for this purpose. The principle behind these tools is that they shall be easy to make and interpret.

3.3.7.1 Graphs based on representation variables

The rationale behind this tool is that the variables used as explanatory variables in the presentation of the estimates, as well as the variables used in poststratification and weighting adjustments, should guide the decision on whether to stop or continue the fieldwork period. In accordance with Holt and Smith (1979): "… it is the structure of the population, rather than the sample design, which an estimator should reflect." This means that some variables must be considered to be more relevant in terms of representativity than others. These variables should be decided upon in the planning process, and then indicate the amount of bias throughout the fieldwork period. The variables should be presented both univariate and multivariate. See the box.

Informed reduction of representation bias during the field work

1. Divide the fieldwork period into milestones, at least 2 and not over 10.

2. Decide representation variables. Keep the number and the categories to a minimum. Representation variables of interest will vary among surveys, but the following seems to be of general interest:
   - Gender
   - Education (high, middle, low)
   - Place of residence (urban, not urban)
   - Age group (under 30, 30-67, over 67)
   - Household size (one person, more than one)
   - Employed (employed, not employed)
   - Income (high, middle, low)

3. Measure the divergence on the variable of interest between the gross sample and the net sample at each milestone.

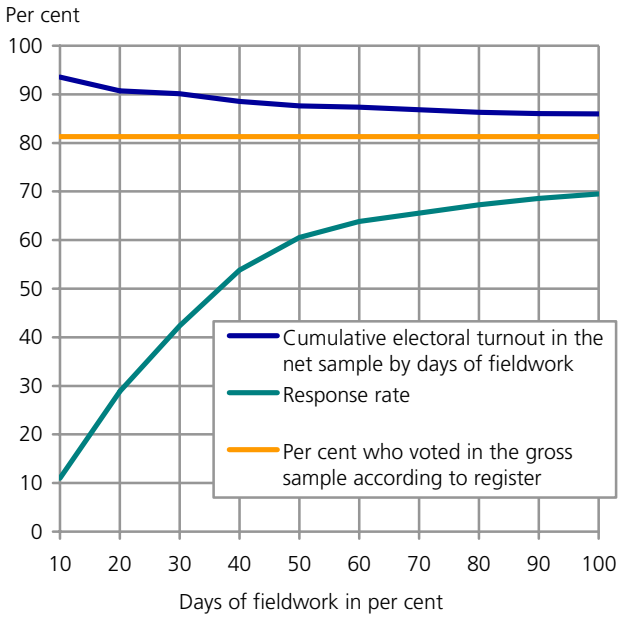
4. Measure the multivariate divergence between the gross sample and the net sample at each milestone (a lower Chi square of the divergence is a good thing).

5. Allocate the interviewer resources where they will reduce the bias the most.

3.3.7.2 Graphs based on target variables

In the same way as for representation variables, it should be decided which target variables are the most important, and these should be measured in the fieldwork period. In figure 3.12, we see how such a process efficiency graph will look after the fieldwork is finished. Electoral turnout is an important indicator of target variables in the Election survey.

Figure 3.12. Cumulative electoral turnout in the net sample by days of fieldwork. General election survey 1997



4. Business surveys

4.1. Introduction

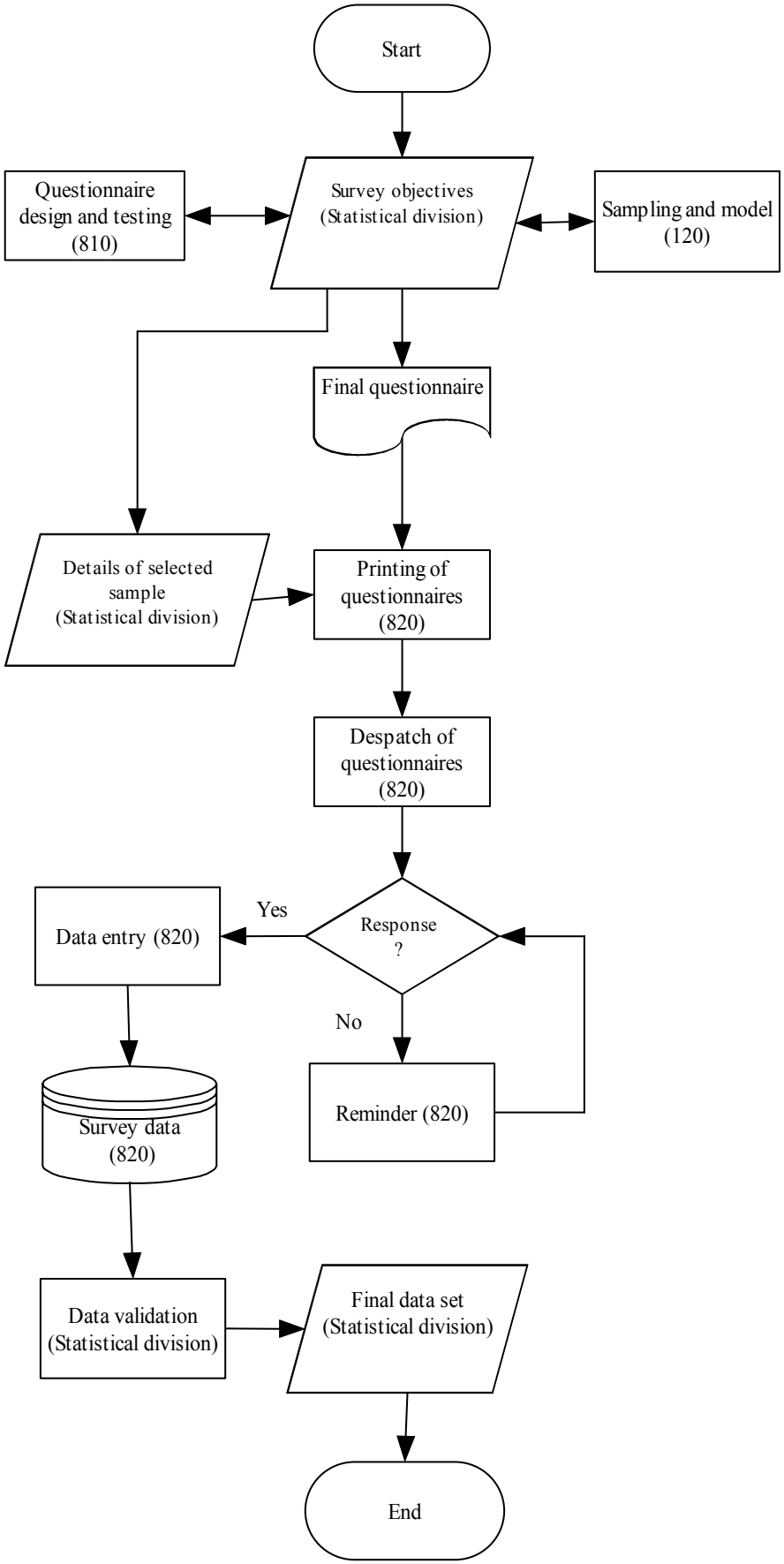
In this section we will give a short presentation of the organisation of business surveys in Statistics Norway. Furthermore, we will use the process approach to investigate nonresponse in some specific cases. Nonresponse in business surveys differs from household surveys because some of the reporting units are of much greater importance than others. Often the number of employees or turnover is used as a measure of size in relation to the stratification, sampling and calculation of population estimates. The impact of nonresponse will therefore to a great extent depend on which units in the sample are missing when the statistics are produced. This section will not give an exhaustive analysis of all effects of nonresponse, but will follow the ideas introduced in chapter 2 and present a systematic approach to the problem, and indicate some tools that can be used to shed light on unit nonresponse related issues.

As we will show later, the detection and storage of metadata connected to the data collection are crucial in order to be able to conduct analysis both during and after the survey period (post survey analysis). For controlling and monitoring nonresponse, these process data will typically be the point in time when the questionnaires were received, what type of questionnaire was used (in the case of mixed mode survey collection) or how many questionnaires were returned due to an incorrect postal address. When these process data are registered and linked to the different units in the sample, they can be used to monitor the different data collection processes and show early signs of critical deviations. The registration of process data is also the key to evaluate possible changes in the processes. To be able to measure the effects of a change in a data collection process that is expected to result in improved quality or reduced response time, we need a basis for comparison from the time prior to the process being altered. Only then can we quantify the effects of a quality improvement initiative.

4.2. Organisation of business surveys

Different statistical divisions in the Department of Economic Statistics (200) and the Department of Industry Statistics (400) conduct business surveys. Much of the data collection process is centralised and performed by the Department of IT and Data Collection (800) in the Division for Data Processing (820). The processes carried out by this division may vary between different surveys, but in most cases it handles the production of questionnaires, dispatch, registration of received questionnaires, data entry and reminders to nonresponse units. In addition, the statistical divisions responsible for business surveys have support from the Division for Data Collection Methods (810) in the preparation of question formulation, questionnaire design and in connection with testing the data collection procedures. The Division for Statistical Methods and Standards (120) also supports the statistical divisions in relation to establishing sampling plans and developing estimation models used to calculate population estimates and accuracy. Figure 4.1 illustrates the interaction between the different divisions. This flow chart shows the establishment of the data collection procedures and the carrying out of a business survey until a final data set is produced. The 'owners' of the different processes are shown with the number of the division in brackets.

Figure 4.1. Organisation of the data collection process in business surveys in Statistics Norway



4.3. Data collection process

In this paragraph we will take a closer look at the different data collection processes established in business surveys. We will also describe some of the measures introduced in an attempt to minimise nonresponse.

4.3.1. The Central Register of Establishments and Enterprises (CRE) as population frame

Common for all business surveys in Statistics Norway are that they use the Central Register of Establishments and Enterprises (CRE) as a population frame. The quality of the CRE is vital for the quality of the different business surveys. This applies to both the correct registration of all active establishments and information concerning the different units. To minimise nonresponse due to the sampling of non-active units or incorrect postal addresses, the statistical divisions and the Division for Data Processing are instructed to send a copy of all received information concerning changes in unit information to the Division for Business Register (410). This division is responsible for updating the CRE with correct information. Some corrections are carried out directly from the statistical divisions (central system for submitting changes). The time it takes from the statistical division receiving new information concerning a unit until the information is correctly updated in the CRE, is crucial for the quality of new samples drawn from the population frame. Updated information on the units in the CRE is also received directly from The Central Register of Legal Entities.

4.3.2. Obligation to provide information and the use of compulsory fines

The Statistics Act of 1989 provides the legal framework for Statistics Norway's activities. This Act provides the authorisation to impose compulsory fines on establishments not responding to compulsory surveys. The majority of business surveys are compulsory and apply compulsory fines when establishments refuse to provide information. There are, however, some surveys that are voluntary. The introduction of compulsory fines in the 1990s has improved the response rate in business surveys, which now often exceeds 95 per cent.

4.3.3. Mixed mode: combined use of Internet and postal questionnaire

As of July 2004 electronic questionnaires are offered for all business surveys in addition to the traditional paper questionnaire. The electronic questionnaires are available through the Internet portal IDUN. When introducing the combined use of electronic and postal questionnaires (mixed mode), the importance of developing recognisable electronic versions of the paper questionnaires has been emphasised. The choice of questionnaire mode may affect the rate of non-response. One of the conditions that need to be met in order to be able to assess the effect that the choice of questionnaire type has on the response rate, is that the time the questionnaires were received and the mode used must be registered. The introduction of Internet questionnaires does not in itself reduce the level of nonresponse but may contribute to reducing the burden the establishments feel when completing questionnaires. In addition to this, the contact person in the establishment has to register his e-mail address to be able to use the Internet questionnaire. This provides the opportunity to send e-mails giving notification that a new survey period is available in the Internet portal and to send reminders to nonresponse units. It can also be used in the data validation process when the respondent has to be contacted. In paragraph 4.4.1, we will take a closer look at the use of process data connected to the choice of questionnaire mode.

4.3.4. Response chasing

With regard to compulsory surveys, follow-up routines for non-received questionnaires are controlled by regulations described in Statistics Norway's handbook number 49 (Obligation to provide information and compulsory fines, only in Norwegian). This handbook presents the instructions for following up unit nonresponse in surveys using compulsory fines. When reminders are sent depends on the survey's frequency (monthly, quarterly, and annual). Please refer to handbook 49 for details concerning this. Information on the obligation to provide information and the fine that will be incurred if the information is not provided is included when the questionnaires are sent. This applies to all surveys using compulsory fines.

Some days or a few weeks (in the case of annual surveys) after the deadline for returning the questionnaire, the 1st postal reminder is sent. This reminder includes information on the survey and the first deadline. It also includes a warning that the unit will be fined if the information is not returned within a new deadline (later referred to as the 1st reminder). If the questionnaire is still not received a few days after the new deadline, a decision is sent explaining that the unit has been fined (2nd postal reminder). The unit is also informed that the fine will be cancelled if the questionnaire is returned within a final deadline.

These routines have turned out to be very effective in generating a very low level of nonresponse in business surveys in Statistics Norway. However, there are a number of establishments that react to the use of these kinds of measures to force out information. Because of these reactions we have to consider whether this kind of pressure on the survey units reduces the data quality. Consequently, it is important to strengthen the follow-up routines based on channels other than the postal reminders. An increasing use of Statistics Norway's Internet questionnaires has led to the registration of a large number of e-mail addresses of

contact persons in the different samples for the different business surveys carried out. Many of the business surveys have started to use e-mail as a new form of contact with the units. This kind of enquiry has a less formal character, and may be perceived as a positive follow-up notice and as help for meeting the deadlines such that the unit avoids being fined for missing questionnaires. The dispatch of e-mails will not cover the entire sample because not all contact persons have registered an e-mail address, but will cover the ones that have previously sent an electronic questionnaire. Today two kinds of e-mails are sent. The first one is sent at the same time the paper questionnaire is sent by post, and informs the respondents that a new survey period is available as well as specifying the deadline for the survey. A link directly to the Internet portal is included in this e-mail. The second e-mail is sent around the first deadline and informs the respondent that the deadline is close and that he will avoid a postal reminder if the questionnaire is returned quickly. A link directly to the Internet portal is also included in this reminder. Experiences from these e-mail enquiries show that establishments are positive to this kind of contact. In addition, a lot of useful information on the units is returned in response to the e-mails sent. This might be information on the status of the establishment (compulsory liquidation, new owner, etc.) or on the contact person (left the company, sick leave, etc.). This type of information is important in order to be able to update the different business surveys' samples and pass on information to the CRE as quickly as possible.

As mentioned in the introduction to this chapter, the importance of different units differs in business surveys. Some units have a major impact on the results due to the fact that they are very large in terms of number of employees, turnover or another variable defining the size of the establishment. In business surveys, the most important units are often characterised as critical units. Nonresponse among these units is often followed up manually by identifying them in the production system of the survey, and subsequently contacting them by telephone. Even if this is a resource-intensive process, this form of response chasing is of great significance to the quality of the statistics.

4.3.5. Flow chart of the data collection process in business surveys

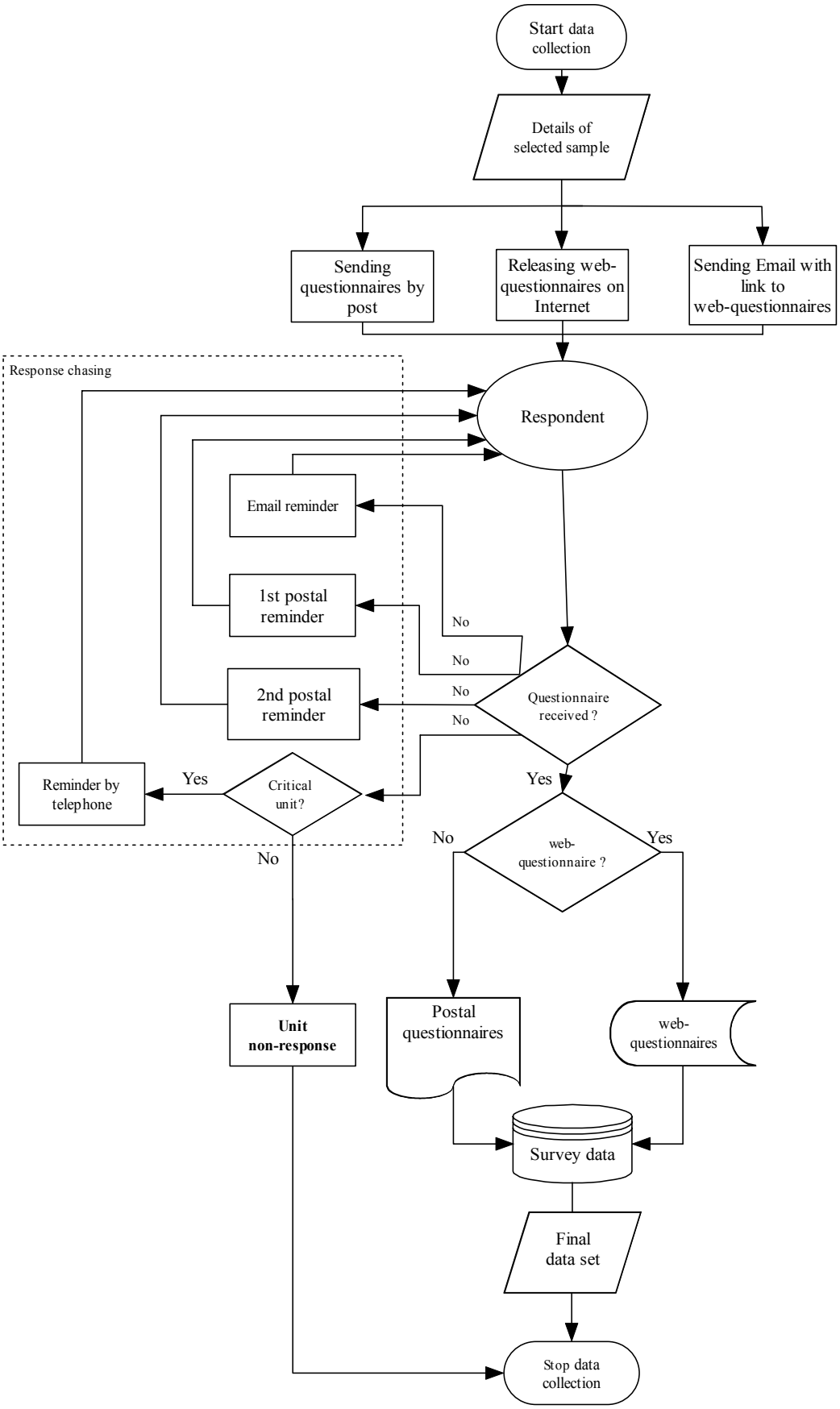
The flow chart in this paragraph illustrates an example of a typical data collection plan for a business survey in Statistics Norway. In this chart, we will disregard which division carries out the process. Deviation from this example may occur, but the fundamental structure will be the same for all compulsory business surveys. The chart will also apply to voluntary surveys, but they do not have the same formal follow-up routine with 1st and 2nd postal reminders. However, these surveys also use a postal reminder for nonresponse units.

The flow chart follows the data collection process for a compulsory business survey using compulsory fines. The process starts with the definition of the selected sample with detailed information on the survey units. Further, paper questionnaires are sent by post and published on the Internet. An e-mail is simultaneously sent to all respondents in the sample with a registered e-mail address, informing them that the questionnaire is now available on the Internet and specifying the deadline for the survey.

What the respondent chooses to do can be measured in the decision box: Questionnaire received? Process data are collected each time a new questionnaire is received. The data registered are the point in time of receipt combined with the identification of the unit. Furthermore, the mode selected by the unit for transferring data is recorded: paper or Internet. These are key process data that may help to gain an insight into the development of nonresponse throughout the survey period or between survey periods. This is illustrated by the use of empirical examples in section 4.4.

In addition to this continuous registration of process data, the flow chart describes four follow-up routines shown within the dotted line defined as response chasing. These routines are carried out on fixed dates and these dates are important process data to record. Usually there will be a fixed number of working days between each follow-up routine. The nonresponse units are identified when one of these 'critical dates' occurs, and new contact is initiated with these respondents. In the first phase, a reminder is sent by e-mail. At the next 'critical date', the 1st postal reminder is dispatched. This reminder contains the warning about the compulsory fine and a new deadline for answering the questionnaire. In the next loop, the 2nd postal reminder is sent with the decision explaining that the unit has been fined, but that they can avoid the fine by returning the data in question by a final deadline. In the last phase, unit nonresponse among the critical units is identified and those classified as critical are contacted one final time by telephone. The number of units included in each follow-up routine will be process data that should be registered and stored for later analysis. By storing these process data, we are able to measure the effects of moving the deadlines or monitoring the data collection process to ensure that an abnormal number of reminders lead to the investigation of the previous processes. An example of this kind of monitoring is given in section 4.4.

Figure 4.2. Data collection process in business surveys



The follow-up routine is cancelled and the questionnaire is regarded as unit nonresponse if the unit is not classified as critical. Unit nonresponse may also occur among critical units, but a great deal of effort is put into ensuring that data from these units are received, and the response chasing does not stop until we have tried to contact the establishment by telephone.

Data from paper and web questionnaires are registered in a database containing both survey data and process data (information about the mode used and time stamp for data entry). The survey data are extracted from this database to a final data set containing survey data for this particular period.

The flow chart in figure 4.2 illustrates how it is possible to systemize and divide the different actions performed throughout the survey period. In addition, it makes it easier to identify the crucial places where process data should be registered. The chart shows the data collection process at a fairly high level. In order to gain a thorough understanding of the different major processes we could have divided certain parts of the chart into lower level charts. This could be applied to a flow chart describing the production and dispatch of questionnaires, or data entry. It is beyond the scope of this handbook to carry out this task, but constructing both high and low level flow charts is recommended when investigating a major process such as the data collection process. This kind of work is currently performed by the Division for Economic Indicators (240), where a systematic documentation process is undertaken by the use of flow charts.

4.4. Identifying and measuring process variables related to nonresponse

The construction of a flow chart for the data collection process makes it easier to separate the different processes and to locate places to collect and measure process data. Measurements of these variables are crucial for evaluating the product quality at a later date. Below is a list of some of the process variables that are important to record in connection with the data collection process:

- Gross sample size at the time of dispatch
- The number of e-mail addresses for the sample units
- The number of returned questionnaires due to incorrect postal address
- The number of e-mail reminders sent
- The number of 1st postal reminders
- The number of 2nd postal reminders
- The number of unit nonresponse among critical units after the final deadline
- The number of critical units contacted by telephone in the last follow-up routine
- Continuous registration of received questionnaires (time and unit identification)
- Continuous registration of questionnaire mode used by the different units (paper or Internet)
- Critical dates (dispatch date, deadlines and reminder dates)

In addition to registering these process variables during the survey period it is important to store these data for comparison between different survey periods at a later date. In this way we can measure the effects that follow changes carried out at a process level. For example, does the introduction of e-mail reminders have a positive effect on the level of unit nonresponse? Do e-mail reminders reduce the number of postal reminders? These questions can be analysed in detail by storing historical process data.

4.4.1. The importance of recording and storing metadata

We will now take a closer look at some specific examples of the registration of process data in connection with the data collection process, and how these data may be used to highlight the development of nonresponse throughout a survey period and across different survey periods. The examples presented are taken from the Quarterly Business Tendency Survey (BTS) and The Quarterly Investment Statistics (QIS), both covering the manufacturing industry. The analysis will be carried out on three different types of process data that are registered and stored for these surveys:

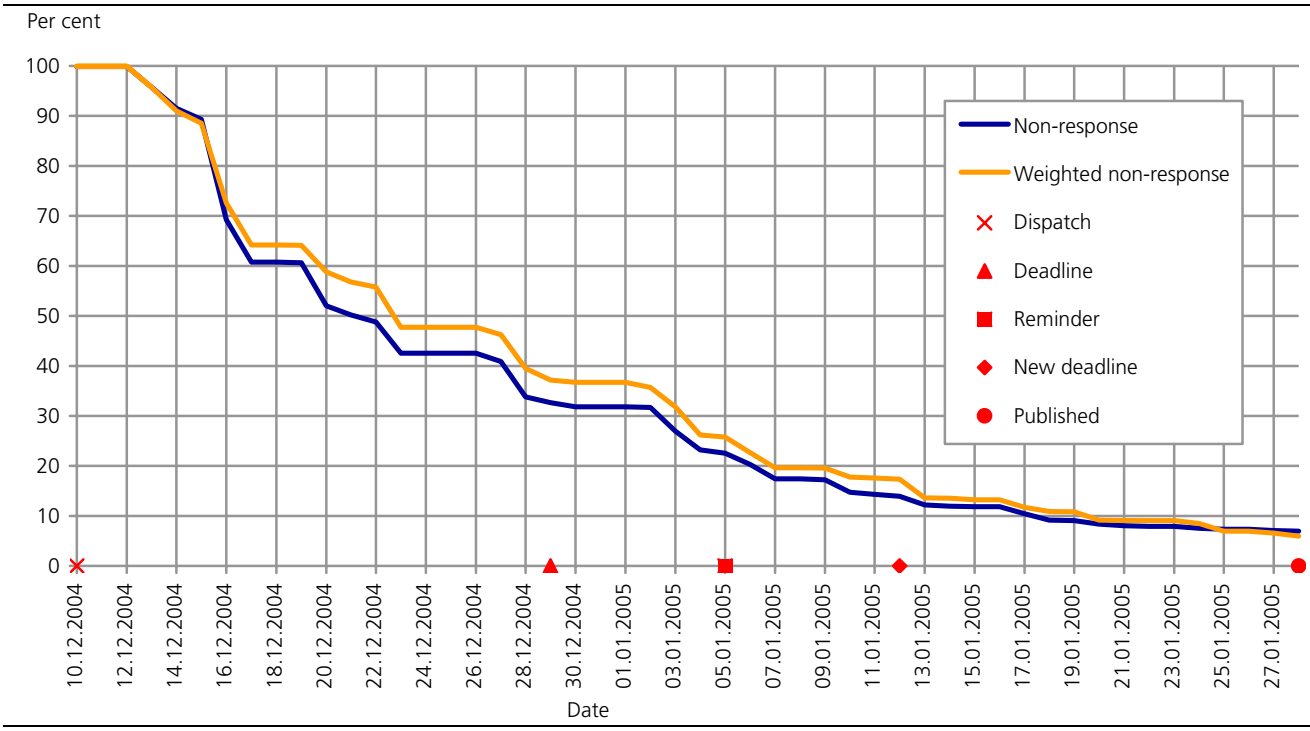
- Time stamp on received questionnaires
- Historical data of nonresponse
- Type of mode used by respondent

These process data are registered continuously throughout the survey periods and are stored in the survey's sample register. The process data can be made available for the active survey period or for previous quarters for comparison across survey periods.

4.4.1.1. Daily measurements of nonresponse during the survey period

In this example, we have used the time stamp from received questionnaires from the BTS for the fourth quarter 2004 to acquire a better understanding of the development of nonresponse throughout a survey period and to measure the effect of the different follow-up routines. In order to be able to carry out this kind of analysis we have to register the point in time each questionnaire are received. This means that the Division for Data Processing (820) signs in all questionnaires received by post in the survey's sample register on a daily basis. In this process it is not necessary to register the survey data, only which unit has delivered and on what date. The survey also has a web questionnaire. Data from this electronic alternative is automatically updated each morning.

Figure 4.3. Weighted and unweighted nonresponse in the Business Tendency Survey. 4th quarter 2004



The BTS is a voluntary business survey that maps out the industrial management leaders' judgement of the business situation and the outlook for a fixed set of indicators. The number of employees is used as a stratification variable in the sampling process. Employment weighted results (response distribution for each question) are estimated for each stratum in each 3-digit industrial group (NACE). The response for each unit is given a weight equal to the number of employees. For aggregation to the industrial group level and totals, the stratum results carry a weight equal to the stratum population employment.

As mentioned earlier, the impact of nonresponse depends on which units are missing. Units with a large number of employees have a greater significance for the quality of the estimates than units with a smaller number of employees. In order to analyse this in greater detail we have calculated both simple unit nonresponse (unweighted) and weighted nonresponse. The calculation of nonresponse is carried out for each day in the survey period, from the dispatch of questionnaires up to the publication date.

Unweighted nonresponse:

$$(4.1) \qquad y_{u,t} = \frac{n_g - n_{n,t}}{n_g}$$

Weighted nonresponse:

$$(4.2) \qquad y_{w,t} = \frac{\sum_{i=n_g} x_i - \sum_{i=n_{n,t}} x_i}{\sum_{i=n_g} x_i}$$

n_g : Number of units in the gross sample

$n_{n,t}$: Number of units at day t in the net sample

x_i : Employment for unit i

Figure 4.3 shows how the rate of nonresponse drops from 100 per cent to 5.9 in the case of weighted nonresponse, during the survey period. By looking at the figure we can determine that the weighted nonresponse was higher than the unweighted in the first part of the survey period. This indicates that there is an overrepresentation of units with few employees that deliver their response early in the survey period. However, this difference is evened out in the latter part of the period, and by the time of publication, the weighted nonresponse is lower than the unweighted: 5.9 compared to 7.0. This suggests that the work carried out to follow up the critical units (the ones with the highest number of employees) by telephone has a positive effect on the nonresponse. It might also suggest that it is more demanding for the critical units to respond to the questionnaire compared with the small businesses. It could be the case that the larger businesses have to wait until key economic variables are available in their internal systems before answering the questions in the BTS questionnaire. Regardless of this, if we look at the two graphs in figure 4.3, the differences of the weighted and unweighted are not particularly large. This means that the received questionnaires are more or less evenly distributed among large and small businesses in relation to the gross sample's composition.

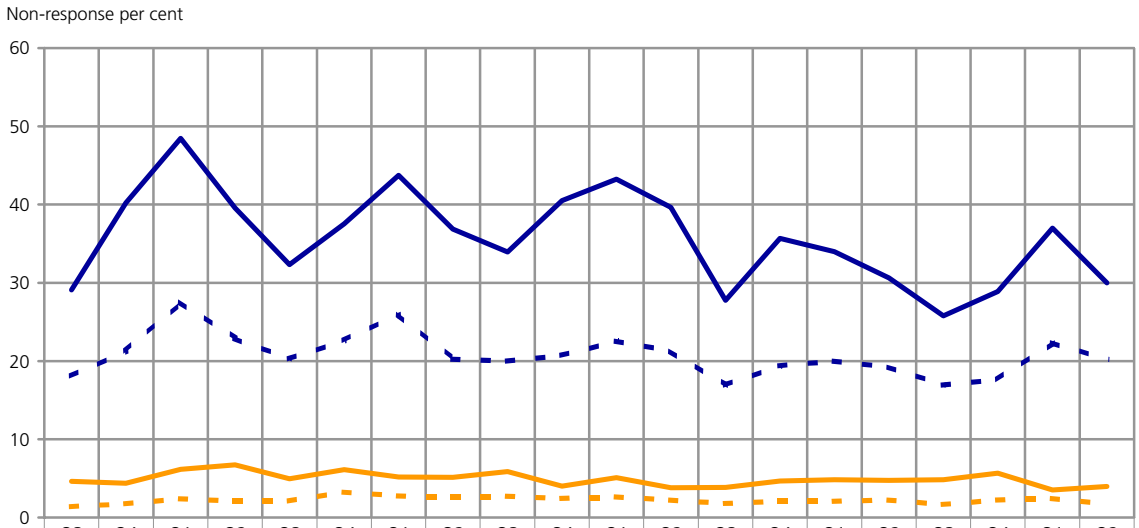
The various critical dates connected to the data collection are marked on the figure's time axis. The development of nonresponse around these dates indicates the impact of the different follow-up routines on the response rate. The figure shows that there is a drop in the rate of nonresponse prior to the first deadline. We also discover that the weighted nonresponse drops from about 26 per cent by the time we send the reminder to about 17 per cent by the time we reach the new deadline stated in the reminder. The reminder by telephone to the critical units is not marked in the figure because this is a process lasting for several days in the latter part of the survey period, but we find that the weighted nonresponse is reduced from 17 per cent at the new deadline to about 6 per cent at the day of publication.

As shown in this empirical example, a simple presentation of nonresponse per day throughout the survey period can provide valuable information on the response process. In chapter 4.5, we will use results from this presentation of process data in a larger context in order to evaluate the impact of nonresponse on product quality, which in this case is the accuracy of the survey estimates. When this type of daily registration of process data is stored appropriately, we will also be able to compare profiles of nonresponse between different survey periods. In this way we will be able to investigate whether there are seasonal effects influencing the profile of nonresponse in connection with holiday periods like Christmas, summer and Easter. We will also be able to compare response profiles before and after an expected quality improvement in the data collection process has been established. This may for example be the introduction of e-mail reminders prior to the deadline. We will be able to measure the real effect of the inclusion of this new follow-up process by comparing survey periods before and after the data collection process has been altered.

4.4.1.2 Detection of data collection problems

In this example we will take a closer look at comparisons of process data between survey periods. Process data for this example is taken from the QIS. The survey measures the level and development in actual and estimated investments for the manufacturing, mining and quarrying, and electricity supply industries, and has a gross sample of about 1 850 establishments. The survey is compulsory and uses compulsory fines for unit nonresponse. In the data collection process, a system has been established for recording sample size when questionnaires are sent and nonresponse when the deadline expires, at the time of the 1st reminder and 2nd reminder and at the time of publication. This registration is carried out each quarter and therefore enables us to investigate how the nonresponse develops between survey periods. The process data recorded is the number of units covered by each action described above, and in this case the weighted nonresponse is not calculated. In figure 4.4, we have plotted unit nonresponse at four critical dates linked to the data collection process: Nonresponse at the deadline, at the 1st and 2nd reminder and at the publication date. The development in these nonresponse rates is presented from third quarter 2000 until the second quarter 2005.

Figure 4.4. Development in nonresponse at critical dates in the data collection process for the QIS



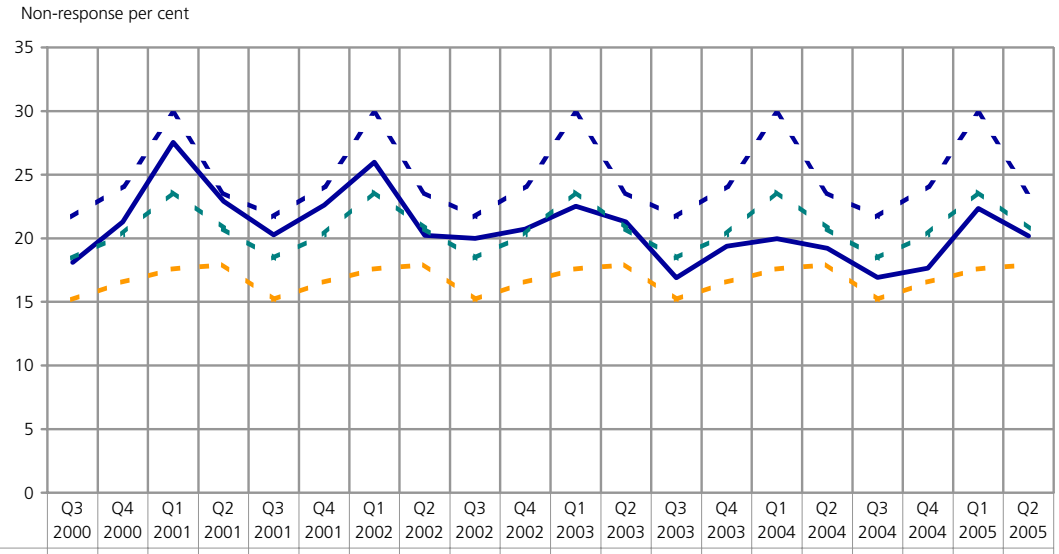
	Q3 2000	Q4 2000	Q1 2001	Q2 2001	Q3 2001	Q4 2001	Q1 2002	Q2 2002	Q3 2002	Q4 2002	Q1 2003	Q2 2003	Q3 2003	Q4 2003	Q1 2004	Q2 2004	Q3 2004	Q4 2004	Q1 2005	Q2 2005
Deadline	29,1	40,2	48,5	39,6	32,3	37,6	43,7	36,9	33,9	40,5	43,2	39,7	27,8	35,7	34,0	30,6	25,8	28,9	37,0	30,0
1st Reminder	18,1	21,3	27,5	22,9	20,3	22,6	26,0	20,2	20,0	20,7	22,5	21,3	16,9	19,4	20,0	19,2	16,9	17,6	22,3	20,2
2nd Reminder	4,6	4,4	6,1	6,7	4,9	6,1	5,2	5,1	5,8	4,0	5,1	3,8	3,9	4,6	4,8	4,7	4,8	5,6	3,5	4,0
Release date	1,4	1,7	2,4	2,1	2,1	3,2	2,8	2,6	2,7	2,4	2,6	2,2	1,8	2,1	2,1	2,2	1,7	2,2	2,4	1,7

Figure 4.4 shows a distinct seasonal pattern in the rate of nonresponse at the deadline, where the nonresponse is clearly at its highest in the first quarter and at its lowest in the third quarter. The reason for this is presumed to be that the deadline for the first quarter is placed early in February. In this quarter, the units are asked to report final investments in the fourth quarter for the previous year, and the establishments cannot report these figures before annual accounts are ready and the results for the fourth quarter can be adjusted to match the previously reported investments reported in the first three quarters. With regard to the low level of nonresponse in the third quarter, this is presumably due to the fact that the questionnaire is sent two weeks earlier than in the other three quarters. This is because of the summer holiday. The deadline is placed at the beginning of August and many of the respondents wish to complete the questionnaire before the holiday period.

There seems to be an equivalent seasonal variation at the time of the 1st reminder, which is sent about one week after the deadline. If we look at the nonresponse rate at the time of the 2nd reminder and when the survey results are released, the seasonal variation is no longer apparent.

If we look at the development over time, the figure shows that the rate of nonresponse at the deadline is reduced during the period from the third quarter 2000 to the second quarter 2005 compared with the same quarter for previous years. We also note that the seasonal variation is less dominant in the quarters following the third quarter 2003. One reason for this might be that the Internet questionnaire was made available for all respondents participating in this survey as of the year 2003. However, even if the nonresponse is reduced at the deadline, the nonresponse at the time of the 2nd reminder and at the publication date was fairly stable throughout the period 2000-2005. In figure 4.5, we present a tool that can be used to monitor the rate of nonresponse at critical dates. The tool is called a *control chart* and defines the outer boundaries for the accepted variation in a process. If a process variable is measured outside one of these boundaries an investigation should be undertaken to disclose the source of the abnormal variation. In this way, we can take measures at an early point in time and try to solve the problems before it is too late. In our example of nonresponse, these abnormal variations may be caused by the fact that the prepaid reply envelopes have not been included when the questionnaire was sent, or that an error in the electronic questionnaire is making it impossible to send data via the Internet. If the nonresponse is monitored at critical dates we may be able to take measures at an early point in time to correct the reason for the extreme level of nonresponse. For example, correct the error in the web questionnaire, or forward the missing reply envelopes to the respondents as an ad-hoc process. In figure 4.5, we use a control chart to investigate the nonresponse rate when the 1st reminder is sent for QIS. In addition to the real nonresponse rate we have calculated the average rate for each quarter and defined the boundaries as two times the standard deviation of the quarterly averages.

Figure 4.5. Control chart for nonresponse in QIS for 1st reminder



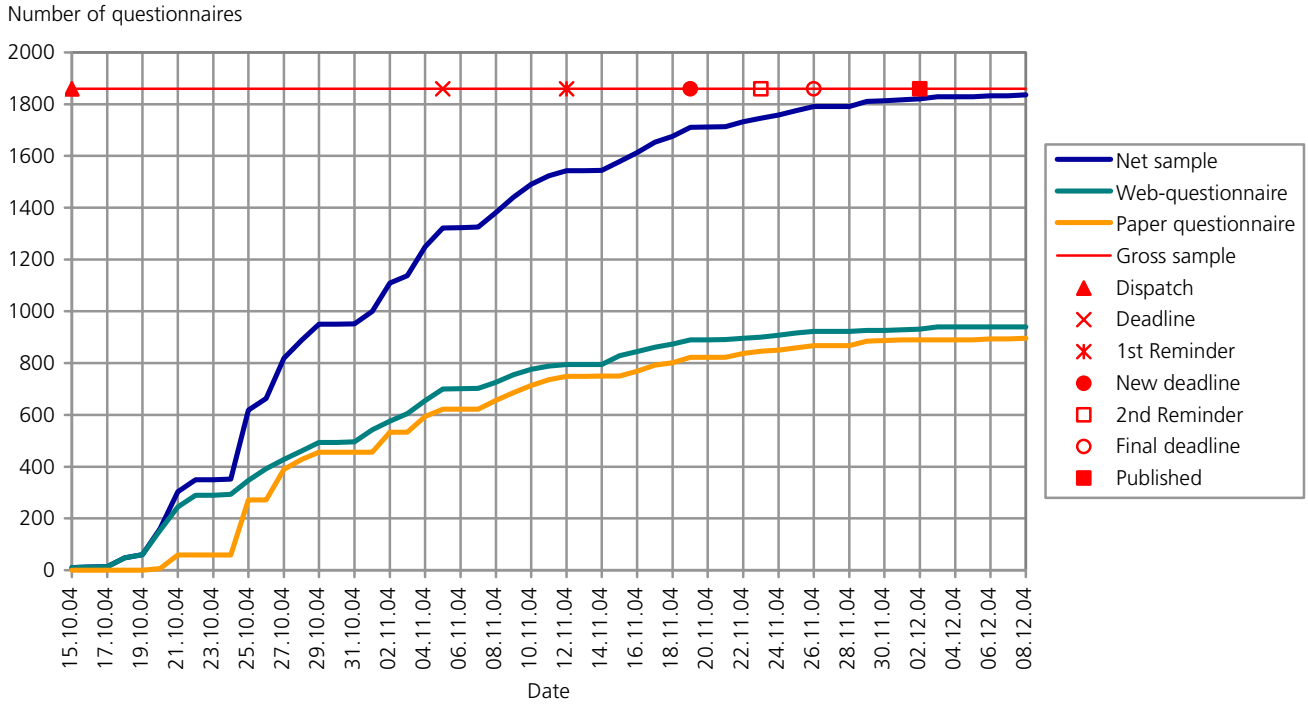
	Q3 2000	Q4 2000	Q1 2001	Q2 2001	Q3 2001	Q4 2001	Q1 2002	Q2 2002	Q3 2002	Q4 2002	Q1 2003	Q2 2003	Q3 2003	Q4 2003	Q1 2004	Q2 2004	Q3 2004	Q4 2004	Q1 2005	Q2 2005
1st Reminder	18,1	21,3	27,5	22,9	20,3	22,6	26,0	20,2	20,0	20,7	22,5	21,3	16,9	19,4	20,0	19,2	16,9	17,6	22,3	20,2
Average + 2*Stdv	21,7	24,1	29,8	23,6	21,7	24,1	29,8	23,6	21,7	24,1	29,8	23,6	21,7	24,1	29,8	23,6	21,7	24,1	29,8	23,6
Quarterly average	18,4	20,3	23,7	20,8	18,4	20,3	23,7	20,8	18,4	20,3	23,7	20,8	18,4	20,3	23,7	20,8	18,4	20,3	23,7	20,8
Average - 2*Stdv	15,2	16,5	17,6	17,9	15,2	16,5	17,6	17,9	15,2	16,5	17,6	17,9	15,2	16,5	17,6	17,9	15,2	16,5	17,6	17,9

Figure 4.5 shows that we have not had any periods where nonresponse rates exceeded the outer boundary of what is tolerated. We also notice that the nonresponse is below the quarterly average in the latter quarters. If we believe that the lower nonresponse has stabilised in the later quarters and that this level constitutes the new lower level we should adjust the boundaries in the control chart by only using nonresponse data from 2003 and onwards in the calculations. The only data needed to establish this tool is the nonresponse rate at critical dates.

4.4.1.3 Mixed mode empirical example

As we have stated earlier, all business surveys offer an electronic alternative in addition to the traditional paper questionnaire. The electronic mode is in most cases a web questionnaire that can be filled in and sent via Statistics Norway's Internet portal, but other modes are also available for certain surveys, such as e-mail response, data on diskette and data extraction directly from the businesses' economic systems. This means that all business surveys offer multiple response modes concurrently. To analyse if the introduction of web questionnaires has influenced the response rate we can use the process data that marks the received data with the type of questionnaire mode used. In the flow chart in figure 4.2, which describes the data collection process, these process data are registered in the decision box: Web questionnaire? In this example, we have connected two types of process data: The registered point in time the questionnaire was received and the mode used for data delivery. In this way we are able to construct a response profile for each response mode. In figure 4.6, we present these process data for the QIS for the survey period fourth quarter 2004. This survey offers two questionnaire modes: paper and Internet. The analysis of these process data can provide useful insight on the effect of mixed mode. We will investigate if the introduction of web questionnaires has resulted in shorter response times, and if users of the electronic alternative have another response profile than those choosing paper questionnaires.

Figure 4.6. Response profile for the QIS 4th quarter 2004 for web and paper questionnaire



The figure displays the number of questionnaires received each day from the day the questionnaires are dispatched until the day the results are released, for the different questionnaire modes. In addition to this, the sum of all questionnaires received (net sample) is included in the figure. The gross sample is marked with the horizontal line and it shows that about 1 850 questionnaires were dispatched. The critical dates in the data collection period are also marked on this line. The survey did not use e-mail reminders at this time.

The figure shows that the receipt of web questionnaires is far greater than the paper version in the first 10 days of the survey period. However, this is due to the fact that the paper questionnaires were not registered in the period 21.10-25.10, but gathered and registered on 26.10. If we disregard this, we see that the response was divided nearly 50/50 between web and paper throughout the survey period. This indicates that the response profile for the respondents for this survey is the same, regardless of the questionnaire mode selected. It is not the case that the users of web questionnaires deliver their responses earlier or later than the ones selecting paper. This also applies to the response rates around the critical dates.

This survey will shortly introduce e-mail reminders. This reminder is to be sent to all respondents with a registered e-mail address who have not replied by the first deadline. This applies mainly to users of the web questionnaires, who have to register their e-mail address in order to be able to use the Internet portal. When this routine is established, it will be crucial to compare the response profiles before and after the e-mail reminder is introduced. In this way we will be able to measure the effect of the change in the data collection process.

4.5. The use of key process variables to evaluate product quality

In section 4.4, we presented examples of measurements of process variables concerning nonresponse and the use of some simple tools for monitoring and analysing the development of nonresponse. In this section we will use measurements of process variables together with measurements of product quality. In this way we can evaluate the effect nonresponse has on the overall product quality. The examples presented are derived from the BTS covering the manufacturing industry.

4.5.1. The impact of nonresponse on product quality during the survey period

The product quality can be expressed as the bias and accuracy of the survey estimates we publish. In this example we will use the results from 4.4.1.1, where we calculated weighted and unweighted nonresponse throughout the survey period for the BTS in the fourth quarter 2004. We will place these results together with with population estimates for one single question in the BTS, calculated at different dates throughout the survey period. To evaluate the accuracy we shall use the estimated sample variance of the estimate, developed by Mevik (2004). The sample uncertainty is estimated on the basis of the net sample received on different dates throughout the survey period. It is assumed that the distribution of nonresponse throughout the survey period is random. This is supported by the findings in Wang (2004).

Before we present the results we will briefly explain some central terms used in the BTS. The results from the BTS are presented as diffusion indices. The diffusion index is a measure of how the units judge a specific variable, for instance the level of their production in the next quarter compared to the current quarter. More precisely the diffusion index for a given quarter is given by

$$(4.3) \quad d = S + \frac{1}{2}U$$

where

$$(4.4) \quad S = \frac{s}{X} \cdot 100 \qquad \text{and} \qquad U = \frac{u}{X} \cdot 100$$

Here s is the number of employees in units that expect a positive development in the level of their production, in the next quarter compared to the current quarter, u is the number of employees in units that expect no change in the next quarter compared to the current quarter, and X is the total number of employees for all units in the population. d denotes the share of employees that are working in a unit that expect the variable in question to increase, plus half the share of employees that are working in a unit that expect the variable not to change.

The diffusion index has a turning point at 50. A common way of interpreting an index value of 50 is to say that half the respondents expects an increase and the other half a fall. An increasing index at or above 50 indicates that the growth rate is increasing, while a falling index above 50 indicates a falling rate of growth. Opposite for an index below 50.

In this example we will estimate the diffusion index at different dates in the survey period for the question "The general outlook for the establishment for the forth-coming quarter[1]". In addition to the population estimates we will calculate a 95 per cent confidence interval surrounding these estimates by using the model-based measure of uncertainty defined in Mevik (2004).

[1] The complete question presented in the questionnaire: How do you judge - generally for the enterprise in this industry - the outlooks for the forthcoming quarter compared with the present quarter? The response alternatives are Better, Unchanged or Worse

Figure 4.7. Process efficiency graph: nonresponse and product quality in the BTS 4th quarter 2004

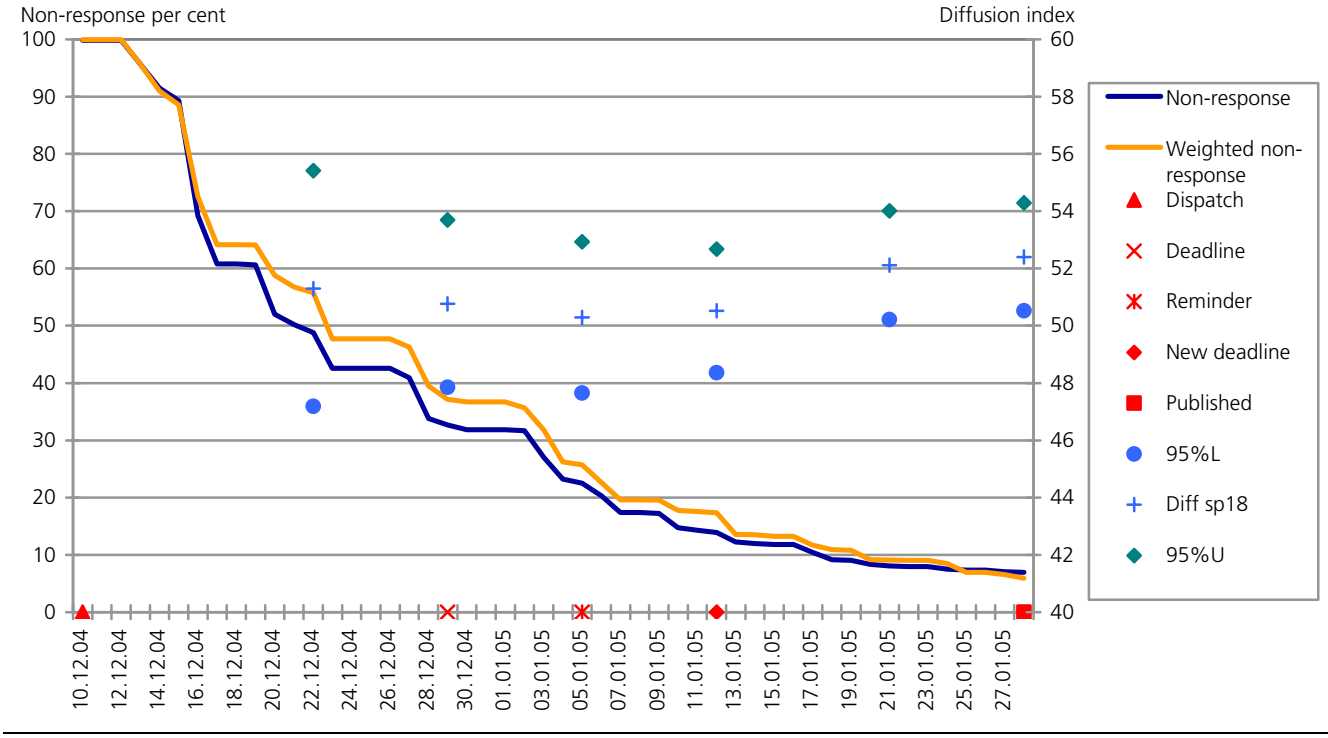


Figure 4.7 depicts the weighted and unweighted nonresponse (on the left axis) and the diffusion index with the 95 per cent confidence interval (on the right axis). In addition, the critical dates are marked on the time axis. We have called this figure a process efficiency graph. It describes the continuous improvement in product quality (increased precision in the estimate) due to the reduction of nonresponse, and in this way explains the efficiency of the data collection process in relation to the product quality. The figure shows that the diffusion index is above 50 for all estimates calculated during the survey period. Further, we see that the confidence interval becomes narrower when the rate of nonresponse is reduced. This is because the sample uncertainty is gradually reduced when we receive new questionnaires and a larger part of the population's total employment is covered by the net sample. Another aspect is that all estimates lie within the first confidence interval calculated. However, it is not before the penultimate estimate (21.01.05), when the rate of nonresponse is below 10 per cent, that the diffusion index is significantly above 50, and we can conclude that the general outlook for the forthcoming quarter is judged as better.

This presentation shows how process data can be used together with a measure for the product quality in order to gain a better understanding of how the dynamics of nonresponse influence the results. The process efficiency graph offers valuable information about the possibility of publishing the survey at an earlier date or the possibility of releasing reliable flash estimates before the final publication. When expected quality improvement measures are introduced, for example increased gross sample, the effect can also be measured by comparing this kind of graph before and after the implementation of the change.

4.5.2. Analysing factors used for calculation of population estimates

The process variable we will now concentrate on is not generated by the data collection process but by the weights used in the stratified ratio estimates in the BTS. The gross sample and the population are held constant throughout the year. Employment is used as weights in the calculation of the response distributions for the different questions, both in relation to the sample unit's answers and between the different employment strata and industries. Each industry (3-digit NACE) is divided into four employment strata:

- Stratum 1 Units with 300 employees and over
- Stratum 2 Units with 200-299 employees
- Stratum 3 Units with 100-199 employees
- Stratum 4 Units with less than 100 employees

In the estimation process it is assumed that the non-response (item and unit) is random within each stratum defined by employment and industry category. By using employment data from the net sample in the estimation, the rate of nonresponse will influence the factors calculated to weigh the different strata together.

Calculations of the proportion of responses on industry level are based on the proportion of responses on the

stratum level. In the transition from stratum to industry, the stratum results are weighted with the population employment to correct for relative differences between the strata in a particular industry. The same principles are used for aggregation from industry to manufacturing total.

The factors calculated to correct for relative differences between the strata in a particular industry are stored in the production system when the survey results are produced. Comparisons between these factors, calculated in the current and the previous quarter, are used in the revision process. The factors are a process variable that represents the net sample's coverage in the individual strata. If the gross sample and the population are held constant throughout the year, then changes in these factors from one quarter to another can only be caused by a changing rate of nonresponse. An increase in nonresponse in a stratum will lead to lower coverage and therefore a higher factor, and a higher factor will produce an increased level of sample uncertainty and reduced product quality (reduced accuracy).

In table 4.1 we present the factors used in the weighting process for industry 22.1 (the publishing industry) in 2004. The table shows that the factor in stratum 1 in the first quarter is higher than in the remaining three quarters. This means that the nonresponse is greater in this stratum in the first quarter compared with the other three quarters.

When the statistics are produced, the production system generates revision tables that list the strata where the population weights are changed compared with the previous quarter. On the basis of this output, the revision process concentrates on following up the nonresponse in the strata where the factors are considerably higher than in the previous quarter. Units in these strata, which delivered data in the previous quarter, are contacted by telephone and asked to return the questionnaire. In this way process data from the production system is used to reduce the effect of nonresponse and increase the product quality.

4.5.3. A nonresponse analysis of the Business Tendency Survey

In the current estimation of the BTS, it is implicitly assumed that nonresponse is missing- at-random (MAR) within each stratum. i.e. that the nonresponse probability depends only on the stratifying variables. Nonresponse is imputed implicitly by treating the net stratum estimates as the gross stratum estimates. Wang (2004) takes a closer look at this assumption to see if it holds, or if it is better to use a more complex modelling of the nonresponse.

The key process variable is now the nonresponse bias. This bias is investigated by using different nonresponse models (weighting) to adjust for nonresponse and different methods of imputation of the nonresponse units.

The results suggest that there is no systematic skewness in the distribution of nonresponse. The estimates of the variable of interest are not significantly influenced by the use of different models for weighting or different imputation techniques. By investigating the different estimates we come to the conclusion that a more complex modelling of the nonresponse does not improve the accuracy of the estimates (product quality).

This is an example of how it is possible to compare different estimation procedures, related to nonresponse, and draw conclusions based on the investigation of key process variables. In this particularly case there was no gain by using a more complex modelling of the nonresponse, therefore the stratified MAR-assumption is still used in the production of the BTS. The same technique may be applied to other surveys, which may give useful insight in the underlying processes generating the nonresponse, and indicating the best method for handling nonresponse.

Table 4.1. Factors used in weighting from stratum to industry level in the BTS 2004 for NACE 22.1

NACE3	Stratum	factor0401	factor0402	factor0403	factor0404	0401-0402	0402-0403	0403-0404
221	1	1.53	1.17	1.17	1.17	0.36	0.00	0.00
221	2	2.38	2.38	2.38	2.38	0.00	0.00	0.00
221	3	2.20	2.22	2.20	2.50	-0.02	0.02	-0.30
221	4	8.66	8.23	9.12	8.23	0.43	-0.88	0.88

5. Weighting for nonresponse bias

5.1. Introduction

In the previous chapters we have presented methods and techniques, which are used to minimize nonresponse rates and/or nonresponse bias during data collection. Once the collection and preparation of data is terminated, the question is to what extend the remaining nonresponse induces biased population estimates, and whether statistical adjustments should be used to reduce these biases. Choice of adjustment method, including "no adjustment", is a survey specific process and the important key process variables are sample mean and sample variance together with nonresponse bias. The purpose of this chapter is to go through the processes, which lead up to the final choice of estimation method. First we present some techniques usually applied to investigate to what extend the nonrespondents differ from the respondents. Then follows a description of the most commonly used weighting methods and some general experiences concerning the effects of weighting on the sampling variance and the nonresponse bias. Finally, we give some brief comments on methods of variances estimation in the presence of nonresponse.

5.2. Use of administrative data to study the effects of nonresponse

For obvious reasons, it is very hard to get a clear and comprehensive picture of the nonresponse bias for all variables in a survey. However, through use of auxiliary information from administrative registers it is possible to gain some insight into the nonresponse bias for some variables. A simple, but useful analysis of the nature of the nonresponse bias, can be done by producing some tables, which compare the distributions of some variables in the selected sample, gross sample with the same distributions in the sample of respondents, net sample. Table 5.1 shows some aspects of the nonresponse bias in the Norwegian Election Survey 2001.

Table 5.1. Effects of nonresponse in the Norwegian Election Survey 2001

	Gross sample			Net sample		
		Gender			Gender	
	All	Males	Females	All	Males	Females
All	100,0	49,3	50,7	100,0	50,1	49,9
Age						
18-30	22,1	11,2	10,9	21,1	10,5	10,6
31-49	38,8	19,0	19,8	39,1	19,4	19,8
50-66	25,5	13,3	12,2	26,3	14,0	12,3
67-80	13,5	5,8	7,8	13,4	6,3	7,2
Education						
Primary School	18,3	8,2	10,1	15,4	7,2	8,3
Second Level	59,5	30,5	29,1	58,9	31,0	27,9
University Level	22,2	10,7	11,5	25,7	12,0	13,7
Municipalities by population size						
Undder 10 000	25,0	12,6	12,3	25,1	13,5	11,7
10 000-19 000	19,7	9,8	9,9	20,4	10,6	9,8
20 000-49 999	23,7	12,0	11,7	23,3	11,3	12,0
50 000 or more	31,7	14,9	16,8	31,2	14,8	16,5

It is seen from table 5.1 that response rates vary between various groups of the population. In particular, it seems as if the better educated are more likely to respond to this survey and therefore the sample of respondents is biased towards the better-educated part of the population.

In order to make table 5.1, it is necessary to have information about respondents as well as non-respondents. In Statistics Norway, a large number of administrative registers are available, which can be linked to the respondents as well as the nonrepondents through an identification number. Tables like table 5.1 are therefore produced routinely for several variables for most surveys. It is generally observed that small, urban households are underrepresented among the respondents.

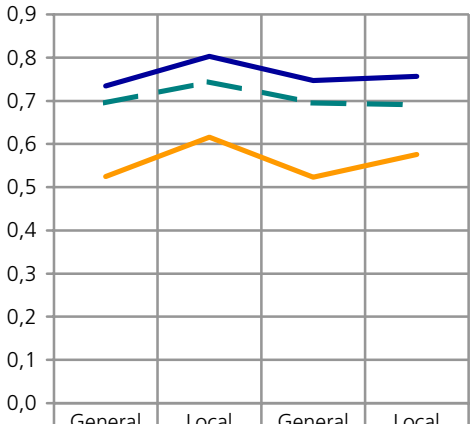
Tables like table 5.1 have essentially 2 purposes:

- In the final report from the survey, similar tables are usually presented in the introduction in order to give the reader some insight into the effects of nonresponse.
- When population estimates are produced, a decision has to be taken whether or not weighting should be used in order to reduce the effects of nonresponse. See chapter 5.4.

Before presenting some of the most commonly used weighting methods, we shall discuss other methods used to gain insight into the effects of nonresponse.

Many surveys are repeated regularly. In such cases, important insight into the nonresponse bias can be obtained by combining information from several rounds of a survey. In table 5.2 response rates are given for voters and non-voters in four election surveys. Information whether a person voted or not is collected from administrative registers established to control the election. It is seen that response rates among voters are systematically higher than among non-voters, leading to a sample of respondents, which is biased towards more politically active persons. This bias is also observed in previous election surveys, see e.g., Thomsen and Siring (1983).

Table 5.2. Response rate among voters and non voters



	General election 1997	Local election 1999	General election 2001	Local election 2003
Voters	0,73	0,80	0,75	0,76
All	0,70	0,74	0,70	0,69
Non voters	0,52	0,62	0,52	0,58

It is also possible to use multivariate techniques like logistic regression to identify variables with significant impact on the response rate. In the Norwegian Health Survey, such an approach was chosen.

Example: The Norwegian Health Survey 1995. (HS 95) (from Belsby and Vedø,1998)
In this survey a sample of 13662 persons was selected to be interviewed, using Computer Assisted Personal Interviewing (CAPI). Interview was obtained from 10 248, giving a nonresponse rate at 25%. In addition to the interview, all persons were asked to fill in a questionnaire containing confidential information about their general health status and return the questionnaire by mail. Among the respondents aged 14-79 to the interview, 14.5% did not return this questionnaire, giving a total unit nonresponse rate of 37%.
In order to study the impact of nonresponse, seven register variables were used as explanatory variables in a logistic model for the response probability p:

$$(5.1) \quad ln\frac{p}{1-p} = \beta_0 + \beta_1 x_1 + \ldots + \beta_7 x_7$$

where $x_1, \ldots, x_7$ are auxiliary variables, which are assumed to influence response rates
and $\beta_0, \beta_1, \ldots, \beta_7$ are unknown parameters.

The following seven variables were chosen as explanatory variables:

- Size of household
- Age of eldest person in the household
- Highest education in household

- Whether or not a household member had received social welfare
- Country of birth
- Place of residence (urban/rural)
- Total household income

All seven variables were found to have statistically significant influence on the response probability.

Table 5.3. P-values for the variables

Variable	P-value
Country of birth	0,0060
Population density	0.0021
Income	0,0001
Age	0,0001
Social welfare	0,0001
Education	0,0001
Size of household	0,0001

P-value of Pearson Chi-square equals 0.1232, indicating an acceptable fit of the model

Based on these findings, it was decided to weight the data in order to reduce the nonresponse bias. The example continues in section 5.4.1.

5.3. Use of information from the interviewers

Another way to gain insight into the effects of nonresponse, which has been mentioned several times in the previous chapters, is to combine responses to the survey with information from the interviewers. A simple method is to divide the nonrespondents into comprehensive and mutually exclusive groups, using reason for nonresponse as dividing variable. Usually it is sufficient to divide into two groups, "refusals" and "not at homes". The rationale is that accessibility and willingness are two important reasons for nonresponse. Using information from administrative registers, the effects of these two potential factors on nonresponse bias can be studied by performing separate analysis on the two groups of nonrespondents. In table 5.4, it is seen that the rate of voting, here called turnout differs among different kinds of nonrespondents. For both "refusals" and "not at home", the turnout is smaller than among respondents. Furthermore, the turnout among "not at homes" is substantially lower than among "refusals". This clearly indicates that the "refusals" are more similar to the general population than the "not at home". Similar results have been found for a number of other variables as shown by Thomsen and Siring (1983).

Table 5.4. Turnout in last elections among respondents, refusals and not at homes/prevented

	Official electoral turnout	Turnout Gross sample	Turnout respondents	Turnout refusals	Turnout Not at homes/ prevented
General election 1997	78 %	81 %	86 %	74 %	62 %
Local election 1999	62 %	68 %	74 %	55 %	47 %
General election 2001	76 %	77 %	83 %	68 %	53 %
Local election 2003	59 %	61 %	67 %	53 %	39 %

In table 5.5, turnout is given as a function of ease of response. The indicator for ease of response is constructed by dividing the respondents into five groups. The first group consists of the first 20% of the respondents from the field; the second group consists of the next 20% of respondents, etc. It is seen that there is a clear relationship between turnout and ease of response. The early respondents have a much higher turnout than the later ones.

Table 5.5. Turnout and ease of response in Election surveys 1997- 2003

Percentage who voted in net sample, time for response and non-respondent in:	Net sample	Answered the survey among 1-20	21-40	41-60	61-80	81-100	Non-respondense
General election survey 1997	86	92	89	89	83	76	71
General election survey 2001	83	86	86	84	81	77	64
Local election survey 2003	67	75	70	67	64	61	47

This simple idea may be generalized in several ways using other indicators of ease of response. This method has several advantages: The effects of nonresponse on all target variables may be studied. Furthermore, the effects of nonresponse on measures of relationship, like regression coefficients, correlations etc, may be studied (Thomsen and Siring, 1983).

It is possible to use this information to adjust the population estimates in order to reduce the non-response bias, as suggested by Bartholomew (1961). Such adjustments have been made on an experimental basis for several surveys in Statistics Norway, but has until now never been used in actual statistics production. The main purpose of such a time indicator of the estimate is to serve as a process variable during the data collection. By following the development of the estimates of the target variables during the whole collection period, it is possible to set up a Process Efficiency Graph in order to determine when to stop data collection or concentrate further efforts within certain subpopulations. See chapter 3 and 4.

5.4. Weighting methods

5.4.1. Sample - and population based weights

To reduce the effects of nonresponse, weighting is often performed. Each observed unit is assigned a weight to adjust for varying response probability. As an example, in Table 5.1 the sample may be divided into three groups by level of education, denoted by $h = 1,2,3$. Then within each group a weight is assigned to each unit. A simple weight is the reciprocal of the group response rate. Clearly this does not change the observed subsample means within each education group, denoted by $\bar{y}_h$. Provided a self-weighting sample, the nonresponse adjusted estimate for the

population mean is given by $\hat{\bar{Y}} = \sum_h \frac{n_h}{n} \bar{y}_h$, where n_h is the size of each education group, and $n = \sum_h n_h$. Whereas the unadjusted estimate is given by $\tilde{\bar{Y}} = \sum_h \frac{n_{hr}}{n_r} \bar{y}_h$, where n_{hr} is the number of respondents within each education group, and $n_r = \sum_h n_{hr}$.

The weights constructed above from table 5.1 are sample based. Sometimes, the relative sizes of the educational groups are known in the whole population from a recent census or administrative registers. In such cases weights can be calculated based on the ratio of population to the observed sample, and are thus population based. The choice between population and sample based weights, is usually based on the availability of data. When possible, population weights are chosen because they often reduce bias as well as the variance of the estimate. In Statistics Norway we have access to a large number of administrative registers, which often makes it possible to choose population based weights. The population based weights are discussed in Section 5.4.2 Poststratification.

Example: The Norwegian Health Survey 1995. (continued)
In this survey we identified seven variables, which all were statistically significant for explaining the response probability. To adjust for the variation in response probability, data were weighted, using the reciprocal of the response probability, estimated by using the response model (5.1). See e.g. Ekholm and Laaksonen (1991). In order to investigate whether nonresponse has effects on the results, weighted and unweighted population estimates are shown for 5 variables from the survey in table 5.6. It is seen that the effects on the final results are small for these variables. These results are typical for most variables from the survey. Clearly this response model alters the estimates very little, which may indicate either that a nonignorable model is necessary to evaluate nonresponse bias or that in this case the nonresponse effect is negligible. Variances were not calculated, but it is to be expected that the variance increased as a result of the weighting.

Table 5.6. Effect of weighting for 5 variables

Variable	No weighting	Weighting
Sleeping problems	27.4	28.4
Smoker	28.4	29.1
Exercises at least 2 days per week	42.6	42.6
Feel lonesome	9.7	9.9
Serious problems during childhood	14.1	14.3

5.4.2 Poststratification.

To carry out poststratification, the sample is divided into groups with known population sizes, N_h. The poststratification estimator of the population mean is defined by

(5.2) $$\sum_h \frac{N_h}{N} \bar{y}_h \ ,$$

where $\bar{y}_h$, $h = 1, 2, ..., H$, is the sample mean in response poststratum h.

It follows that
(5.3) $$\sum_h N_h = N \ .$$
Some important properties of the poststratified estimator are:

- If no nonresponse is present, the poststratification estimator is unbiased. If the poststrata are chosen as homogeneous groups with respect to the target variables, considerable variance reduction may be gained.
- If nonresponse is present, poststratification is often used to reduce the nonresponse bias. Under the assumption that the response probability is constant within each poststratum, the nonresponse bias is reduced to zero by using the poststratified estimator.
- By using the poststratified estimator, the distribution of the marginals of the poststratifying variables is identical in the population and the weighted sample. This property is highly appreciated in official statistics, as it is possible to secure consistency in some distributions, for instance (sex x age groups) between surveys and a census and/or administrative registers.

For these reasons, poststratification is often used in Statistics Norway as a first step in the process to determine the final choice of weights. At this step, feasible auxiliary information is identified, and its impact on the number of poststrata and nonresponse bias is studied. If the number of poststrata is high, a more complicated weighting method, calibration, is used, presented in the next section. Before doing that, we shall present a method that can be used to evaluate the bias of the poststratified estimator.

5.4.2.1. Poststratification and nonignorable response model

As mentioned above, the nonresponse bias of the poststratified estimator vanishes when the response probability is constant within all poststrata. This assumption is hard to verify, and is often clearly wrong. In order to evaluate this assumption, an informative nonresponse model is sometimes introduced. In this model we assume that the response probability only depends on the target variable *y*, i.e., is constant over all poststrata given *y*. This simple nonignorable nonresponse mechanism is useful to assess the potential bias of the poststratified estimator.

Assume we have a simple random sample, and that the target variable *y* and the poststratifying variable *x* are

binary. Let b_{pst} denote the bias of the poststratified estimator of the population mean of y and let b_{srs} denote the bias of the unweighted sample mean. Zhang (1999) shows that under the simple nonignorable response model given above, the ratio $\gamma = b_{pst} / b_{srs}$ can be estimated from the responses alone. As $(\bar{y} - \bar{y}_{pst})$ is an estimate of $(b_{srs} - b_{pst})$, a bias correcting estimator is given by

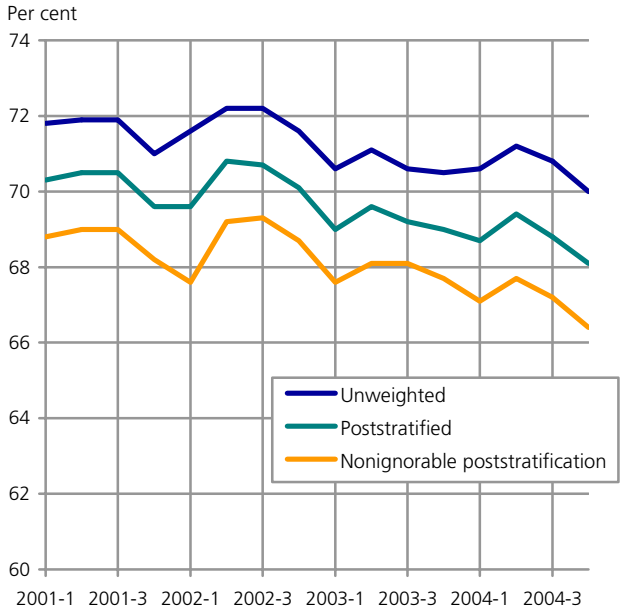
$$(5.4) \quad \bar{y}_{adj} = \bar{y}_{pst} - (\bar{y} - \bar{y}_{pst})\hat{\gamma} / (1 - \hat{\gamma}) = \bar{y}_{pst}(1 - \hat{\gamma})^{-1} - \bar{y}\,\hat{\gamma} / (1 - \hat{\gamma}),$$

Example: The Norwegian Labour Force Survey
To illustrate some of the properties of poststratification, we shall use the Labour Force Surveys (LFS) as example. An important variable in LFS is the employment rate. Poststratification has for a long time been used in connection with these surveys. Register based employment status is used as auxiliary variable, linking every person in the sample to the Employment Register. Due to the high correlation between the LFS employment status and the corresponding variable in the register, poststratification greatly reduces the sampling variance of the target variable, LFS employment rate, see Thomsen and Zhang (2001). In Figure 5.1 it is seen that the poststratified estimate of the number of employed persons is consistently lower than the unweighted estimate. This we believe is due to nonresponse. Based on register information, the response rate among employed persons is clearly higher than the response rate among non-employed persons. We therefore believe that the poststratified estimate has less nonresponse bias than the unadjusted estimate, and the poststratified estimates are used as the basis of official statistics. (It should be mentioned that the poststratified estimates are further adjusted by calibration before being published as official statistics. The effects on the population estimates and their variances of this adjustment is however very small. The purpose of calibration is to facilitate the publication of county estimates at a lower aggregation level.)

In Figure 5.1 it is seen that weighting affects the level of employment, but seems to have little effect on the development over time. This is comforting as the employment rate is an indicator used to monitor the economic development of the society.

As nonresponse induces a bias in the employment rate, it is important to evaluate to what extent poststratification reduces this bias. It seems clear that the sample of respondents is biased towards the employed part of the population. The direction of the adjustment towards a lower employment rate is therefore as expected. As mentioned earlier, if the response rate only depends on the register based employment status, the poststratified estimate completely removes the nonresponse bias. This assumption is hardly correct, and to evaluate the possible bias of the poststratified estimate we introduce a kind of "worst-case" response model. In this model we assume that the response rate depends on the actual employment status in the LFS and not on the register based status. Applying the approach of Section 5.4.2.1, estimates $\bar{y}_{adj}$ under this nonignorable reponse model are shown in Figure 5.1 (as nonignorable poststratification). It is seen that these estimates are consistently lower than the poststratified estimates through the whole period. These numbers are not published, but are interpreted as an indicator that poststratification removes between 50% and 100% of the nonresponse bias.

Figure 5.1. Employed persons estimated in the Labour force survey (2001-2004) using three different weighting methods. Percent



5.4.3 Calibration

A general framework for weighting is calibration (Deville and Särndal, 1992). The method may be described as follows:

- The calibrated sampling weights should be as close to the initial sample weights as possible. The initial weights are typically the inverse of the selection probabilities or weights from the poststratified estimator.
- The weighted sample distribution of the auxiliary variables should be equal to their distribution in the population.

A distance measure, D is established to measure the distance between the final calibrated weights and the initial sample weights. Minimizing D under the second condition now finds calibration weights.

Calibration has proven a very useful estimation tool. It is more flexible than poststratification for several reasons. Typically, calibration is carried out with respect to the marginal distribution of the auxiliary variables; whereas poststratification requires the simultaneous distribution of the same auxiliary variables. This means that calibration requires less extensive auxiliary information at the population level. In addition, where poststratification breaks down due to empty sample poststrata, calibration remains feasible provided there are no empty margins. See Zhang (2000). Calibration is also able to make combined use of auxiliary information associated with different units. An example is weighting based on both characteristics of households and persons (e.g., Heldal, 1992).

However, care is needed when designing the calibration. It is important that the statistician keeps control over the final weights, as these may have a large variation or even be negative. In such cases, it is important to adjust the weighting method by choosing other auxiliary variables or redefine them. During this process, poststratification has proven a very useful tool.

Lundstrøm and Särndal (1999), Särndal and Lundstrøm (2005) focused on the use of calibration in the presence of nonresponse. However, it is not immediately clear as to what extend the nonresponse bias is affected by calibration. For instance, it is not intuitive that the response probability can be expressed as a linear function of the auxiliary variables. Logistic regression model such as (5.1) is preferable. This is another reason for using poststratification as a first step in nonresponse adjustment when designing the calibration. For example, by using the informative response model suggested in 5.4.2.1, it is possible to get an idea of the nonresponse bias in the poststratification estimator. Comparison between the calibration and poststratification estimates could then indicate the impact of calibration on the nonresponse bias.

Example: The Norwegian Consumer Expenditure Survey 2001.
Household expenditure surveys are special surveys in many ways. In our context, the very high nonresponse rate is of special concern. In 2001 the nonresponse rate was 51%. Considerable efforts have therefore been made during several years in order to find good adjustment methods to reduce any nonresponse bias, see Belsby (1995, 2003). In this example, we compare two weighting methods.

The first weight is the reciprocal of the inclusion probability, which is equal to the product of the selection probability and the response probability. The selection probability is known, while the response probability is estimated as a function of the age of the eldest person in the household. Spline regression is used for this estimation. For details, see Belsby (2003).

The second weight is constructed by calibrating the first weight, using type of household as calibration variable. Information concerning type of household was available from the Population Census 2000.

To compare the performance of the two weighting methods, their effect on population estimates of the age - and income distribution is given in table 5.7 and 5.8.

Table 5.7. Age distribution estimated by using the reciprocal of the inclusion probability as weight and calibrated weights

Age	Population	No calibration	Calibration
00-10	13.8	14.3	14.9
10-19	13.2	13.4	13.8
20-29	13.6	12.2	12.4
30-39	16.0	17.0	16.5
40-49	14.6	15.8	15.8
50-59	13.3	12.3	12.7
60-69	8.2	8.8	8.5
70-79	7.3	6.0	5.4

Table 5.8. Income distributions estimated by using the reciprocal of the inclusion probability as weight and calibrated weights

Deciles	Population	No Calibration	Calibration
1-3	30.0	30.1	31.1
4	10.0	7.3	7.1
5	10.0	9.0	8.7
6	10.0	8.7	8.5
7	10.0	10.2	10.0
8	10.0	10.5	10.5
9	10.0	10.5	10.3
10	10.0	11.0	10.8

The results in table 5.7 and 5.8 show that for distribution of income and age, the two weighting methods give similar results. No variance calculations were made. It is, however to be expected that calibration will reduce the sampling variance on most estimates.

In future expenditure surveys (and most other household surveys), information from the population census 2000 cannot be used as calibration variable, primarily because it will be outdated. An alternative, which seems reasonable, is to use the newly established register of households and families.

5.4.4. Other weighting methods

Other weighting methods are also used. A method with similar properties as calibration, is raking or iterative proportional fitting. It was first used in the U.S. 1940 census and introduced by Deming and Stephan (1940). It is particularly useful in connection with categorical data. The method uses an iterative algorithm in order to find weights such that the population estimates of the marginal of some auxiliary variables are equal to the population marginal taken from some external source.

In a few cases, tailor-made, more complex models are used on an experimental basis. In Belsby, Bjørnstad and Zhang (2005) a completely model-based approach is used to estimate the distribution of households by size. This model takes into account that the nonresponse mechanism may be nonignorable. They also find that the poststratified estimate of the total of single-person households is about half way between the unadjusted estimate and the model-based estimate. The adjusting effect of poststratification here is similar to that of the employment rate from the LFS (Figure 5.1).

5.4.5. Some general comments

To day, weighting of sample surveys is very common in most household surveys in Statistics Norway. Usually poststratification or calibration is used, often in combination. There are two reasons for this:

- The sampling error is reduced in a large majority of surveys.
- The weighted sample distributions of some important auxiliary variables are equal from one survey to another and equal to the distribution in the population.

For a statistical agency, it is important to secure consistency between various surveys, and weighting is an efficient way to achieve this goal. (It is sometimes useful to perform weighting in two steps. In the first step the weighted estimates are made equal to the marginals in the selected sample and in the second step the weighted estimates are made equal to the population marginals. The reason is that the first weighting gives a good indicator of the effects on the nonresponse bias, while the second weighting is a variance reducing weighting.)

Concerning its impact on nonresponse bias, the situation is less clear. There are good reasons to believe that weighting has some impact on the bias of the target variables, especially when the weighting method is chosen after a careful analysis of the nonresponse bias and its relationship with various auxiliary variables. To quantify how much the total bias is reduced by weighting, it is necessary to choose a realistic, often nonignorable, nonresponse model. We believe that more work along the line of informative and non-informative poststratification is useful to throw more light on this question. See Zhang (2001) for a generalization of the binary case in Section 5.4.2.1.

5.5. Variance estimation in the presence of nonresponse

Direct or linearization-based variance estimation methods, that are readily available in the case of full response, require additional separate development in the presence of nonresponse. Resampling methods provide an attractive alternative approach. See Shao (1996) for an excellent overview. For example, Belsby, Bjørnstad and Zhang (2005) applied the bootstrap for household survey estimators under various nonresponse models. Other resampling methods include the jackknife and the balanced repeated replication (BRR). The resampling methods can be applied in situations with unit nonresponse that is handled by reweighting, as well as item nonresponse that is handled by imputation, or both. Each resampling method uses the same scheme for any estimator of interest, and we can estimate the variances of many estimators based on the same set of pseudo-replicate samples. Asymptotic consistency of the various resampling methods is established under suitable regularity conditions for the stratified multistage sampling design (e.g., Shao and Tu, 1995, Chapter 6).

A few remarks, partly based on our own experiences, can be noted. (a) The BRR is readily applicable only for the special, though important, case of two primary sampling units (PSUs) per stratum. The jackknife can be understood as a linearization of the bootstrap. The bootstrap is the most general approach, and is applicable both for smooth and non-smooth statistics. Under single-stage sampling or multistage sampling with large number of PSUs, the bootstrap requires in fact fewer replications than the jackknife. (b) A possible difficulty arises when the estimation uses random imputation and the stratum numbers of PSUs are small. Independent re-imputation for each pseudo-replicate sample can lead to overestimation of the variance. Saigo, Shao and Sitter (2001) proposed a modification for the bootstrap. (c) Another possible violation of the regularity conditions is that the first-stage sampling fraction is non-negligible, in which case the PSUs cannot be treated as if they were selected with replacement. A related issue is when the first-stage selection probabilities are very unequal. See Berger and Skinner (2005) for a discussion. (d) Finally, the nonresponse models studied in the literature of bootstrap for variance estimation (e.g. Shao and Sitter, 1996) are typically of the so-called missing-at-random type (Rubin, 1976). Although the bootstrap can readily be carried through for nonignorable nonresponse, general theoretical results or comprehensive empirical evaluation studies are lacking at the moment. For an application of bootstrap under nonignorable nonresponse, see Zhang (2001).

Appendix 1

Interviewer-administered surveys in Statistics Norvay

Survey	Mixed mode	Main mode Second mode	Multimode	Length of interview Minutes	Sample size	Panel	Frequency	Planned field period in days	Season	Other
Level of living	Yes	Face to face telephone	Drop off	30	5 000	No	Every year	129	Okt- feb	
Survey on income and living conditions	Yes	Telephone face to face	No	21	8 500	Rotating, 8 times	Every year	98	March-may	
Survey of Consumer expenditure	Yes	Face to face telephone	1. Interview, 2. diary 3. interview	1. int: 10 2. int 45	2 200	20% of the prev. sample	Consecutive	*Fieldwork period is 1 year. Reference period is 14 days*		Household is unit
Labour Force Survey	Yes	Telephone face to face	No	5	24 000	Rotating 8 times	Consecutive	10	Quarterly	Mandatory. Proxy allowed
Rental survey	No	Telephone	No	5	1 600	For a year 12 times	Consecutive	14	Every month	Dwelling is unit
Rent marked survey	No	Telephone	No	5	21 000	No	Every year	62	Mai -july	Dwelling is unit
Travel and holiday	Yes	Telephone and face to face	No	20	8 000	No	Quarterly	32	Jan, April, July, Okt	
Survey on Media use	No	Telephone	No	20	2 600	No	Quarterly	28	Mar, June, Sept, Des	Only telephone listed people
Election survey	Yes	Face to face and telephone	No	60	3 000	Rotating 2 times	Every 4. year	64	Sept- nov	
European Social survey	No	Face to face	Drop off	70	2 750	No	Every 2. year	70	Sept- nov	

References

Bartholomew, D.J. (1961). A Method of Allowing for "Not-at-Home" Bias in Sample Surveys. Applied Statistics, **10,** pp. 52 – 59

Belsby, L. (1995). Forbruksundersøkelsen. Vekt-metoder, Frafallskorrigering og Intervjuer-Effekt. Notater 95/18 Statististisk sentralbyrå-Statistics Norway

Belsby, L. (2003). Vekter i Forbruksundersøkelsen. Notater 2003/58. Statistisk sentralbyrå-Statistics Norway

Belsby, L., Bjørnstad, J.F. and Zhang, L.-C. (2005). Modeling and Estimation Methods for Household Size in the Presence of Nonignorable Nonresponse. Applied to the Norwegian Consumer Expenditure Survey. Survey Methodology, **31,** pp. 197 - 212.

Belsby, L. and Vedø, A. (1998). Frafallsanalyse av Helseundersøkelsen 1995. Notater 98/3, Statistisk sentralbyrå- Statistics Norway

Berger, Y.G. and Skinner, C. (2005). A Jackknife Variance Estimator for Unequal Probability Sampling. Journal of the Royal Statistical Society, Series B, **67**, pp. 79 - 89.

Cialdini, R.B. (1990),.Deriving Psychological Concepts Relevant to Survey Participation from the Litteratures on Compliance, Helping, and Persuasion. Paper presented at the Workshop on Houshold Survey Nonresponse at Statistics Sweden.

Cobben, F., Groves, B. and Schouten, B. (2006). R-indexes: Empirical Studies. Statistics Netherlands.

Cochran, W. (1977). Sampling Techniques. New York. Wiley.

De Heer, W. (1999). International Response Trends, Results of an International Survey, Journal of Official Statistics, **15**, pp. 129-142.

Deming, W.E. (1991). Out of the Crisis. Cambridge University Press.

Deming, W.E. and Stephan,F.F. (1940). On a least squares adjustment of a sample frequency table when expected marginal totals are known. Annals of Mathematical Statistics, **11,** pp. 427-444.

De Leeuw, E. and De Heer, W. (2002). Trends in Household Survey Nonresponse: A longitudinal and International Comparison. in Groves, R, Dillmann, D.A.,Eltinge, J.L., and Little, R. J. A. eds. Survey Nonresponse. New York, Wiley.

De Leeuw, Edith, Joop Hox, Elly Korendijk, Gerty Lensvelt-Mulders (2005). The Influence of Advance Letters on Response in Telephone Surveys: A Meta-Analysis. Paper presented at the 16th International Workshoop on Houshold Survey Nonresponse, SCB, Dalarne, Sweeden.

Deville, J.C. and Särndal, C.E. (1992). Calibration Estimators in Survey Sampling. Journal of the American Statistical Association, **82**, pp. 376-382.

Dillman, D. (1978). Mail and Telephone Surveys: The Total Design Method, New York: Wiley

Ekholm, A. and Laaksonen, S. (1991). Weighting via Response Modelling in the Finnish Household Budget Survey. Journal of Official Statistics, **7**, pp. 325-337.

Eurostat (2002). Quality in the European Statistical System- The Way Forward. Office for Official Publications of the European Communities: Luxembourg.

Eurostat (2005). Handbook on Improving Quality by Analysis of Process Variables.

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

Groves, R.M., Caildini, R.B., and Couper, M.P. (1992). Understanding the Decision to Participate in a Survey. Public Opinion Quartely **56(**4), pp. 475-495.

Groves, R.M. and Couper, M.P. (1998). Nonresponse In Houshold Interviewer Surveys, New York: Wiley.

Groves, R.M., Dillmann, D.A., Eltinge J. L. and Little, R.J. A. (2002). Survey Nonresponse. New York, Wiley.

Groves, R.M. (2005). Research Synthesis: Nonresponse Rates and Nonresponse Bias in Household Surveys. Paper presented at the 16. International Workshop on Household Surveys, Tallberg, 28-31 August 2005.

Harris-Kojetin and Tucker (1999). Exploring the Relations of Economic and Political Conditions with Refusal Rates to a Government Survey. Journal of Official Statistics, **15**, pp. 167-184.

Heldal, J. (1992). A Method for Calibration of Weights in Sample Surveys. Working Paper 3/1992, Statistics Norway.

Holt, D. and Smith, T.M.F. (1979). Post Stratification. Journal of the Royal Statistical Society, Series A, **142**, pp. 33-46.

Kish, L. (1965). Survey Sampling. New York: Wiley

Kish, L. (1995). The Hundred Years' War of Survey Sampling. Statistics in Transition, **2**, pp. 813-830.

Lundström, S. and Särndal, C.-E. (1999). Calibration as a Standard Method for Treatment of Nonresponse. Journal of Official Statistics, **15**, pp. 305-327.

Luppes, M. (1995). A Content Analysis of Advance Letters from Expenditure Surveys of Seven Countries. Journal of Official Statistics, **11**, pp. 461-480.

Madow, W.G., Nisselson, H. and Olkin,I. (eds) (1983). Incomplete Data in Sample Surveys. Vol **3**. New York Academic Press.

Mevik, A. K. (2004). Uncertainty in the Norwegian Business Tendency Survey. Documents 2004/10, Statistics Norway.

Rubin, D. (1976). Inference and Missing Data. Biometrika, **63**, pp. 581 – 592.

Saigo, H., Shao, J. and Sitter, R. (2001). A Repeated Half-Sample Bootstrap and Balanced Repeated Replications for Randomly Imputed Data. Survey Methodology, **27**, pp. 189 – 196.

Särndal, C.-E. and Lundström, S. (2005). Estimation in Surveys with Nonresponse. New York: Wiley.

Shao, J. (1996). Resampling Methods in Sample Surveys (with Discussions). Statistics, **27**, pp. 203-254.

Shao, J. and Tu, D. (1995). The Jackknife and the Bootstrap. New Tork: Springer-Verlag.

Shao, J. and Sitter, R. (1996). Bootstrap for Imputed Survey Data. Journal of the American Statistical Association, **91**, pp. 1278 – 1288.

Singer, E. (1999). The effects of Incentives on Response rates in Interviewer-Mediated Surveys. Journal of Official Statistics, **15**, pp. 217-230.

Singer, E (2002). The use of Incentives to Reduce Nonresponse in Houshold Surveys, in Groves, R., Dillman, D., Eltinge, J., and Little, R. (eds), Survey Nonresponse, pp. 275-288. New York: Wiley.

Smith, T.W. (2002). Developing Nonresponse Standards, in Groves,R., Dillmann, D., Eltinge,J., and Little, R. (eds), Survey Nonresponse, pp. 27-40. New York, Wiley.

Solheim, L., Håland, I. and Lagerstrøm, B.O. (2001). Proxy Interview and Measurement Error in the Norwegian Labour Force Survey. Paper presented at the 12. International Workshop on Household Survey Nonresponse, Oslo 12-14 September 2001.

Stoop, I.A.L. (2005). The Hunt for the Last Respondent. Social and Cultural Planning Office, The Hague.

Sæbø, H.V., Byfuglien,J. and Johannessen R. (2003). Quality Issues at Statistics Norway. Journal of Official Statistics, **19**, pp. 287-303.

Thomsen, I. and Siring, E. (1983). On the Causes and Effects of Nonresponse: Norwegian Experiences, in W. G. Madow and I Olkin (eds), Incomplete Data in Sample Surveys, Vol. **3**. New York; Academic Press.

Thomsen, I. and Zhang, L.-C. (2001). The Effect of Using Administrative Registers in Economic Short Term Statistics: The Norwegian Labour Force Survey as a Case Study. Journal of Official Statistics, **17**, pp. 285-294.

Wang, J. H. (2004). Non-response in the Norwegian Business Tendency Survey. Documents 2004/3, Statistics Norway.

Zhang, L.-C. (1999). A Note on Poststratification when Analysing Binary Survey Data subject to Nonresponse. Journal of Official Statistics, **15**, pp. 329-334.

Zhang, L.-C. (2000). Post-stratification and Calibration – A Synthesis. The American Statistician, **54**, pp. 178-184.

Zhang, L.-C. (2001). A Method of Weighting Adjustment for Survey Data Subject to Nonignorable Nonresponse. Discussion Papers 311, Statistics Norway.

Tidligere utgitt på emneområdet

Previously issued on the subject

Notater

1999/82: Øyvin Kleven, Elisabeth Dalheim og Dag Roll-Hansen: Innvanderes utdanning - en pilotundersøkelse.

2003/81: Jan Henrik Wang : Frafall i konjunkturbarometeret

Discussion Papers (DP)

311. Li-Chun Zhang: A method of weighting adjustment for survey data subject to nonignorable nonresponse

Artikler

121: Ib Thomsen and Erling Siring: On the causes and effects of non-response : Norwegian experiences

De sist utgitte publikasjonene i serien Rapporter

Recent publications in the series Reports

2006/6 P.R. Johansen og J-A. Jørgensen: Virkningsberegninger på KVARTS. 2006. 60s. 180 kr inkl. mva. ISBN 82-537-6937-7

2006/7 F. Foyn og K.A. Kjesbu: Brukernes vurdering av SkatteFUNN-ordningen. Delrapport i evalueringen av SkatteFUNN-ordningen. 2006. 62s. 180 kr inkl. mva. ISBN 82-537-6938-5

2006/8 A. Langørgen, T.A. Galloway og R. Aaberge: Gruppering av kommuner etter folkemengde og økonomiske rammebetingelser 2003. 2006. 35s. 155 kr inkl. mva. ISBN 82-537-6953-9

2006/9 Y. Lohne og H. Nome Næsheim: Eldre i arbeidslivet. 22s. 2006. 155 kr inkl. mva. ISBN 82-537-6955-5

2006/10 A. Langørgen og R. Aaberge: Inntektselastisiteter for kommunale tjenester. 2006. 19s. 155 kr inkl. mva. ISBN 82-537-6960-1

2006/11 G. Berge og J.K. Undelstvedt: Kommunal avløpssektor: Gebyrer 2005 - Utslipp, rensing og slamdisponering 2004. 2006. 66s. 155 kr inkl. mva. ISBN 82-537-6964-4

2006/12 T. Hægeland, K.A. Kjesbu og J. Møen: Fører SkatteFUNN-ordningen til økt FoU-innsats? Foreløpig rapport om innsatsaddisjonalitet. 2006. 27s. 155 kr inkl. mva. ISBN 82-537-6967-9

2006/13 G. Dahl, A. Walstad Elnes. T. Jørgensen og C. Trewin: Langtidsmottakere av økonomisk sosialhjelp. 2006. 41s. 155 kr inkl. mva. ISBN 82.537-6971-7

2006/14 N.K. Buskoven: Undersøkelse om bibliotekbruk. 2006. 71s. 180 kr inkl. mva. ISBN 82-537-6974-1

2006/15 A. Barstad, E. Havnen, T. Skarðhamar og K. Sørlie: Levekår og flyttemønstre i Oslo indre øst. 2006. 260s. ISBN 82-537-6977-6

2006/16 T. Hægeland, L.J. Kirkebøen og O. Raaum: Resultatforskjeller mellom videregående skoler. En analyse basert på karakterdata fra skoleåret 2003-2004. 2006. 64s. 180 kr inkl. mva. ISBN 82-537-69779-2

2006/17 Å.Cappelen, E. Fjærli, F. Foyn, T. Hægeland, K.A. Kjesbu, J. Møen og A. Raknerud: SkatteFUNN-evalueringen - Årsrapport 2005. 2006. 46s. 180 kr inkl. mva. ISBN 82-537-6990-3

2006/18 J. Ramm: Inntektsforholdene til grupper med nedsatt funksjonsevne - inntektsregnskap for personer. 2006. 31s. 155 kr inkl. mva. ISBN 82-537-6994-6

2006/19 L. Allertsen og T. Kalve. Innvandrerbarn i barnevernet 2004. 2006. 50s. 180 kr inkl. mva. ISBN 82-537-6996-2

2006/20 T.A. Galloway og M. Mogstad: Årlig og kronisk fattigdom i Norge. En empirisk analyse av perioden 1993-2001. 2006. 37s. 155 kr inkl. mva. ISBN 82-537-7002-2

2006/21 B. Langset, E. Holmøy og L. Lerskau: Et grånende Norge: Betydningen av økt behov for eldreomsorg for makroøkonomi og offentlige finanser mot 2050. 2006. 24s. 155 kr inkl. mva. ISBN 82-537-7018-9

2006/22 E. Ugreninov og O. F. Vaage: Studenters levekår 2005. 2006. 86s. 180 kr inkl. mva. ISBN 82-537-7020-0

2006/23 V. Aalandslid: Valgdeltakelsen blant norske statsborgere med ikke-vestlig innvandrerbakgrunn ved Stortingsvalget 2005. 2006. 37s. 155 kr inkl. mva. ISBN 82-537-7022-7

2006/24 G. Dahl og S. Lien: Uførepensjon og sosialhjelp/introduksjonsstønad blant innvandrere. 2006. 45s. 155 kr inkl.mva. ISBN 82-537-7027-8

2006/25 S. Lien og E. Nørgaard: Rapportering fra krisesentrene 2005. 2006. 37s. 155 kr inkl. mva. ISBN 82-537-7029-4

2006/26 G. Daugstad: Omfang av bruk av kontantstøtte blant barn med ikke-vestlig innvandrerbakgrunn. 2006. 66s. 180 kr inkl. mva. ISBN 82-537-7033-2

2006/27 J.I. Hamre: Evaluering av ordinære arbeidsmarkedstiltak påbegynt 4. kvartal 2004. Dokumentasjon og analyse av effekter november 2005. 2006. 76s. 180 kr inkl. mva. ISBN 82-537-7044-8