

Statistics Norway  
Department of Social Statistics

*Arne Andersen, Tor Morten Normann  
and Elisabeth Ugreninov*

**EU-SILC: Pilot Survey**  
Quality Report from Statistics Norway

**Contents**

- 1 Background and purpose ..... 3**
- 2 Sample ..... 4**
- 3 Data collection ..... 4**
- 4 Non-response ..... 6**
- 5 Timing of the interviews..... 7**
- 6 Interview method and data entry ..... 8**
- 7 Reports from the interviewers on recruitment of respondents ..... 9**
- 8 Field procedures: household questionnaire and proxy interview ..... 9**
  - 8.1 Proxy interview ..... 10
  - 8.2 Household questionnaire..... 10
- 9 Evaluation of the questionnaire ..... 11**
  - 9.1 Household questionnaire..... 11
  - 9.2 Personal questionnaire ..... 14
- 10 construction of income target variables ..... 15**
  - 10.1 Data sources ..... 16
  - 10.2 Income concepts..... 17
  - 10.3 Data quality ..... 21
- 11 Integration with existing surveys..... 22**
- 12 Use of register information on employment ..... 22**
- 13 Conclusion..... 23**
- Appendix..... 25**
- Recent publications in the series documents..... 30**



## 1 BACKGROUND AND PURPOSE

Since 1997 Statistics Norway has been conducting an annual longitudinal survey on living conditions in Norway. Together with the annual theme rotating cross sectional survey, these constitute the coordinated surveys on living conditions in Statistics Norway. It is now the intention to change this longitudinal survey from the year 2003. From then on, it is supposed to be a part of a EU/Eurostat led cooperation called EU-SILC (Statistics on Income and Living Conditions). This both implies a considerable change of the questionnaire, and hence the preparation of the output data, and an enlargement of the sample.

The EU-SILC pilot survey, in Norway called "Levekår i Europa" (Living conditions in Europe), was conducted to prepare the change of the panel survey. The main purposes of the pilot have been testing several aspects of the questionnaire, the fieldwork, and establishing routines concerning the preparation and delivering of data. Testing of response rates was subordinate, but will nevertheless be mentioned in the report. The final aim is of course to establish a foundation for the new panel survey from 2003 on.

**Table 1: Key figures for the survey**

	<b>Number</b>	<b>Percent</b>
<b>Sample</b> (persons selected for interview)	400	100,0
<b>Non-eligible</b> (dead, persons living abroad)	4	
<b>Gross sample</b>	396	100,0
<b>Non-response</b>	158	39,9
<b>Net sample</b> (persons interviewed)	238	60,1
<hr/>		
<b>Total number of respondents</b> (including household members)	478	
<b>Number of persons interviewed by proxy</b>	129	
<b>Method of data collection:</b>	Personal interview, telephone	
<b>Duration of interview:</b>	Approximately 22 min	
<b>Period of data collection:</b>	June 5 - 27, 2002	

## 2 SAMPLE

An important aspect in this pilot was to achieve a pre specified number of interviews, that is a net sample of no less than 200 respondents. In order to have a sufficient number of respondents in our gross sample, we draw a sample of 400 respondents. We normally expect a significantly higher response rate than 50 percent. The main argument for selecting as many as 400 in this survey, when no more than 200 interviews were needed, was that we wanted to make the period of data collection short. By putting a minimum of effort into getting in contact with the respondents we felt that we could utilize our interviewer resources to the maximum extent by selecting a large sample. This aspect must be taken into consideration when commenting on non-response (ch. 4).

Originally, there were no given upper age limit for respondents in this survey, but by a mistake the sample was drawn from BEBAS<sup>1</sup> among persons aged 16 to 79 years old by Dec 31, 2001. This is due to the fact that some of the routines from the sampling for the existing longitudinal survey were followed, and will of course be corrected in the main survey.

The sample was drawn from the entire country, without using the sample frame of Statistics Norway. There is therefore no design effect in this survey.

## 3 DATA COLLECTION

This survey was conducted by telephone interviewing. If necessary, face-to-face interviews were allowed. Nonetheless, none of the respondents were interviewed face-to-face. The interviews were carried out in the period between June 5 and 27, 2002. On June 27, we stopped all further interviewing since the required number of interviews was achieved.

From figure 1 we can observe how the response rate developed during the period of data collection. Once again, one should notice that the response rate is low in this survey since the interviewing was interrupted when the required numbers of interviews were achieved. This is also the reason for the falling response rate towards the end of the period, after being quite high earlier in the period. Normally, more resources are used in contacting all respondents in the gross sample.

---

<sup>1</sup> BEBAS the central database for demography/population fo Statistics Norway. The database is updated several times a month with information from the Norwegian population register.

**Figure 1: The development of response rates**

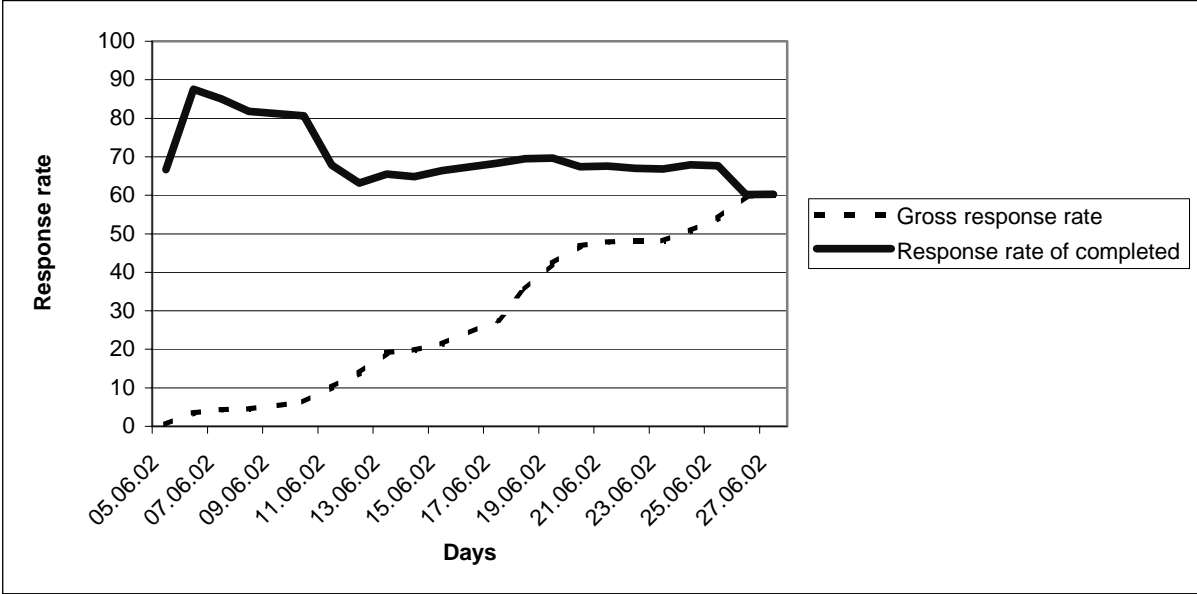
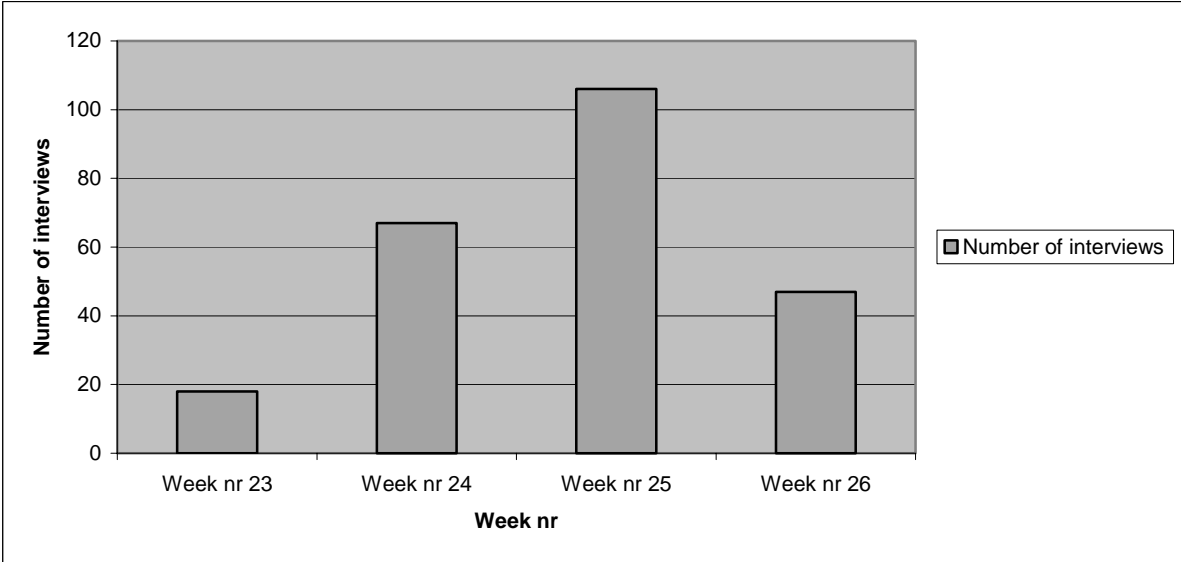


Figure 2 shows the number of interviews conducted each week of the data collection period. As expected, the two full weeks of interviewing shows the highest number of conducted interviews. The first and the last week of interviewing were not full weeks, hence a smaller number of interviews.

**Figure 2: Number of interviews per week**



## 4 NON-RESPONSE

The main reason for looking at non-response in this pilot is to see whether experiences can be used in the forthcoming main EU-SILC survey. Non-response may lead to bias if the distribution of a specific characteristic among the actual respondents (net sample) differs from the distribution among those selected for interview (gross sample). A common assumption is that high non-response increases the risk of biases. We therefore put effort into minimizing non-response.

In this case, there are a couple of aspects that makes the transfer of experiences concerning non-response from the pilot to the main survey difficult. First, the period of data collection in the pilot was in June, while the main survey will be conducted in a period from February to May. By experience, we assume that early spring is better for interviewing than summer. Second, as already pointed out, we stopped interviewing when a sufficient number of interviews were reached. Not all of the respondents in the gross sample were contacted within the data collection period. They were all automatically given status as non-respondents. Therefore, both the number of non-respondents and the distribution of different reasons for non-response differ from what we may have expected with a normal period of data collection.

Nevertheless, we can never expect that all of those selected in a gross sample for a voluntarily survey actually participates in the survey. The respondent may refuse to participate, or for some reason be prevented from participating or we may not be able to get in contact with the respondent during the period of data collection. All of this constitutes non-response in the survey. In this survey 158 persons were non-respondents. This is 39,9 percent of the gross sample.

Table 2 shows how the gross sample is distributed (per cent) by interview and different reasons for non-response by gender, age and region. The most common reason for non-response is "no contact". We cannot put much weight on this due to the interruption of the interviewing at 200 completed interviews. A total number of 70 refused to take part in the survey. These constitute 44,3 percent of the non-respondents and 17,7 percent of the gross sample. In this pilot, this is the most interesting part of the non-response. We may make comparisons with the theme rotating survey on living conditions, in which the refusals constitutes 19,2 and 19,9 percent in 2000 and 2001 respectively.<sup>2</sup> Hence the refusal rate in the pilot is a bit lower. If the data collection had been conducted in a normal manner, we would expect that some of those now classified as "no contact" would be refusals, thus increasing the refusal rate in the pilot. We must also take into consideration that the respondents in the existing longitudinal survey on living condition will be included in the main EU-SILC survey (appr 3600 respondents). This will affect response rates since both the refusal rate and the no contact is low in this group.

We should also notice that the non-response rate is higher among men than among women, and that an important reason is the high refusal rate among men. This is something new compared to the theme rotating survey on living conditions where there is no difference between genders. Again, caution is required when interpreting this finding, but we should

---

<sup>2</sup> Since the pilot must be seen as a cross sectional survey, and not a panel survey, it falls more natural to compare it to the theme rotating survey on living conditions, which also is a cross sectional survey, than the panel survey.

evaluate the presentation of the survey to see whether it may appeal more to women than to men.

Finally, non-response caused by "prevented" and "other reasons" constitutes a total of 1,8 percent of the gross sample.

**Table 2. Gross sample distributed by interview and different reasons for non-response by gender, age and region. Percent.**

	Total	Interview	Refusals	Prevented	No contact	Other reasons for non-response	N
<b>Total</b>	100,0	60,1	17,7	1,5	20,5	0,3	396
<b>Gender</b>							
Males	100,0	54,7	21,4	0,5	23,4	-	201
Females	100,0	65,6	13,8	2,6	17,4	0,5	195
<b>Age</b>							
16-24 years	100,0	69,2	11,5	-	19,2	-	52
25-44 years	100,0	56,4	14,1	-	28,9	0,7	149
45-66 years	100,0	62,3	20,3	0,7	16,7	-	138
67-79 years	100,0	56,1	26,3	8,8	8,8	-	57
<b>Region</b>							
Akershus and Oslo	100,0	59,5	13,5	1,4	24,3	1,4	74
Hedmark and Oppland	100,0	67,7	9,7	3,2	19,4	-	31
Østlandet (other)	100,0	50,6	26,0	2,6	20,8	-	77
Agder and Rogaland	100,0	66,1	16,1	1,8	16,1	-	56
Vestlandet	100,0	57,0	20,3	-	22,8	-	79
Trøndelag	100,0	68,6	11,4	2,9	17,1	-	35
Nord-Norge	100,0	63,6	18,2	-	18,2	-	44

## 5 TIMING OF THE INTERVIEWS

One important aspect of this pilot was to estimate the length of the interview. We were interested in the total interview time as well as the length of different parts of the interview. In advance, the interview was estimated to take 30 minutes. On basis of the 238 interviews completed (including members of the respondent's households), we have to conclude that it took less time than expected. Using a trimmed average, removing the lengthiest 5 percent and the shortest 5 percent of the interviews, we find an average of approximately 22 minutes. If the same questionnaire and design is used in the main survey, this could be an estimate. Two things should be taken into account, though. By a mistake, we excluded persons above the age of 80 from our pilot. Including these in the main survey will probably affect the timing. Decisions taken on proxy interviewing will also affect the timing.

We also wanted to find out more about the time used on different parts of the interview. We were especially interested in parts of the interview assumed to be difficult and time consuming. The most interesting part was perhaps the part concerning the household's total housing costs, which we assumed could be both difficult and time consuming. In the questionnaire we programmed a time test running from Likn1 to Avg3b. In this part we ask



about loan, rent and other housing costs. A trimmed average indicates that this part took approximately 2 minutes. Which questions are asked in this part is related to whether the household owns or rents the dwelling, and whether or not the household has a housing loan. In our sample, 80 percent of the household own their dwelling. About 60 percent of the owners have loans. Most questions are asked to those who own their dwelling, have loans and pay rent. Looking at the reports from the interviewers, we see that many report on this part of the questionnaire, particularly the questions about insurance and service costs. These two questions are asked only to owners with no shared expenses. These are quite a large part of our sample, and there is a potential for improvement in both data quality and time consumption.

Another interesting point concerning timing is that the questions about health seemed to work smoothly. The trimmed average for this part is less than one minute. In this part of the questionnaire, the number of questions asked is also dependent on the answers given. Respondents with good health are asked fewer questions than those with health problems. In our net sample 53 persons (22 percent) were asked all questions in this part. The estimated length of this part is dependent upon whether or not we expect this share to be stable in the main survey.

Finally, there is timing of questions concerning labour. The average for all respondents, both selected respondent and household members is approximately 3 minutes and 40 seconds. If we look only at the selected respondent the average is approximately 4 minutes and 40 seconds. The average for household members who answered the question themselves is approximately 3 minutes and 15 second, while for those interviewed by proxy the average is one minute less.

## **6 INTERVIEW METHOD AND DATA ENTRY**

Statistics Norway's local interviewers conducted this survey. The method for data collection was Computer Assisted Telephone Interviewing (CATI). The software used by Statistics Norway is Blaise. When the questionnaire is programmed, the routing is automatically programmed in order to ensure that all respondents get the right questions. We also programmed some checks, both to control consistency between answers given, but also to control extremes. These checks are either in form of a warning, which can be suppressed by the interviewer, or in form of an absolute control, which requires that one or more answers have to be corrected.

During the interview, the Blaise programme was programmed to recode answers given into the required target variables. In cases where this could not be done, the target variables were established during the file handling. One problem to point out here is that further specification of how to handle item non-response in cases where one item is one small part of a target variable is needed.

## **7 REPORTS FROM THE INTERVIEWERS ON RECRUITMENT OF RESPONDENTS**

Since this was a pilot survey where the aim was get experiences, which can be used in the coming main survey in 2003, it was important to get information from the interviewers about different aspects of the survey. We therefore included some questions to the interviewer at the end of every interview. These questions concerned recruiting respondents to participate in the survey, whether certain questions were difficult to answer for the respondent or difficult to understand, and finally whether proxy interviewing seemed to create any troubles.

The first question to the interviewer concerned refusals. We asked the interviewer if they had the impression that the respondent had decided to refuse in advance, or if they had the impression that the reason behind the refusal was more circumstantial. Although the answers given by the interviewer must be subjective, we may get an impression of how theme and advance information as opposed to circumstance affects the respondents' willingness to participate.

Our material is limited; we only have information concerning 62 of the refusals. The interviewers had the impression that 49 of these had made up their mind about refusing in advance of the contact. This is 79 percent of the refusals. The share among men is slightly higher than among women, 81 and 76 percent respectively. This may lead to two assumptions. First, we can improve the design of the advance information in order to motivate respondents to participate. Second, interviewers are professional in their contact with respondents and thus minimize circumstantial refusals. But there is an alternative interpretation, and that is that interviewers here have put little effort into persuasion, and accepted refusals easily, thus misinterpreted it to be an advance decision by the respondent. Nevertheless, we must keep focus on improving our advance information.

The second question to the interviewers was concerning those respondents who were interviewed. The question was whether the respondent had to be persuaded or not to participate. Of the 235 respondents we have information about, only 42 had to be persuaded to participate, the rest had decided to participate in advance. Again, there is a higher share of men than of women that has to be persuaded, but the numbers are too small to base any conclusion upon. This, and the previous point, may indicate that the advance information led the respondents to make a decision about participating or not in advance of the contact. But, as mentioned above, it may also indicate that interviewers did not put much effort into persuasion.

## **8 FIELD PROCEDURES: HOUSEHOLD QUESTIONNAIRE AND PROXY INTERVIEW**

An important aim of the pilot was to test the field procedures, especially procedures for interviewing other adult household members than the selected person. How many proxy interviews did the interviewers make? How do proxy interviews affect the quality of the responses? We were also interested in the degree to which the interviewers follow the

instructions on which person in the household shall answer the household questionnaire. Also this may have implications for quality.

**8.1 Proxy interview**

One of the questions to the interviewer was about how the interviewer felt about conducting a proxy interview. In this survey, we require data about employment for all household members above the age of 15. To avoid unnecessary non-response on this part of the survey, we allowed for proxy interviewing, where another member of the household could answer the questions. In each case of proxy interview, we asked the interviewers about if they felt that the person answering the questions could answer them easily or not. We received 99 responses to this question, of which only eight said that the respondent had difficulties answering the questions. Again, these are small numbers and we rely upon the interviewers subjective evaluation of the situation, but this nevertheless indicates that proxy interviewing does not pose a serious threat to data quality. The only way to really evaluate this is to control the data by re-interviews. Regrettably, this was not possible within the pilot. Another aspect to be mentioned is that allowing for proxy interviewing might save resources by reducing the use of time by the interviewer to administer the field work. This is most evident in cases where only a few questions are asked the household members.

We have direct interview with all the 238 selected respondents. Out of 255 adult household members we have information from direct or proxy interview from 240 persons. Of these 47 per cent were interviewed by direct interview.

Table 3 shows the percentage interviewed by direct or proxy interview by relationship to the selected person. Direct interview is most common for spouse or cohabitant. It is least common when the non-selected person is a parent, parent-in-law or grandparent. Usually this is the situation when the selected respondent is a young member of the household.

**Table 3. Type of interview, by relationship to the selected respondent. Per cent.**

Relationship to the selected respondent	Direct interview	Proxy interview	Total	N
Spouse or co-habitant	56	44	100	151
Parent, parent-in-law or grandparent	24	76	100	51
Child, stepchild or brother/sister or other	39	61	100	38
				240

**8.2 Household questionnaire**

It is important for the quality of the information in the household questionnaire that the person who has the best knowledge about the themes asked answer the household questionnaire. It has been decided that the owner/renter of the dwelling should answer these questions.

In 200 households or 84 per cent of the households the person who answered the household questionnaire was owner or renter. In 187 households or approximately 80 per cent of all households the selected respondent is (co)owner/renter.

Table 4 shows which person in the household answered the household questionnaire. The most problematic are the 24 selected respondents (10 per cent) who are not owners/renters. The vast majority of these are young members of the household, and they will probably have problems answering all the household questions. We don't know if an interview with the selected respondent was the only possibility (the owner/renter was temporarily absent) or if the interviewer just didn't follow the instructions.

**Table 4. Persons in the household who answered the household questionnaire.  
Number of respondents**

Family relations to selected respondent	Selected respondent owns/rents the residence alone	Selected respondent owns/rents the residence along with others	Other than selected respondent owns the residence	N
Selected respondent	104	83	24	211
Spouse or co-habitant	3	8	2	13
Parent, parent-in-law or grandparent	-	-	11	11
Child, brother/ sister or other	1	1	1	3
<b>N</b>	108	92	38	238

## 9 EVALUATION OF THE QUESTIONNAIRE

### 9.1 Household questionnaire

We have evaluated the questions mainly through comments from the interviewers and through the partial non-response.

First some comments from the interviewers. We asked the interviewers to comment on questions, which were time-consuming and/or difficult to answer for some reason or other. We focus here on those questions, which were mentioned most frequently by the interviewers. Let us first point out that comment saying either that one question was time consuming or difficult may actually be overlapping. One particular question may be time consuming because it is difficult.

The first sequence causing some problems was the one about loans for housing and interests paid on these loans. (LUtg - LRent in the Norwegian questionnaire). We have eight reports saying that respondents used a long time to answer both single questions and/or the entire sequence. The problems are related to questions on interest and repayment in total, and especially on interests. This is a problem that is difficult to solve. As long as we require this information we have to ask the questions. A possible improvement could be to prepare the respondent for the question by mentioning it in the advance information, and thus giving the respondents a possibility to prepare for the question.

The next two questions creating problems are the ones about insurance on dwelling and municipal service charges (Avg1 and Avg2). Based on report from the interviewers, these questions were the most difficult ones. We received fourteen reports on the question about insurance costs, mainly because the respondents had to use long time to answer the question. On the municipal service charges question, we received sixteen reports, also mainly because the question was time consuming. A possible explanation to this may be that these are costs paid once a year, and that most people may have difficulties remembering the exact cost if the bills were paid a while ago. Insurance is often sold as part of a package, and to single out one specific part of the insurance may cause problems. These troubles are also known from previous surveys, but we have yet not been able to find any satisfying solutions to the problem. Again, a possible way out could be to prepare the respondent by mentioning the question in the advance information, thus giving the respondent the possibility to prepare by looking up bills or receipts necessary to answer the question.

We have also received some reports concerning the questions about maintenance of the dwelling and expenses for such maintenance (Avg3a and Avg3b). These questions may be time consuming. The same goes for the question about to what degree the total housing cost is perceived as a financial burden (Tyng).

There were three reports on the question which asked the respondent to estimate the market value of the dwelling (irrespective of whether one owns or rents the dwelling). We assumed that this question could be problematic, since this is a question requiring a spontaneous subjective evaluation of a complex issue. Thus, only three reports on this question does not seem disturbing.

The next question creating a few reactions is the one in which the respondent is asked to estimate the lowest monthly income the household needs to make ends meet (End2). This question has both been perceived as difficult to answer (5 respondents) and as time consuming (10 respondents). Still, we have not reports saying the respondent reacted negatively to the question. Concerning this question as well as the previous one mentioned, we assumed that the question could create problems. This is a highly hypothetical question since most households adjust their consumption to their income. Estimating the minimum income needed is therefore difficult since "needs" could be interpreted in different ways.

We received a couple of reports on the questions about whether someone in the household made or received regular payments to/from a different household (Gitt1 and Faatt1), we have received a couple of reports. One negative reaction to the question about making regular payments, and one reaction saying that it was difficult to answer the question about receiving

payments. Since these questions could indicate difficulties in the financial situation in some households, hence being sensitive questions, reactions to these questions are not unexpected.

We now turn to the results of the analysis of partial non-response. There are only two target variables from the household questionnaire with high partial non-response. Both are mentioned above: Total housing cost ( HH070 ) and minimum income required to make ends meet ( HS130 ).

For 30 households (13 per cent) total housing cost is missing. Total housing cost is a variable constructed from several of the questions in the questionnaire. The questions contributing most to missing in HH070 is partial non-response on the question on interest on mortgage. Only a few do not know the total amount paid (including both interest and repayment of principal), but many have problems specifying interest. However also partial non-response to questions on insurance and service charges contribute to missing in HH070.

Missing is most common among selected persons who own the dwelling along with others (15 of 83 households). Two thirds of these are women. Table 5 also shows the rather surprising result that missing is not high when the respondent is a selected person who does not own the dwelling. This of course does not mean that the answers are of good quality. It may however indicate that in the cases when the selected person is responding on the household questionnaire even if he does not own the dwelling he may have had some help from others in the household.

**Table 5. Number of households by position in the household of the person who answered the household questionnaire and number of households with missing information on total housing costs (in parenthesis).**

Family relations to selected person	Selected person is responsible for the accommodation alone	Selected person is responsible for the accommodation with others	Other than selected person is responsible	N
Selected person	104 (9)	83 (15)	24 (2)	211 (26)
Spouse or cohabitant	3 (1)	8 (2)	2 (0)	13 (3)
Parent, parent-in-law or grandparent	-	-	11 (1)	11 (1)
Child, brother/sister or other	1	1	1	3 (0)
N	108 (10)	92 (17)	38 (3)	238 (30)

The variable minimum income needed is a single question. The high partial non-response (16 per cent) is a symptom that the question is not well understood. Table 6 indicates that there are only small differences in percentage of missing according to which person in the

household answers the question. Does this indicate that knowledge about the economy of the household is not the most important determinant of the percentage missing?

**Table 6. Number of households by position in the household of the person who answered the household questionnaire and number of households with missing information on minimum income required to make ends meet (in parenthesis).**

Family relations to selected respondent	Selected person owns the residence alone	Selected person owns the residence along with others	Other than selected persons owns the residence	N
Selected person	104 (17)	83 (12)	24 (5)	211 (34)
Spouse, partner or co-habitee	3 (2)	8 (0)	2 (0)	13 (2)
Parents, parents-in-law or grandparents	0	0	11 (1)	11 (1)
Children, stepchildren, (half) brother and sisters or other	1	1	1	3 (0)
N	108 (19)	92 (12)	38 (6)	238 (37)

## 9.2 Personal questionnaire

The personal questions are evaluated mainly on basis of partial non-response and response from the interviewers.

In Norway there is a main difference between basic labour questions, which are answered by all adult members of the household and other personal questions answered by the selected person only. A large proportion of the interviews, which are done with other household members than the selected person, are proxy interviews. The information from proxy interviews must be assumed to be of poorer quality.

Only a few of the personal questions are mentioned by the interviewers as problematic and none are mentioned by more than one or two interviewers. The questions mentioned are PH060 (unmet need for dentist consultation), PY040 (monthly earnings) and PL190 (when began first regular job). One respondent had problems answering this last question, one used long time to give an answer. We assume that this is due to problems with remembering back in time.

Judged from partial non-response the questions on health (PH010 - PH070) seems to function. Partial non-response is below 2 per cent. There are some indications that the respondents do not distinguish between medical specialists and other doctors. In Norway patients must be referred to a medical specialist by a general practitioner. Looking at the answers from the six respondents with an unmet need for a specialist who answered Other reasons, one answered

'Was not ill enough', and two that they 'thought there was no cure for their illness'. One can suspect that these respondents have not even seen a general practitioner.

The detailed labour questions (PL110 -PL210) have been answered only by the selected respondent. Partial non-response for these questions is very low, 2 percent or lower.

OBS There is a mistake in the Norwegian questionnaire. PL200 and PL210 have been asked to all adult household members. This has not been corrected on the files. The basic labour questions have been asked to all adult household members. There has been one mistake. PL100 on secondary job has been asked only to the selected respondent.

Partial non-response in basic labour questions is generally higher than for the detailed labour questions. This is due to higher non-response among other household members than the selected respondent. Among the selected respondents only PL010 had a partial non-response above 1 percent.

**Table 7. Percentage partial non-response among selected respondents and other household members \* Per cent**

Target variable	All	Partial non-response	
		Selected respondent	Non-selected respondent
PE010	1,8	0,8	2,8
PE030	4,3	0,8	7,5
PL010	4,3	1,7	6,7
PL015	0,2	0	0,4
PL020	2,0	1,3	2,8
PL030	3,5	0,8	5,9
PL060	0,6	0,6	0,6
PL070-090	3,5	0,8	5,9

\* Other target variables have no partial non-response

It is tempting to assume that the high partial non-response among the non-selected respondents is due to indirect interviewing. Surprisingly enough this is not necessarily the case. F.i. the 15 household members with partial non-response on variable PL030 have all been interviewed directly.

## 10 CONSTRUCTION OF INCOME TARGET VARIABLES

All income data in the pilot survey of the Norwegian SILC are collected from different administrative registers. Income data refers to the year 2000. In this chapter we give a brief description of which data sources (registers) have been used, how income was defined at component level, and to what extent this income concept is in agreement with the SILC



income definition. In addition we compare the income data in SILC with that of the national income statistics.

## **10.1 Data sources**

In order to construct total household income and disposable income data on income have been collected from several administrative registers. In SILC the following data sources have been used:

(a) Tax Return Register

(Employee and self-employment income, property income, taxable pensions and benefits)

(b) Tax Register for Personal Tax Payers

(Taxes on income and wealth, social security contributions)

(c) Register for end-of-the-year certificates

(Unemployment benefits, company car, early retirement pension)

(d) The State Housing Bank

(Housing allowances)

(e) The State Educational Loan Fund

(Education related benefits)

(f) National Insurance Administration

(Old-age and disability pension, survivor's benefit, family related allowances)

(g) Social statistics

(Social assistance)

A comprehensive data file on income is thereupon established by linking the total resident population to all the different income registers. The key that links the individual to the registers is the Personal Identification Number.

## 10.2 Income concepts

This section gives an overview of how different income data from registers have been organised in order to be comparable to the income concepts outlined in the SILC guidelines. In addition references are made to any departures from these guidelines.

All income data derived from registers are recorded gross at component level. Furthermore, all income data are collected at the individual level (i.e. the person registered as the receiver of the income). This also concerns typically 'household' related incomes such as dwelling support or social assistance. Register data also includes the income of children aged 13-16 years at the individual level. The income of children aged 12 or younger are, however, included in their parent's income (e.g. interest received).

### Gross Employee Cash income - (PY010)

Defined as the sum of all taxable wages and salaries including overtime, holiday pay, tips and bonuses. The wage concept also includes non-cash income such as free telephone and newspapers, low-interest loans etc. However, company car is not included in the cash wage concept.

Deviation from the SILC concept:

- payments to foster parents (included in wages, cannot be separated from wages)
- severance and termination pay (----- "-----)
- sickness benefit received after 15 days or more of sickness (-----"-----)
- fringe benefits other than company car (included in the wage concept).

With the exception of sickness payments these deviations from the SILC definition are expected to be of a minor importance. (In the main survey of SILC (2003) fringe benefits other than company car may be identified from registers).

### Gross Non-Cash employee income - (PY020)

This income item only includes the tax-assessed benefit from using a *company car*.

It may however at a later stage be possible to collect data on the list price of the car from registers.

### Gross cash profit/loss from self-employment - (PY050)

Defined as net entrepreneurial income from self-employment after depreciation and deficit. In addition rental income is included.

Deviation from the SILC concept:

According to SILC interest payments on business loans should be deducted from operating surplus. In register data no such deductions are made in entrepreneurial income. Instead total interest payment can be collected from the tax return, but it is not possible to distinguish business interests from private interests.

Gross cash losses from self-employment - (PY060)

Entrepreneurial income is collected *net* in register data. Gross cash losses thus appear as negative amounts in variable PY050: Gross cash profit.

Non-cash Income from self-employment - value of own goods for own consumption- (PY070)

The tax-assessed benefit from consuming own goods (estimated by the tax authorities) is included in Gross cash income from self-employment.

Interest, dividends, profit from capital investment in unincorporated business - (HY090)

Interest and dividends are taxable income available in the tax return. In addition some minor income from property are included, for instance profit from life insurance.

Income from rental of property or land - (HY040)

Not relevant. All rental incomes are included in self-employment income.

Regular pension from private scheme - (PY080)

Data on private pension payments received are available from the tax return register.

Family related allowances - (HY050)

Includes the following income components:

- family allowance
- maternity allowance (birth grant)
- parents' tax deduction
- cash-for-care benefit
- child support for single parents (childcare and education)
- transitional benefit for single parents

Deviation from the SILC concept:

It is not possible to identify daily cash benefit related to *parental leave benefit (birth or adoption)*, which is part of *wages*.

Housing allowances - (HY070)

Includes dwelling support in cash to renters and owner-occupiers.

Deviation from SILC concept:

The benefit from renting a subsidised dwelling is not included in the income concept. There are, however, few such dwellings in Norway.

#### Unemployment benefits - (PY090)

Includes unemployment benefit for employees and unemployment benefit for the self-employed.

Deviation from SILC concept:

No data available on benefit (in-kind) related to vocational training.

#### Old-age function - (PY100)

Includes old-age pension from the social security system and occupational pensions.

Deviation from SILC concept:

It was not possible to split *occupational pensions*, i.e. social insurance benefits from employers scheme, into different types of beneficiaries, e.g. old-age pensioners, disabled or survivors. Instead all types of occupational pensions have been included under the old-age function.

#### Survivors' function - (PY110)

Includes survivors' pension from the National Insurance. In addition several minor income items have been included that are received mainly by survivors, e.g. tax-free wage income and holiday pay earned by the deceased.

Deviation from SILC concept:

Not possible to include funeral grant in the income concept. This benefit is transferred directly to the firm of undertakers.

#### Social benefits in the sickness - (PY120)

All sickness benefits are included in wages and salaries and cannot be specified in registers.

#### Invalidity benefits - (PY130)

Includes disability pension from the National Insurance, early retirement pension (AFP), basic and additional benefits and compensation for occupational injuries.

Deviation from SILC concept:

No data available on economic integration of the handicapped.

#### Education related allowance - (PY140)

Includes scholarship from the State Educational Loan Fund.

#### Social assistance - (HY060)

Includes the total amount received in social assistance (benefits and loans).

Regular inter-household cash transfers received - (HY080)

Includes alimonies received from former spouse.

Deviation from SILC concept:

No data available on private cash support from parents to children living in a separate household (e.g. students), or vice versa.

Other income - (PY150)

Includes social security benefits that cannot be specified, for instance additional payments of pensions or payments in advance.

Total gross household income - (HY010)

The sum of all income components specified above.

Total Tax on income and social contribution - (HY140)

Includes assessed income and wealth taxes and social contributions.

Deviation from SILC concept:

This variable includes both income and wealth taxes. It is difficult to specify income or wealth taxes in tax files because all taxes are recorded *net*, after special tax deductions (e.g. parent's tax deduction, special tax deductions for residents of Finnmark etc.).

Regular taxes on wealth - (HY120)

Included in HY140: Total tax on income.

Regular inter-household cash transfers paid - (HY130)

Includes paid maintenance to children and former spouse (alimony). These payments appear as deductions in the tax return.

Deviation from SILC concept:

No data available on private cash support e.g. from parents to children living in another household.

Total disposable income - (HY020)

Defined as Total gross income minus current transfers paid.

### 10.3 Data quality

All income data that are used in the construction of the SILC income variables are collected from administrative registers. Although the quality of register data is considered to be high in Norway, there are potential sources of errors.

One source of error concerns errors in the reporting of income in the tax return. We do not believe that this is a substantial problem, however, due to routinely reporting of taxable income from the employer, the Social Security Office, banks and credit institutions to the tax authorities. The only exception may be in respect to self-employment income, where there is less information available for such controls. In addition all income from informal work ('black income') is missing in register data.

There will always be imperfections in data from an administrative register, for instance due to processing errors etc. Great care has been taken to minimise the effect of such errors. For instance, outliers are frequently discovered and corrected by comparing the reported income amounts in one register with tax figures for the same individuals in another register. Lack of consistency is often discovered by comparing aggregated figures for taxable income with the sum of individual income entries. However, because of the large amount of data in registers it will not be possible to perform detailed consistency checks for all individual records.

The register data used to construct the SILC income concept are the same that are used in the production of national income statistics. The definition of income, particularly at the component level, are, however, slightly different.

Table 8 presents the sum of total gross income and disposable income for the resident population of Norway as of 31 December 2000.

**Table 8. Total gross income and disposable income. Billion NOK. 2000**

	SILC definition	National definition	Difference
Total Gross income	832,2	850,2	18,0
Disposable income <sup>1</sup>	615,9	633,3	17,4

<sup>1</sup> In the national definition this income concept refers to 'After-tax income'.

As can be seen there is a difference between the national definition and the SILC definition of approximately 17-18 billion NOK in 2000, for both disposable and gross income.

The main difference can be explained by differences in the definitions of property income. The national definition of income includes *capital gains*. In 2000 this income item amounted to 17 billion NOK. Capital gains are, however, not included in the SILC definition of property

income. The remaining difference can be attributed to difference in the population unit. The national figures also include the income and taxes of residents of Spitsbergen (Svalbard), while the SILC figures only include residents of mainland Norway.

## **11 INTEGRATION WITH EXISTING SURVEYS**

Statistics Norway has since 1996/1997 carried out two surveys on living conditions each year, one cross-sectional survey and (from 1997) one longitudinal survey. The longitudinal survey started with a sample of 5000 persons aged 16-79 years. Each year after 1997 the sample was supplemented with new 16-year old persons and new immigrants so that the sample of the survey in year T is representative of the population in year T. The interview time should be less than 30 minutes, but information from interview is supplemented with information from registers, f.i. on income.

The plan is to build EU-SILC, both the cross-sectional and the longitudinal part, on this longitudinal survey. To do this the sample must be adjusted so that it covers the population required in EU-SILC and adjusted with respect to sample size. The interview time of EU-SILC is estimated to approximately 30 minutes. To integrate the EU-SILC questionnaire with the existing questionnaire would mean that the interview time would be considerably longer (but not 60 minutes). We consider it important that the interview time is kept at about 30 minutes. There are two reasons for this: first for economic reasons it is considered essential to make the interviews mainly by telephone, and second the burden on the household must be kept on an acceptable level so that attrition in the longitudinal survey is low. Hence a considerable part of the information in the existing longitudinal survey on living conditions will be discontinued. Still there is a considerable overlap between the content of EU-SILC and the existing longitudinal survey (and the register information will of course be common).

We consider it important that information asked in both EU-SILC and the existing longitudinal survey will be comparable with both sources as far as possible. This has therefore resulted in some adjustments of the EU-SILC questions. These are documented in the annex.

## **12 USE OF REGISTER INFORMATION ON EMPLOYMENT**

The Register of Employees contains information on occupation, working hours for all jobs the employee has, industry and size of the company.

There are however differences between the information from register and from interview which must be kept clear. The register is not complete and there may be a lag in the register information. Perhaps the most serious difference is in working hours. The working hours in the register are the working hours agreed with the employer, not the usual working hours. Hence overtime work will usually not be counted.

Of minor importance is that the register has no definition of the main job and of secondary jobs. However based on information on working hours agreed we will probably come quite close to the definition of main job by the respondent.

The most useful information from the Register of Employees is the information on industry and in the future also on occupation. We plan to use information on industry in the main survey. The plan is to link information on the company's name and address from the register and ask in the interview if this is correct. This will save costs by reducing the coding work and will probably increase the quality of the NACE coding of industry. Information on occupation from the register is not 100 per cent complete now. When in the future the information on occupation is good enough the plan is to use register information on occupation in the same way as we now plan for industry.

### **13 CONCLUSION**

It was not an aim to test non-response, partly because we have a lot of experience with comparable surveys, partly because the main survey will be conducted under different circumstances. Most important a major part of the sample will be the sample of the existing longitudinal survey on living conditions. The response rate was low. However we attribute this to the circumstances in which the fieldwork was carried on. There are no other indications that the response rate will be low.

An important aim was to test the field procedures. The instructions on which person should answer the household questionnaire have been followed in 90 per cent of the households. We don't know what is the reason is for not following the instructions. In the main survey we will ensure that all interviewers know the instruction. Surprisingly there are no indications that the quality of the responses from the households that did not follow the instructions is lower than in other households.

All adult members of the household should answer the basic labour questions, preferably by direct interviews. A little less than half (47 per cent) of the interviews with other household members than the selected respondent were direct interviews. Item non-response is higher for non-selected respondents. We have not had time yet to examine systematically if this is due to proxy interviews, but we have indications that this is not necessarily the case. On the whole our conclusion is that interview with other household members does not constitute a big problem. We got interview with most adult household members. We would prefer that a larger part of the interviews were direct interviews, and in the main survey we will try to motivate the interviewers to increase the percentage of direct interviews.

Another important aim was to get an estimate of the interview time. The estimate is approximately 22 minutes, lower than expected. The fact that the eldest were not included and that 'difficult' respondents were often not interviewed must be considered.

The most important aim was to test the quality of the responses to the questions. In the household questionnaire there seems to be serious problems with two of the target variables: total housing cost and minimum income required to make ends meet.

In the personal questionnaire there were no indications from reports from the interviewers and from item non-response of any serious problems. In advance we expected that 'Age completed initial education' and 'Unmet need for a specialist' would cause problems. We there decided to split the question on initial education into several questions. This seems to have been



successful. The percentage missing on PE030 is 4 per cent. It is among the highest of the personal questions. However we consider it reasonable. The questions on unmet needs seem to have worked well judged from reports from the interviewers and from percentage missing. We have not used a designed that makes it possible to analyse how meaningful the answers are.

The registers on income that Statistics Norway use gives very detailed information of high quality on income. We consider that the gross and net income concepts of Eurostat can be approximated very well by our register information. For some of the income components there are some deviations. Technically the linking of the pilot with the register caused no problems. We also has long experience in this kind of linking.

### COMPARISON OF EUROSTAT QUESTIONNAIRE AND NORWEGIAN EU-SILC

If there are no comments to a question it means that we consider them directly comparable.

#### HOUSEHOLD QUESTIONNAIRE

##### **Eurostat Q. Norwegian Q.**

Q1 Hus, Hus1-6

The standard Norwegian question is much more detailed, but most categories are easily translated to Eurostat categories. To construct the Eurostat categories we added a question on number of apartments/flats in the building.

Q2 Bol1

Only rooms which are at least 6 sqm are counted. The consequences for comparability are very small.

Q3b Bol3

'For the sole use of the household' is not included in the Norwegian questionnaire. We have interpreted this to mean that the bath-room is inside the dwelling.

Q4a Bol6a,b

We have split this question in two: Rot in windows or floor and Leaking roof, damp walls or floor.

Q4b Bol6c

The Norwegian q. asks 'not enough **daylight**'.

Q5 Eie, leie

The Norwegian q. is more detailed. However it is quite clear how to aggregate categories to construct the Eurostat categories of owners and tenants. To distinguish between tenants paying rent at or below market price we asked whether the rent that is paid is market rent (question Husleie2). To distinguish household with a rent-free accommodation we asked whether the household pay rent (question Husleie1).

Q6 Laan

Q7

Not asked in Norwegian q. Respondents who don't know interest are filtered to Q9.

Q8 Lrent1

This is asked after Lutg2, the equivalent to Q12.

Q9, 10 Lrent2

Q11, Q13

These questions are not asked. The question on interest (Q7) asks for the gross amount. In Norway interests are deducted from income. Except for low incomes the tax relief will be 28 per cent of the interest. OBS In the calculation of HH070 the tax relief has not been deducted. It is possible however to make a good approximation by imputing tax relief. We didn't have time for this in the pilot.

Q12 Lutg1,2

We first ask if the respondent want to give the amount per month, quarter or year, then we ask for the amount.

Q14-17 Lrent5-8

Q18, 21, 22, 23

This information is taken from register. Information is given in HY070.

Q19 Husleie3, 4

Husleie3 ask whether the respondent wants to give the amount per month, quarter or year.

Husleie4 is equal to Q19 but specify somewhat more what has and what has not to be included. The specifications are in accordance with the definitions.

The question asks for the gross amount.

Q20a-e Husleie5-7

The Norwegian q. asks for electricity, other heating and hot water. Charges on (cold) water will always be included. Gas is virtually nonexistent in this context and is therefore not asked about.

Q24 Avg1, 2,3a,3b

Avg1 asks about insurance. Avg2 asks about charges for local services like water, sewage and refuse removal and other municipal charges (excluding municipal tax on houses/dwellings).

Avg3a, b asks for regular maintenance costs.

Q25 Tyng

HH070

The Norwegian questionnaire asks some questions which are not specified in the Eurostat questionnaire. Futg1-3 are asked to what we call self-owners. They do not pay rent, but

common costs (local charges, maintenance etc) are paid as a fixed monthly amount.

Likn1, 2 are asked because some self-employed (e.g. farmers) cannot separate loan on the house from loan on their business. For these households the value of living in the house is calculated during the tax assessment. This value is incorporated in HH070 instead of interest.

Q26 Disp1, 1a 2, 2a...,5, 5a

Q27 AndLaan1

Q28 AndLaan2

Q29 Raad1 - 4

Q30 Prob1 - 4

Q31 End1

Q32 End2

Q46, 47 Gitt1, 2, 3

The Norwegian q. differs because it excludes alimonies to former spouse/children. Information on alimonies are taken from register. HY130 is therefore calculated as a sum of information from register and from Gitt3.

Q48

This question is not asked. Alimonies to former spouse are not taxed on the person that pays them. Other private payments are not deducted, and hence they are taxed.

Q50, 51 Faat1, 2, 3

The same as for Q46, 47. HY080 is calculated as a sum of information from register and from Faat3.

Q52

This question is not asked. Alimonies from former spouse are taxable. Other private transfers are (very seldom?) taxed.

## **PERSONAL QUESTIONNAIRE**

PQ1 Arb14

The only difference is that the Norwegian question is only asked respondents working less than 32 hours a week. The categories are read by the interviewer instead of using av card.

PQ2, 3 Arb1, 2

Arb1 tells explicitly that family workers shall be included. In Arb2 specifications are given in the interviewer manual.

PQ4 Arb3

Arb3 asks for attempts during the last four weeks (or more precisely since a date four weeks before the date of the interview).

PQ5 Arb5 - I

PQ6 Arb6

PQ7 Arb4

Arb4 refers to the last four weeks.

PQ7c Arb7a

PQ7d, 8 Arb7b

PQ7e, 10 Arb7c, d  
Arb7d asks self-employed if they have any employees.

PQ7f, 14 Arb7e

PQ7g, 15 Arb7f  
Arb7f asks 'Did you supervise or manage any personnel **or did you in other ways have a superior position**'.

PQ7h, 16 Arb7g

PQ9 Arb8a-c  
Arb8a-c asks for the name and address of the firm. Industry is coded from register information on the firm.

PQ11 Arb12a, b

PQ12 Arb13  
Arb13 explicitly mentions that paid overtime and extra work at home shall be included.

PQ13 Arb15a, b  
Arb15a asks the respondent to tell if he prefer to give the amount per hour, week, two-week period, month or year.

PQ17 Arb19

PQ18 Arb20

PQ19 Arb21a

PQ20 Arb21b

PQ22 Arb22

PQ106 Hels1

PQ107 Hels2  
In addition to chronic illness Hels2 asks for 'any consequence of injury or any disability'.

PQ108 Hels3a1, a2, 3b  
Arb3a1: ' Does this (chronic illness) lead to limitations in your daily activities'  
Arb3a2: ' Have these limitations lasted for at least six months'  
Arb3b: ' Would you say that you are strongly limited or somewhat limited'?

PQ109 Hels4a

PQ110 Hels4b, 4bSp

PQ111 Hels5a

PQ112 Hels5b, 5bSp

PQ113 - 116 Register

PQ117 Sivstat

PQ118 Siv  
Categories: Yes married, Yes cohabitating, No

PQ119 - 123 Constructed from household roster

PQ124 Arb23a

PQ125 Register

PQ126 Arb23c-f  
To get information of good quality this question was split into four questions.  
Arb23c: ' Have you got any education in addition to the obligatory'?'  
Arb23d: ' Have you interrupted this education because of work or birth/childcare'?'  
Arb23e (if yes): ' How old were you when you interrupted the education the first time because of work or birth/childcare'?'  
Arb23f (if no): ' How old were you when you finished education '?

PQ127 Register

PQ128 Arb24

PQ129 Arb25

PQ130 Arb26a - 1  
Arb26a: Main activity in July 2001  
Arb26: Was the main activity the same during the whole period July 2001 to June 2002 ' ?  
Arb26b: Main activity in August 2001.  
Etc.

## Recent publications in the series Documents

- 2001/2 K.Ø. Sørensen, J.L. Hass, H. Sjølie, P. Tønjum and K. Erlandsen: Norwegian Economic and Environment Accounts (NOREEA) Phase 2
- 2001/3 G. Haakonsen, K. Rypdal, P. Schøning and S.E. Stave: Towards a National Indicator for Noise Exposure and Annoyance: Part I: Building a Model for Traffic Noise Emissions and Exposure
- 2001/4 T. Åvitsland: User Costs of Real Capital
- 2001/5 H.V. Sæbø: Strategic planning and management in Instituto Nacional de Estatística, Mozambique: Short Term Mission Report, November 2 - 9, 2000
- 2001/6 H.V. Sæbø, J. Byfuglien and R. Johannessen: Quality issues in Statistics Norway
- 2001/7 J. Byfuglien: Comparability of income data: Project report
- 2001/8 A.S. Abrahamsen, M.Q. Andersen and R. Ragnarsøn: Project: Quality of Manufacturing Statistics and Evaluation of Methods for Producing Preliminary and Final Figures. Final Report
- 2001/9 T. Skoglund: Employment in the Norwegian National Accounts
- 2001/10 O.C. Moen: Nordic Economic Growth in Light of New Theory: Overoptimism about R&D and Human Capital?
- 2001/11 H.V. Sæbø: Strategic Planning and Management in Instituto Nacional de Estatística, Mozambique: Short Term Mission Report, August 27 - September 6, 2001
- 2001/12 B. Hoem: Environmental Pressure Information System (EPIS) for the household sector in Norway
- 2001/13 H. Brunborg, I. Bowler, A.Y. Choudhury and M. Nasreen: Appraisal of the Birth and Death Registration Project in Bangladesh
- 2001/14 K. Rypdal: CO<sub>2</sub> Emission Estimates for Norway. Methodological Difficulties
- 2001/15 E. Røed Larsen: Bridging the Gap between Micro and Macro: Interdependence, Contagious Beliefs and Consumer Confidence
- 2001/16 L. Rogstad: GIS-projects in Statistics Norway 2000/2001
- 2002/1 B. Hoem, K. Erlandsen og T. Smith: Comparisons between two Calculation Methods: LCA using EPIS-data and Input-Output Analysis using Norway's NAMEA-Air Data
- 2002/2 R. Bjørnstad: The Major Debates in Macroeconomic Thought - a Historical Outline
- 2002/3 J. L. Hass and T. Smith: Methodology Work for Environmental Protection Investment and Current Expenditures in the Manufacturing Industry. Final Report to Eurostat.
- 2002/4 R. Bjørnstad, Å. Cappelen, I. Holm and T. Skjerpen: Past and Future Changes in the Structure of Wages and Skills
- 2002/5 P. Boug, Å. Cappelen and A. Rygh Swensen: Expectations and Regime Robustness in Price Formation: Evidence from VAR Models and Recursive Methods
- 2002/6 B.J. Eriksson, A.B. Dahle, R. Haugan, L. E. Legernes, J. Myklebust and E. Skauen: Price Indices for Capital Goods. Part 2 - A Status Report
- 2002/7 R. Kjeldstad and M. Rønsen: Welfare, Rules, Business Cycles and the Employment of Single Parents
- 2002/8 B.K. Wold, I.T. Olsen and S. Opdahl: Basic Social Policy Data. Basic Data to Monitor Status & Intended Policy Effects with Focus on Social Sectors incorporating Millennium Development Goals and Indicators
- 2002/9 T.A. Bye: Climate Change and Energy Consequences.
- 2002/10 B. Halvorsen: Philosophical Issues Concerning Applied Cost-Benefit Analysis
- 2002/11 E. Røed Larsen: An Introductory Guide to the Economics of Sustainable Tourism
- 2002/12 B. Halvorsen and R. Nesbakken: Distributional Effects of Household Electricity Taxation
- 2002/13 H. Hungnes: Private Investments in Norway and the User Cost of Capital
- 2002/14 H. Hungnes: Causality of Macroeconomics: Identifying Causal Relationships from Policy Instruments to Target Variables
- 2002/15 J.L. Hass, K.Ø. Sørensen and K. Erlandsen: Norwegian Economic and Environment Accounts (NOREEA) Project Report -2001
- 2002/16 E.H. Nymo: Influence of Migrants on Regional Variations of Cerebrovascular Disease Mortality in Norway. 1991-1994
- 2002/17 H.V. Sæbø, R. Glørsen and D. Sve: Electronic Data Collection in Statistics Norway
- 2002/18 T. Lappegård: Education attainment and fertility pattern among Norwegian women.