Statistics Netherlands

Division of Social and Spatial Statistics
Statistical analysis department Heerlen

Heerlen
The Netherlands

Final quality report

EU-SILC 2009

The Netherlands

Preface

In recent years, Statistics Netherlands has focused on an increased use of register data instead of survey data in the production process of statistical information. By making efficient use of register data, Statistics Netherlands intends to improve the accuracy of the statistical information and, at the same time, to decrease the response burden on households. Examples of administrative registrations are the Population Register (the municipal basic registration of population data; in Dutch: Gemeentelijke BasisAdministratie - GBA), data on social security and tax data. The Population Register (GBA) contains information on age, sex, ethnicity, place of birth, place of residence, marital status and other information for all (registered) persons living in the Netherlands. This registration has been available from 1995 onwards, and is updated monthly. The Labour Force Survey (LFS) is one of the surveys that are linked to the GBA. The design of the LFS is based on a face-to-face interview (CAPI), followed by a four-wave panel by telephone interview (CATI).

In the Netherlands, 2005 was the initial year for the EU-SILC survey. And for various reasons (costs, response burden, available information), it was decided to use the fifth wave LFS-respondents as the EU-SILC sampling frame. In doing so, a relatively short telephone-interview (on average 15 minutes) was sufficient to collect the additional EU-SILC information. After the fieldwork, all information based on the Population Register, register data on income and the LFS was matched to to the EU-SILC respondents.

Statistics Netherlands implemented the integrated four-year rotational design which means that the cross-sectional en longitudinal EU-SILC data are based on the same set of sample observations. Rotational design refers to the sample selection based on a number of subsamples or replications. Once the system is fully established (from EU-SILC 2008 onwards) the sample for any one year consists of four replications which have been in the survey for 1, 2, 3 or 4 years. Each year one of the four replications is dropped and replaced by a new one. The new group consists of new sample persons who were drawn from the Labour Force Survey.

Preface

1.   Common Indicators

1.1   Common longitudinal European Union indicators EU-SILC 2009

The longitudinal dataset 2006-2009 of the EU-SILC operation comprises a panel of four years (2006-2009). The main objective of the four-year panel rotation is to deliver an adequate data basis for the calculation of the persistent-at-risk-of-poverty indicator. Persistent at risk of poverty occurs if a respondent is at risk of poverty (income below 60 % of median income) in the last wave of the four-year panel and has been at risk of poverty at least two times during the preceding waves (2006, 2007, and 2008).

Table 1.1: Persistent at-risk-poverty rate EU-SILC 2009

Indicator	Value
Persistent at-risk-of-poverty rate after social transfers - total	4,7
Male	5.4
Female	4.1
At-risk-of-poverty rate after social transfers - 0-17 years	4,9
Male	6,5
Female	3,3
At-risk-of-poverty rate after social transfers - 18-64 years	4,8
Male	5,2
Female	4,4
At-risk-of-poverty rate after social transfers – 65+ years	4,2
Male	4,7
Female	3,9

2. Accuracy

2.1 Sampling design

The EU-SILC survey is an annual survey with a four-year rotational panel and has been carried out as an integrated survey, covering both cross-sectional and longitudinal primary target variables by a single operation. The cross-sectional sample of SILC 2009, the fifth year of EU-SILC in the Netherlands, consists of four rotational groups. Group R1' has entered the survey in 2006 and sample persons in group R2' were interviewed for the first time in 2007. Group R3' has entered the survey in 2008 and group R4' consists of new sample persons who were drawn from the Labour Force Survey in 2009.

Figure 2.1. Rotational design EU-SILC

EU-SILC 2005	R1	R2	R3	R4				
EU-SILC 2006		R2	R3	R4	$R1^1$			
EU-SILC 2007			R3	R4	$R1^1$	$R2^1$		
EU-SILC 2008				R4	$R1^1$	$R2^1$	$R3^1$	
EU-SILC 2009					$R1^1$	$R2^1$	$R3^1$	$R4^1$
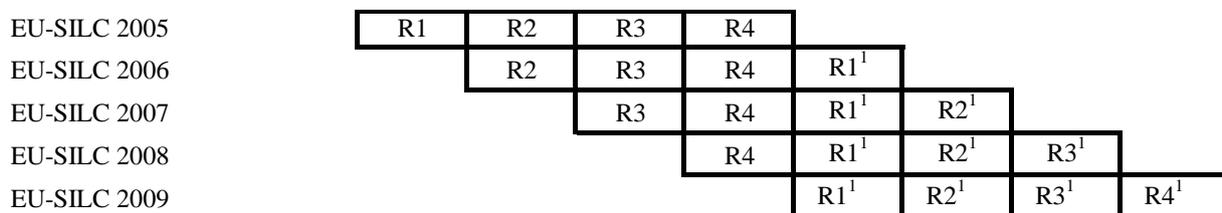
2.1.1 Type of sampling

Sample persons in new rotational groups were partly drawn from the Labour Force Survey (LFS). The sampling frame of the LFS is the Dutch municipality administration (Gemeentelijke Basisadministratie or GBA). The LFS-sampling design can be classified as a stratified two-stage sampling design, with municipalities as primary sampling units and addresses as secondary sampling units. The sampling of first stage elements is with probability proportional to size (number of addresses per municipality). Municipalities with 7,300 addresses ore more are always in the sample. The second stage elements are selected with simple random sampling such that the total sampling design becomes self-weighting. From these addresses further sampling units are constructed: households. For the collection of detailed information on social variables one member of the household aged 16 or older is selected (the so-called selected respondent).

2.1.2 Sampling units

The sampling units are addresses that are registered in the sampling frame. All households on selected addresses are eligible for the survey, up to a maximum of three households per address.

2.1.3   Stratification criteria

Stratification involves the division of the population into sub-groups, or strata, from which independent samples are taken. The stratification variables are the 40 COROP-regions (NUTS3). These are regional areas within the Netherlands and are used for analytical purposes by, among others, Statistics Netherlands. The Dutch abbreviation stands for Coördinatiecommissie Regionaal Onderzoeksprogramma, literally the Coordination Commission Regional Research Programme. Applying this type of stratification allows for representative samples on a regional level.

Figure 2.2. COROP regions in the Netherlands

2.1.4   Sample size and allocation criteria

The four-year component of EU-SILC 2009 consists of 1,286 households in group R1' with accepted household interviews in 2006, 2007, 2008 and 2009.

Figure 2.3. Rotational design EU-SILC

EU-SILC 2005	R1	R2	R3	R4				
EU-SILC 2006		R2	R3	R4	R1[1]			
EU-SILC 2007			R3	R4	R1[1]	R2[1]		
EU-SILC 2008				R4	R1[1]	R2[1]	R3[1]	
EU-SILC 2009					R1[1]	R2[1]	R3[1]	R4[1]

4-year longitudinal subsample

EU-SILC 2005	R1	R2	R3	R4				
EU-SILC 2006		R2	R3	R4	2.339			
EU-SILC 2007			R3	R4	1.876	3.731		
EU-SILC 2008				R4	1.552	2.839	3.621	
EU-SILC 2009					1.286	2.449	2.914	R4[1]
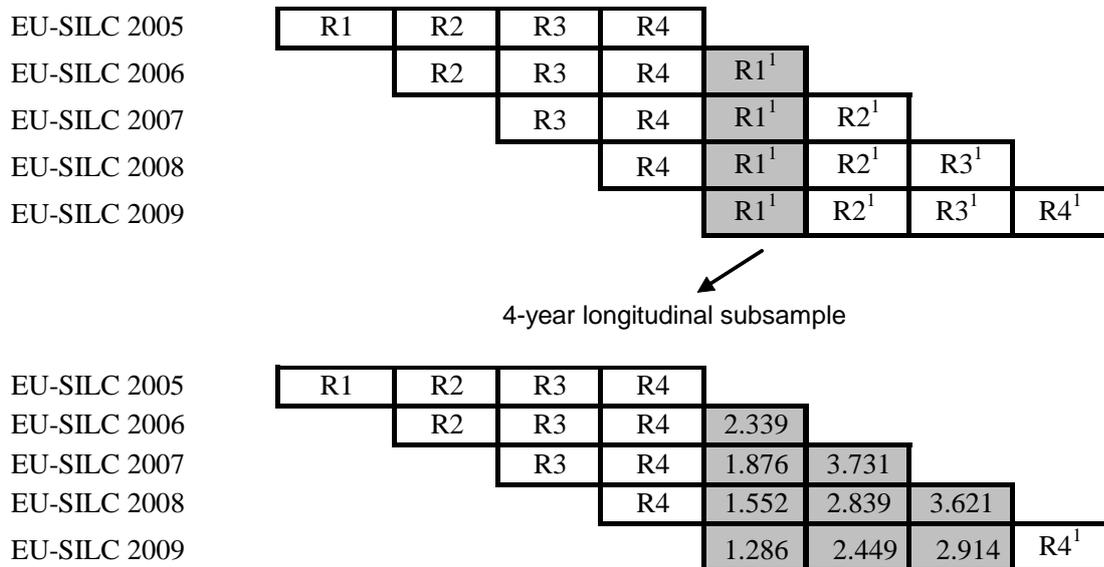
Member states have to achieve a minimum effective sample size for the cross-sectional and longitudinal sample. For the Netherlands the net cross-sectional sample size is 6,500 households and 6,500 selected respondents. Correcting for estimated design effects, the minimum achieved sample size should be larger. Similar considerations apply to the longitudinal sample: in this case the net effective sample size for the Netherlands is 5,000 households and 5,000 selected respondents.

Combining rotational group R1', R2' and R3' 6,649 households have accepted interviews in two consecutive years (2008 and 2009).

2.1.5   Sample selection scheme

Primary sampling units are selected by means of systematic sampling with probability proportional to size. Therefore the ordering of these units in the strata is relevant: the primary sampling units in each of the strata are randomly ordered. The secondary sampling units are selected with simple random sampling in order that the total sampling design becomes self-weighting.

Addresses corresponding to institutions, addresses that have been part of a survey sample in the previous year, and addresses in some small regions of the national territory (West Frisian Islands) are removed from the sample. These addresses are not part of the reference population.

2.1.6   Sample distribution over time

The following table provide for the 4-year longitudinal sample an overview of the cumulative sample development during the fieldwork period .

Table 2.1: cumulative sample size over time

Fieldwork in:	2006	2007	2008	2009
June	1,009	819	692	557
July	1,130	886	721	612
August	200	171	139	117
Total	2,339	1,876	1,552	1,286

2.1.7   Renewal of samples: rotational groups

In the Netherlands, 2005 was the initial year of EU-SILC. A new sample was constructed and divided into four rotational groups. One of the subsamples was purely cross-sectional and was not followed up in 2006. Respondents in the second subsample participated two years, in the third subsample three years, and in the fourth subsample four years. Because accurate panel attrition rates were not available at the start of the EU-SILC survey, the subsample sizes are chosen to be of quite different sizes in order to guarantee a longitudinal sample of sufficient size.

2.1.8   Weighting

The longitudinal data set for individuals in EU SILC 2009 contains information on the eligible individuals traced from original sample households in EU SILC 2006, EU-SILC 2007 or EU-SILC 2008.
Three sets of longitudinal weights have been calculated for the persons in the participating panels in the relevant period according to the EUROSTAT document EU-SILC 065. These weights are RB062 (two years), RB063 (three years) and RB064 (four years). Common starting point of these longitudinal weights is the base weight RB060.

2.1.8.1   Design factor

The design factor (or design effect) expresses the loss in precision due to the actual sampling design, as compared to a single random sampling (SRS) design. As such, it plays an important role in determining the required sample size. The design factor can be calculated as the ratio of the variance (of a particular estimator), obtained under the actual design, to the variance obtained by SRS. Here, the design factor for the total at-risk-of-poverty rate is presented. The calculation of the design factor proceeds as follows. The variance obtained under the actual design is found by squaring the corresponding standard error listed in table 2.6 (see section 2.2.1). Next, in order to compute the variance that would have been obtained from a single random sample, a resampling method is used to simulate such a sample from the actual sample file. The simulated single random sample is subsequently used to infer the SRS variance, following the same strategy as outlined in section 2.2.1. With the thus found variance, the resulting design factor for the at-risk-of-poverty rate was 1.24 (based on EU-SILC 2006 cross-sectional data).

The design factor calculated here is in reasonable agreement with a preliminary estimate of the design factor, on the basis of which the total sample size was chosen. Calculating backwards, the effective longitudinal sample size is 6,649/1.24 = 5,362 households. This figure amply meets the requirement by the EU-SILC Regulation, which stipulates a minimum longitudinal effective sample size of 5000 households for any pair of consecutive years.

2.1.8.2   Non-response Adjustments-first wave

Non-response adjustments are necessary because of the bias introduced by selective non-response on the household level. Selective non response affects the inclusion probabilities of the sampling units. Ideally the inclusion probability can be calculated by multiplying the inclusion probabilities of the sampling design with the exact response probabilities. Unfortunately, in practice these response probabilities are unknown and some kind of approximation has to be made. The method of logistic regression was adopted to approximate the response probabilities for the new rotational group. The response probabilities were modelled by the explanatory variables age, degree of urbanisation, type of household, and labour force status.

2.1.8.3   Adjustments to external data- first wave 2006

For the new rotational group adjustments to external data were made to calculate the base weights. The calibration was performed on household and personal level using linear consistent weighting, so that individuals within the household have identical weights.

The following variables were included in the calibration scheme:

- Household size : 1 household member, 2 household members 3 household members, 4 or more household members
- Sex:
- Age class : 0 – 15 , 16 – 19, 20 – 24 , 25 – 29, 30 – 34 , 35 – 39, 40 – 44, 45 – 49, 50 – 54, 55 – 59 , 60 – 64 , 65 – 69 , 70 – 74 , 75 years or older.

2.1.8.4. Final longitudinal weights – first wave 2006
.
In the first wave, the personal longitudinal base weights (RB060) are the weights resulting from the design-weights after non-response adjustment and the calibration.

2.1.8.5 Non-response adjustments – subsequent waves

For the second, third and fourth wave the initial base weights of year t +1 are based on the cross-sectional weights (RB050) of year t. These initial weights have been adjusted for non-response and scaled to the population in scope. For the subsequent waves a proper model to determine the response probabilities could not be fitted using logistic regression. Therefore the response probabilities were considered equal for all persons in the response.

2.1.8.6 Adjustments to external data- subsequent waves

For each separate rotational group and each wave, the sum of the weights RB060 should be equal to the size of the longitudinal population in scope.

The following variables were included in the calibration scheme:

- Household size : 1 household member, 2 household members 3 household members, 4 or more household members
- Sex:
- Age class : 0 – 15 , 16 – 19, 20 – 24 , 25 – 29, 30 – 34 , 35 – 39, 40 – 44, 45 – 49, 50 – 54, 55 – 59 , 60 – 64 , 65 – 69 , 70 – 74 , 75 years or older.

With respect to rotational group R1' (first wave in 2006) the sum of the base weights in 2009 is equal to the size of the longitudinal population in scope from 2006 to 2009. Household Members with RB110 = 3, 5, 6 or 7 (moved into from outside sample, moved out, died or not in register) have a zero weight and members with RB110=4 (newly born) received the weight of their mother.

2.1.8.7. Final longitudinal weights- subsequent waves

For the 2009-operation, three sets of longitudinal weights have been calculated for the persons in the participating panels in the relevant period. These weights are rb062 (two years), rb063 (three years) and rb064 (four years). These two-, three- and four-year duration longitudinal personal weights have only values for 2009 as this correspond to the last wave in the file. With respect to RB062, the sum of the weights, all rotational groups together, is equal to the size of the longitudinal population of individuals in scope in 2008 and 2009. Concerning RB064 this sum is equal to the size of the longitudinal population 2006-2009. Individuals in scope are the ones with RB110 = 1 or 2. Members with RB110 = 3, 4, 5, 6 or 7 (moved into from outside sample, newly born, moved out, died or not in register) have a zero weight.

The basis for the weight DB090 in the longitudinal files is the weight DB090 from the cross-sectional files. However, these weights are summing to a total less than the household population, because the longitudinal file consists of only three rotational groups (R1', R2' and R3'). Therefore, variable DB090 had to be expanded with a scale factor. With this scale factor the sum of the weights for the three rotational groups together is equal to the cross-sectional household population size.

Table 2.2: Household weight in longitudinal file 2006-2009

	2006	2007	2008	2009
Accepted household interviews (R1', R2', R3')	2,339	5,607	8,066	6,649
Sum of cross-sectional weights	**2,357,686**	**4,221,351**	**7,020,381**	**5,034,230**
Number of households in population	7,146,088	7,190,543	7,242,202	7,312,579
Scale factor	3.03	1.70	1.03	1.45

2.1.8.8. Final household cross-sectional weight

Final cross-sectional weights were obtained by a calibration of the joint cross-sectional and longitudinal sample. Adjustments made by calibration schemes in general improve the accuracy of the data (mean square error). Three good reasons for using calibration schemes are: 1) the estimates of variables that are used in the calibration scheme are made consistent with those of more reliable sources. 2) the standard error of the estimates is reduced if the calibration variables correlate with target variables. 3) non-response bias is reduced if the calibration variables correlate with both target variables and response probabilities.

Two external data sources were used in the calibration procedure:
1. the Population Register (GBA), and
2. the register on income data based on integral data from the tax authorities.

The adjustments were made on the basis of the base weights: the product of the design weights with the inverse of the response probabilities (non-response weights). The calibration was performed on household and personal level using linear consistent weighting, so that individuals within the household have identical weights equal to the household weight. The set of variables used for calibration includes the smaller subset suggested by Eurostat in document EU-SILC 065/04. Additional calibration variables that correlate strongly with the target variables were added: income data and data on tenure status from the income register. The following variables were included in the calibration scheme:
- sex,
- age in years, 0,1,2,3,4 thru 84 and 85 years and over,
- household size: four categories (1, 2, 3, 4 and more household members),
- region: 12 categories, one for each of the provinces (nuts 2),
- tenure status, in two classifications (owner, tenant)
- equivalized disposable income (CBS-definition) in deciles
- main source of income (employee, self-employed, unemployed, social assistance, disabled, retired aged under 65, retired aged 65 years or older, student, no income).
- low income category, in three classifications (non target population, low income and other income).
- at-risk of poverty-rate (based on Income Panel Survey, national definition)

Taking into account consistency requirements and the correlation of weighting terms with important target variables (Laeken indicators), the following weighting terms were constructed:

weighting model terms at household level:
- household size,
- region (nuts 2),
- tenure status
- low income category.

weighting model terms at personal level:
- sex x age,
- equivalized income (decile groups),
- main source of income
- at-risk-of poverty-rate IPS

The household cross-sectional weight DB090 and the personal cross-sectional weight RB050 are the direct result of the linear consistent weighting procedure. PB040 equals this weight for persons of 16 years and older. PB040 equals 0 for people younger than 16 years.

Finally, the cross-sectional weights for the selected respondent are determined by adjusting the weight PB040 for the probability with which this respondent is selected within the household. This probability is equal for all persons that are older than 16. This probability is four times as large for persons that are exactly 16 years.

2.1.8.9 Substitutions

Not applicable.

2.2   Sampling errors

2.2.1. Standard errors and effective sample size

The subsequent tables present means, number of observations and standard errors for the cross sectional component 2009 and for each wave of the longitudinal component. The standard errors have been calculated with the use of the software package Bascula which has been developed by the methodology department at Statistics Netherlands. Using Bascula one can calculate (weighted) totals, means, ratios and the standard errors of target variables for a variety of sampling designs and weighting models.

Table 2.3: Mean, number of Observations, and standard errors for household income components EU-SILC cross-sectional 2009

Gross income components at household level	Mean (euro)	N	Standard Error
Total household gross income (HY010)	53,051	9,728	260
Total disposable household income (HY020)	35,008	9,728	160
Total disposable household income before social transfers other than old age and survivors' benefits (HY022)	32,064	9,728	163
Total disposable household income before social transfers including old age and survivors' benefits (HY023)	25,597	9,728	168
Gross income components at household level			
Imputed Rent (HY030G)	2,455	6,882	15
Income from rental of property or land (HY040G)	8,283	368	1,126
Family/child related allowances (HY050G)	2,063	3,256	15
Social exclusion not elsewhere classified (HY060G)	5,277	1,449	179
Housing allowances (HY070G)	1,754	946	34
Regular inter-household cash transfer received (HY080G)	3,939	639	293
Interest, dividends, profit from capital investments (HY090G)	2,461	8,597	153
Interest repayments on mortgage (HY100G)	8,229	6,012	140
Income received by people aged under 16 (HY110G)	813	187	172
Regular taxes on wealth (HY120G)	-	-	-
Regular inter-household cash transfer paid (HY130G)	4,387	1,203	333
Tax on income and social contributions (HY140G)	17,558	9,727	130

Table 2.4: Mean, number of observations, and standard errors for personal income components, EU-SILC cross-sectional 2009

Gross income components at personal level	Mean (euro)	N	Standard Error
Employee cash or near cash income (PY010G)	29,580	12,520	197
Non-cash employee income (PY020G) Company Car (PY021G)	5,614	990	126
Contributions to individual private pension plans (PY035G)	2,410	2,568	120
Cash benefits or losses from self-employment (PY050G)	19,681	1,887	1,017
Value of goods produced for own-consumption (PY070G)	-	-	-
Pension from individual private plans (PY080G)	10,014	85	1,152
Unemployment benefits (PY090G)	8,339	542	380
Old-age benefits (PY100G)	18,416	4,274	242
Survivor's benefits (PY110G)	10,880	148	433
Sickness benefits (PY120G)	3,596	243	438
Disability benefits (PY130G)	13,719	828	462
Education-related allowances (PY140G)	2,742	1,015	80

Table 2.5: Mean, number of observations, and standard error for the equivalized disposable income, cross-sectional 2009).

Equivalized disposable income	Mean	Number of Observations	Standard Error
Population by household size			
1 household member	19,128	2,660	269
2 household members	25,608	6,902	326
3 household members	23,461	3,605	331
4 and more household members	21,701	10,358	229
Population by age groups			
<25	20,605	7,333	157
25-34	22,693	2,411	296
35-44	23,483	3,727	247
45-54	25,639	3,815	387
55-64	25,583	3,355	477
65+	20,845	2,884	345
Population by sex			
Male	23,044	11,571	118
Female	22,420	11,954	122
Total	22,730	23,525	98

Table 2.6: Mean, number of observations, and standard errors for income components and equivalized disposable income, rotational group R1', 2007-2009.

	2009			2008			2007		
	Mean	Number of observations	standard error	Mean	Number of observations	standard error	Mean	Number of observations	standard error
HY010	53,568	1,286	998	51,191	1,552	857	50,158	1,876	1,007
HY020	35,642	1,286	575	34,463	1,552	533	34,391	1,876	739
HY022	32,669	1,286	644	31,861	1,552	572	31,774	1,876	765
HY023	25,913	1,286	619	25,328	1,552	558	25,608	1,876	748
HY030G	2,429	949	45	2,407	1,104	39	2,489	1,285	35
HY040G	9,926	41	2,389	12,865	51	5,606	12,894	50	3,631
HY050G	2,095	416	53	1,785	494	37	1,698	599	28
HY060G	5,025	203	776	6,764	93	1,064	6,545	117	748
HY070G	1,724	119	94	1,659	156	68	1,705	170	71
HY080G	4,507	71	991	4,026	81	706	4,851	104	506
HY090G	2,159	1,157	391	1,845	1,400	354	2,993	1,643	692
HY100G	8,038	820	290	7,598	992	220	7,293	1,156	180
HY110G	496	32	134	657	25	151	622	26	136
HY130G	3,237	140	321	5,203	160	933	4,306	240	427
HY140G	17,645	1,286	474	16,257	1,552	369	15,260	1,876	358
PY010G	29,989	1,645	675	29,324	1,952	518	28,061	2,321	523
PY021G	6,114	137	427	5,375	166	316	5,148	202	439
PY030G	4,993	1,696	132	4,925	2,019	100	5,814	2,548	107
PY035G	2,363	343	251	2,457	422	245	2,190	504	188
PY050G	17,774	255	2,236	16,760	311	1,711	16,378	338	2,313
PY080G	13,037	12	2,970	10,891	15	2,509	10,368	18	2,897
PY090G	5,881	81	863	7,032	107	749	7,958	160	738
PY100G	19,115	595	681	18659	653	495	18,138	772	503
PY110G	10,055	17	1,410	8,972	27	926	9,678	31	741
PY120G	2,844	29	764	4596	45	1,090	3,196	53	672
PY130G	13,697	122	1,372	14245	142	789	13,127	175	694
PY140G	2,549	121	233	3189	136	206	3,078	184	163
Equivalized disposable income									
Population by household size									
1 household member	20,843	330	654	19,396	423	566	21,158	509	1,477
2 household members	24,494	959	826	24,245	1,074	629	23,790	1,334	649
3 household members	24,258	446	1,084	22,659	565	802	22,744	741	1,151
4 and more members	22,020	1,405	608	21,823	1,636	852	20,690	1,926	682
Population by age groups									
<25	21,174	950	493	20,549	1,119	681	19,806	1,414	539
25-34	22,533	278	939	23,362	355	681	23,160	459	490
35-44	23,937	470	707	23,838	577	863	22,608	740	605
45-54	25,780	569	825	24,365	726	613	23,632	766	725
55-64	24,949	449	988	23,971	456	875	27,133	587	2,083
65+	21,019	424	912	19,574	465	499	18,590	544	385
Population by sex									
Male	23,445	1,548	441	22,872	1,820	408	22,435	2,223	475
Female	22,442	1,592	377	21,722	1,878	434	21,581	2,287	538
Total	22,938	3,140	373	22,289	3,698	395	22,004	4,510	435

2.3   Non-sampling errors

2.3.1   *Sampling frame and coverage errors*

As mentioned in paragraph 2.1.1, the sampling frame of addresses is constructed from the Population Register. First a complete list of addresses is made and then divided into 10 disjoint groups: A0, A1, A2 …, A9. Each of these subsets contains 10% of all the addresses in the Population Register. Subset A0 is used as an address sampling frame for the years 2000, 2010, 2020, …, subset A1 is used as an address sampling frame for the years 2001, 2011, and so on. With this kind of approach the sampling frames of ten subsequent years are disjoint and addresses that are contacted within one particular year will not be part of another address survey sample for the next nine years. This approach is in compliancy with the policy of Statistics Netherlands to reduce respondent burden in all surveys. Finally, additional information on the type of address and number of postal delivery points is added to the sampling frame using data from the Geographical Municipal Registration (in Dutch: Geografisch BasisRegister – GBR). The result is a set of disjoint sampling frames (one for each year) with address information and personal information of all individuals that are registered in a Dutch municipality.

Each year in September the sampling frames for the next year are constructed. The sampling frame of addresses is updated monthly for changes related to births, deaths, migration, new addresses, and vacancies. Also taken into account are changes in municipality boundaries and postal codes. At the date of sample drawing the entries of the sampling frame are therefore practically equal to those in the Population Register (GBA). As the fieldwork period starts six weeks later, coverage errors may occur: during the six weeks between drawing and application of the sample new addresses will be established and some addresses have become vacant or have been demolished.

Institutional addresses are removed after drawing the sample by comparing the sample addresses with entries in the register of institutional addresses. This register is updated once a year, so a small number of over-coverage errors are to be expected.

2.3.2   *Measurement and processing errors*

Measurement errors originate from four basic sources:
   (a) the questionnaire (effects of the design, content and wording);
   (b) the data collection method (effects of the modes of interviewing);
   (c) the interviewer (effects of the interviewer on the response to a question including errors of the interviewer);
   (d) the respondents (effects of the respondent on the interpretation of items).

Statistics Netherlands implemented a number of measures to reduce such errors.

- put in specialised expertise in developing questionnaires;
- routings in the questionnaires to provoke only the relevant questions for the respondent;
- cognitive laboratory experiments with focus groups and depth interviewing.
- there is an opportunity to make remarks in the questionnaire;
- evaluations of the questionnaire
- a stable automation system of data communication and production;
- monitoring system;
- each record contains interview accounts as well as interview data;

- extended interviewer instructions and regularly refreshing courses on basic skills and on EU-SILC;
- Interviewer manual;

In a first step in 2002 part of the EU-SILC questionnaire has been tested extensively in a pre-test and a field-test (Snijkers, Beukenhorst and Huynen, 2002).

Statistics Netherlands uses the CATI–method for the EU-SILC interview. The questionnaires were programmed in Blaise with several data entry and coding controls to reduce processing errors. Finally the EU-SILC files were transformed into Eurostats' standard format and tested using the checking programs developed by Eurostat.

2.3.3 Non-response errors

2.3.3.1 Achieved sample size

In 2005 a new sample was constructed and divided in four rotational groups. In table 2.11a it is shown that the four groups differ in size to compensate for panel attrition. The first group did only participate for one year (purely cross-sectional), the second for two years, the third for three years and the fourth for four years. Consequently the sample size for the first group (R1) was smaller than the sample size for the second group (R2), followed by the third (R3) and the fourth group (R4). The first group has been replaced by a new group R1' in EU-SILC 2006 (tabel 2.9b). Group R2' consists of sample persons who were drawn in 2007. Sample persons in group R3' entered the EU-SILC survey in 2008.

Table 2.7a: Sample Size and accepted Interviews EU-SILC 2006

	Total	R1'	R2'	R3'	R4
Persons 16 years and older	17,392	4,395	2,082	4,522	6,393
Number of sample persons	8,986	2,399	1,051	2,311	3,285
Number of accepted personal questionnaires	17,392	4,395	2,082	4,522	6,393
Accepted household interviews	8,986	2,339	1,051	2,311	3,285

Table 2.7b: Sample Size and accepted Interviews EU-SILC 2007

	Total	R1'	R2'	R3	R4
Persons 16 years and older	19,623	3,555	6,979	3,736	5,353
Number of sample persons	10,219	1,876	3,731	1,909	2,703
Number of accepted personal questionnaires	19,623	3,555	6,979	3,736	5,353
Accepted household interviews	10,219	1,876	3,731	1,909	2,703

Table 2.7c: Sample Size and accepted Interviews EU-SILC 2008

	Total	R1'	R2'	R3'	R4
Persons 16 years and older	19,519	2,957	5,437	6,614	4,511
Number of sample persons	10,337	1,552	2,893	3,621	2,271
Number of accepted personal questionnaires	19,519	2,957	5,437	6,614	4,511
Accepted household interviews	10,337	1,552	2,893	3,621	2,271

Table 2.7d: Sample Size and accepted Interviews EU-SILC 2009

	Total	R1'	R2'	R3'	R4'
Persons 16 years and older	18,254	2,467	4,622	5,339	5,826
Number of sample persons	9,728	1,286	2,449	2,914	3,079
Number of accepted personal questionnaires	18,254	2,467	4,622	5,339	5,826
Accepted household interviews	9,728	1,286	2,449	2,914	3,079

Table 2.8: accepted interviews , longitudinal sample EU-SILC 2006-2009

Longitudinal sample 2006-2009	2006	2007	2008	2009	Total
	n	n	n	n	n
DB135=1: Interview accepted for database	2,339	5,607	8,066	6,649	22,661
R1'	2,339	1,876	1,552	1,286	7,053
R2'	-	3,731	2,893	2,449	9,037
R3'	-	-	3,621	2,914	6,535
Personal interviews accepted	4,395	10,534	15,008	12,428	42,365
R1'	4,395	3,555	2,957	2,467	13,374
R2'	-	6,979	5,437	4,622	17,038
R3'	-	-	6,614	5,339	11,953

2.3.3.2   *Unit non-response*

Indicators of unit non-response in the first wave are included in table 2.19. The overall household non response rate was 22% in 2006 (group R').

Statistics Netherlands has focused on an increased use of register data instead of survey data in the production process of statistical information. Most of the present administrative Registers are provided with a unique link key. This is the so-called Social security and Fiscal number (SoFi-number). This SoFi-number is a personal identifier for every (registered) Dutch inhabitant and for those living abroad who receive an income from activities in the Netherlands and consequently have to pay tax over their earnings to the Dutch fiscal authorities. A few SoFi-numbers may be registered with incorrect values in the data-files, in which case linkage with other files is doomed to fail. However, in general, the percentage of matches is close to 100 percent.

In surveys records do not have a SoFi-number. This is also true for EU-SILC in which data are collected by interviews. For those records an alternative link key must be used, which is often built up by combining a set of identifying variables (address, sex and date of birth). This sort of link key will in most cases be successful in distinguishing people. However, it is not a 100 percent unique combination of identifiers. When linking the Population Register as well as the records from EU-SILC with this alternative key – and tolerating a variation between sources in at most one of the variables sex, year of birth, month of birth or day of birth – it reveals that 99 percent of the EU-SILC respondents can be linked to the Population register to obtain their SoFi-number. This is a very good result, though we should not exclude a danger of selectivity in this micro-linking process. EU-SILC respondents that could not be linked to the population register and their household members have been rejected from the database. Consequently, there's no partial unit non-response on income in the EU-SILC database. This is acceptable because the number of unlinked records is very low and the developing of imputation methods for these households is high. However, this method implies a loss of efficiency of the survey and the non response

bias is difficulty controllable. If the unlinked records belong to a selective subpopulation, then estimates based on the linked records may be biased, because they do not represent the total population. Analysis in the past has indicated that the young people, the 15–24 age group, show a lower linking rate in household sample surveys than other age groups. The explanation for this phenomenon is that they move more frequently and therefore they are often registered at the wrong address (e.g. students). However, in using a weighting model which includes age, any selectivity in the database has been solved accordingly.

Table 2.9: Indicators on Unit Non-response, first wave of the longitudinal component (rotational group R1', 2006)

	R1': 2006
Addresses successfully contacted	3,016
Valid addresses selected	3,053
RA address contact rate	0,99
Number of household interviews accepted	2,339
RH (proportion of completed household interviews accepted)	0,78
NRh (Household non-response rate) %	22.3
Personal interviews completed	4,395
Number of eligible individuals	4,395
Rp 1)	1
Individual non response rate (%)	0
Overall individual non-response (%)	22.3

1) proportion of complete interviews within the households accepted for the database

Table 2.10: Household response rate: Comparison of result codes between wave 2 (2007) and wave 1 (2006) , R1'

	Sample outcome in 2007								
	DB135=1	DB135=2	DB120=22	DB130=23	DB130=24	DB130=21	NC	DB120=23	Total
Sample outcome 2006									
R1'									
DB135=1	1,876	4	1	19	50	83	305	1	2,339
DB135=2		33		1	0	0	21		55
Total	**1,876**	**37**	**1**	**20**	**50**	**83**	**326**	**1**	**2,394**

Table 2.11: Response rates for households between wave 2 (2007) and wave 1 (2006), (R1')

	R1'
Wave response rate (%) (A/T-K)	78.4
Refusal rate (%) (G/T-K)	3.5
No contacted and others (%)	17.3
Longitudinal follow-up rate (%)	81.5
Achieved sample size ratio (%)	80.2

Table 2.12 : Household response rate: Comparison of result codes between wave 3 (2008) and wave 2 (2007), R1'

	Sample outcome in 2008								
	DB135=1	DB135=2	DB120=22	DB130=23	DB130=24	DB130=21	NC	DB120=23	Total
Sample outcome 2007									
DB135=1	1,552	4	4	15	39	106	155	1	1,876
DB135=2		2				1		1	4
DB120=22							1		1
DB130=22									
DB130=23							20		20
DB130=24							50		50
Total	**1,552**	**6**	**4**	**15**	**39**	**107**	**226**	**2**	**1,951**

Table 2.13: Response rates for households between wave 3 (2008) and wave 2 (2007), R1'

	R1'
Wave response rate (%) (A/T-K)	79.6
Refusal rate (%) (G/T-K)	5.5
No contacted and others (%)	14.1
Longitudinal follow-up rate (%)	82.7
Achieved sample size ratio (%)	82.7

Table 2.14 : Household response rate: Comparison of result codes between wave 4 (2009) and wave 3 (2008), R1'

| | **Sample outcome in 2009** | | | | | | | | |
	DB135=1	DB135=2	DB120=22	DB130=23	DB130=24	DB130=21	NC	DB120=23	Total
Sample outcome 2008									
DB135=1	1,286	4	2	18	34	91	115	2	1,552
DB135=2		3				1			4
DB120=22							4		4
DB130=22									0
DB130=23							15		15
DB130=24		2			1	1	35		39
Total	**1,286**	**9**	**2**	**18**	**35**	**93**	**169**	**2**	**1,614**

Table 2.15: Response rates for households between wave 4 (2009) and wave 3 (2008), R1'

	R1'
Wave response rate (%) (A/T-K)	79.8
Refusal rate (%) (G/T-K)	5.8
No contacted and others (%)	13.3
Longitudinal follow-up rate (%)	83.3
Achieved sample size ratio (%)	82.9

2.3.3.3 Distribution of households by household status (DB110), by record contact at address (DB120), by household questionnaire result (DB130) and by household interview acceptance (DB135), R1', 2006-2009

Table 2.16: Distibution by household status, by record at address, by household questionnaire result and by interview acceptance, 2006-2009, rorartional group R1'

	2006	2007	2008	2009
DB110 –Household status				
Household from previous wave				
At the same address at last interview		1,952	1,650	1,388
Entire household moved to a private household within the country		116	74	57
Household no longer in-scope				
Entire household moved to a collective household or institution		3	2	2
Household moved outside the country		2	1	
Entire household died		10	4	
Household does not contain sample person				
Non-contacted				
Unable to access		183		
Lost		128	220	167
New household for this wave				
Split-off household				
New address added to the sample this wave or first wave	3,062			
fusion				
Total	3,062	2,394	1,951	1,614
DB120 –Contact at address				
Address contacted	3,016	114	69	53
Address unable to access	37	1	4	2
Address does not exist	9	1	1	2
Total	3,062	116	74	57
DB130- Household questionnaire result				
Household questionnaire completed	2,394	1,913	1,558	1,295
Refusal to cooperate	304	83	107	93
Entire household temporary away				
Household unable to respond	79	20	15	18
Other reasons	239	50	39	35
Total	3,016	2.066	1,719	1,441
DB135- Household interview acceptance				
Interview accepted for database	2,339	1,876	1,552	1,286
Interview rejected	55	37	6	9

2.3.3.4 Distribution of persons by membership status (RB110)

Table: 2.17 Distribution of persons by membership status (RB110), 2006-2009, R1'

Current household members

	RB110=1	RB110=2	RB110=3	RB110=4	RB110=5	RB110=6
2006	5,596	0	0	0	0	0
2007	4,441	0	38	60	83	4
2008	3,698	0	30	47	61	0
2009	3,098	0	22	32	55	2

Because of the EU-SILC sampling design with the selected respondent in the Netherlands the table for personal interview response rates is not provided. Only the selected respondents, one sample person per household, are followed from wave t to t+1. The co-residents are not followed from wave to wave. They may belong to the household of the selected respondent in year t+1 or they (or the selected respondent) may have moved to another household. The table for personal interview response rates for the sample persons will be quite similar to the table with the household response rates.

2.3.3.5 Item non-response

As income data are based on register information, the income variables do not consist item non-response. However, some income components are not available in the tax registers because they are not taxable. This concerns the inter-household transfers and the income from rental of a property or land. These amounts are asked for in the EU-SILC questionnaire.

Table: 2.18 Item non-response household income components, EU-SILC 2009, cross-sectional

	households having received an amount		With full information		With non or partial information	
	count	%	count	%	count	%
HY010 Total household gross income	9,728	100	9,590	99	138	1
HY020 Total disposable household income	9,728	100	9,461	97	267	3
HY022 HY020 before transfers (except pensions)	9,728	100	9,461	97	267	3
HY023 HY020 before transfers including pensions	9,727	100	9,461	97	267	3
HY030G Imputed rent	6,882	71	6,882	71	-	-
HY040G Income from rental of a property or land	368	4	298	3	70	1
HY050G Family/Children related allowances	3,256	33	3,256	33	-	-
HY060G Social exclusion not elsewhere classified	1,499	15	1,499	15	-	-
HY070G Housing allowances	946	10	946	10	-	-
HY080G Regular inter-household cash tansfer received	639	7	570	6	69	1
HY090G Interest, dividends, profit from capital gain	8,597	88	8,597	88	-	-
HY100G Interest repayments on mortgage	6,012	62	6,012	62	-	-
HY110G Income received by people under 16	187	2	187	2	-	-
HY130G Regular inter-household cash transfer paid	1,203	12	1,065	11	138	1
HY140G Tax on income and social contibutions	9,727	100	9,727	100	-	-

Table: 2.18 Item non-response personal income components, EU-SILC 2009 cross-sectional

	Persons (16+) having received an amount		With full information		With non or partial information	
	count	%	count	%	count	%
PY010G Employee cash or near cash income	12,520	69	12,520	69	-	-
PY020G Non-Cash employee income	-	-	-	-	-	-
PY021G Company car	990	5	990	5	-	-
PY030G Employer's social insurance contribution	12,764	70	12,764	70		
PY035G Contributions to individual private pension plans	2,568	14	2,568	14	-	-
PY050G Cash benefits/losses from self-employment	1,887	10	1,887	10	-	-
PY080G Pension from individual private plans	85	0.5	85	0.5	-	-
PY090G Unemployment benefits	542	3	542	3	-	-
PY100G Old-age benefits	4,325	24	4,325	24	-	-
PY110G Survivor' benefits	148	1	148	1	-	-
PY120G Sickness benefits	243	1	243	1	-	-
PY130G Disability benefits	828	5	828	5	-	-
PY140G Education-related allowances	1,015	6	1,015	6	-	-

2.4   Mode of data collection

The response part of Labour Force Survey has been used as the sampling frame for EU-SILC. The income target variables have been derived from Registers. As a result, a substantial reduction of the questionnaire has been achieved. This enabled Statistics Netherlands to use Computer Assisted Telephone Interview (CATI) as interview mode.

Table 2.20: Distribution of RB245, RB250 and RB260, 2006-2009, R1'

	2006	2007	2008	2009
RB245-Respondent Status				
Household member aged 16 and over				
- selected respondent	2,339	1,876	1,552	1,286
-not selected respondent	2,056	1,679	1,405	1,181
RB250- data Status				
Information completed only from registers (11)	12	6	7	5
Information completed from both interview and registers (13)	4,383	3,549	2,950	2,462
Total	4,395	3,549	2,957	2,467
RB260 – Type of interview (selected respondent)				
CATI (3)	1,751	1,688	1,536	1,269
Proxy interview (5)	588	188	16	17

2.5 Imputation procedure

As income data are based on register information – except for the questions concerning some inter-household transfers (paid and received) and the income from rental of a property or land– the income variables do not consist of partial unit non-reponse or item non-response. If the household respondent refused to answer or did not know the amount of the inter-household transfers or the income from rental mean value imputation was used..

2.6 Imputed rent

For estimating the equivalent market rents in EU-SILC, the parameter estimates have been calculated based on another survey, the Survey on Household Expenditures. A regression model was applied on the estimates of market rents of owner-occupiers by real estate agents. This model includes the market value of the dwelling, region, level of urbanisation and household type. The total market rent is calculated by the National Account Statistics. Next the distribution of the market rent over the households is based on the results of the regression model.

2.7 Company car

The estimation of the value of 'company car' has been specified by the amount of benefit for which the recipient is assessed for tax purposes. The calculation of the employee income component 'company car' follows the rules of the tax authorities. As a general rule one has to add 25% of the value of the car to the income. Important are the original price of the company car and the intensity (kilometres) of private use.

3.   Comparability

This chapter reports on the differences between Eurostat definitions and the definitions Statistics Netherlands applied in EU-SILC 2009. It also reports in the impact of these differences on the comparability.

3.1 Basic concepts and definitions

(a) Reference population

The reference population of EU-SILC is all private households and their current members residing in the Netherlands at the time of data collection. The West Frisian Islands with the exception of Texel were excluded from the target population. This is also true for persons living in collective households and in institutions.

(b) Private household

No difference to the common definition.

(c) Household membership

There are some minor differences in the treatment of special categories like lodgers or people temporarily away (e.g. students). These people are only included as a household member if they are registered at the households' address.According to the EU-definitions resident boarders, lodgers and tenants should be included if they share expenses, have no private address elsewhere or their actual/intended duration of stay must be six months or more. Statistics Netherlands does not apply this limit of six months.

(d) Income reference period(s)

The income data of EU-SILC refer to the previous calendar. The income data were mainly collected from registers.

(e) The period for taxes on income and social insurance contributions

Taxes on income and social contributions are based on the 'income received' in the income reference year (accrual basis) and do not refer to the amounts actually paid in the income reference year.

(f) The reference period for taxes on wealth

There are no taxes on wealth in the Netherlands.

(g) The lag between the income reference period and current variables

The EU-SILC fieldwork period starts in June and ends at 30 September. Therefore the lag is at minimum 5 months and at maximum 9 months.

(h) The total duration of the data collection of the sample

The total duration of the data collection was approximately 4 months.

(i) Basic information on activity status during the income reference period

The monthly activity status during the income reference period is mainly based on register data on the main income source. The distinction between full-time and part-time work is based on the survey part of EU-SILC and the LFS.

3.2 Components of income

There are some differences in the definition of total gross income and disposable income based on the national definition and the SILC definition.

According to the Commission Regulation:

- *Interest paid on consumer debts is not considered as part of income definition in EU-SILC. In Statistics Netherlands' statistics on disposable household income interest payments on consumer debts are deducted to calculate the disposable income.*

- *Contributions to individual private pension plans (PY035) are classified under items which are not to be considered as income. In Statistics Netherlands' statistics on disposable household income, regular contributions to and benefits from private insurance schemes covering the risk of income loss are treated similarly as regular contributions to and benefits from (mandatory) social insurance and pension insurance schemes. This implies that contributions are deducted from and benefits are added to disposable income.*

3.2.1 Differences in definitions of the income target variables

Income variables with no differences from standard EU-SILC definitions are not mentioned.

Total household gross income and disposable income (HY010 and HY020);
The total household income (gross/disposable) has been computed without taking account the interest paid on mortgage, the imputed rent, the contibutions to and benefits from individual private pension plans. Subsequently the payable tax on income and social insurance contributions have been corrected to get the fictitious amounts that should have been paid if these components were not received/paid.

Total disposable household income before social transfers except old-age and survivor's benefits (HY022);
In order to calculate HY022 Statistics Netherlands calculated the taxable income without the income components:

PY090G + PY120G +PY130G + PY140G + HY050G +HY060G +HY070G.

Subsequently the payable tax on income and social insurance contributions have been corrected. The reason for this adaptation – the exclusion of these income components – is to calculate the fictitious amounts that should have been paid if such social transfers were not received.

Total disposable household income before social transfers including old-age and survivor's benefits (HY023);

Like HY022, but the income components PY100G and PY110G were also excluded.

Family/children-related allowances (HY050);
Maternity and parental leave benefits are not included in HY050 as those benefits cannot be separated from wages. These components are included in variable PY010.

Regular inter-household cash transfers received - (HY080);
Alimonies received from former spouse are available in the Tax Administration. Other transfers like payments received from parents living in a separate household (e.g. students) and child alimony are collected in the EU-SILC- interview.

Regular taxes on wealth (HY120);
There are no taxes on wealth in the Netherlands.

Regular inter-household cash transfers paid (HY130);
Maintenance allowances to former spouse were collected form the Tax Administration. Other transfers like child alimony are collected in the EU-SILC interview.

Total tax on income and social contribution (HY140);
When calculating disposable income some components were excluded (interest repayments on mortgage, imputed rent). Therefore, this variable refers to the fictitious amounts that have to be paid as if there were no (tax deductible) interest repayments on mortgage.

Gross employee cash income (PY010G);
Allowances for transport to or from work are not included in PY010. Severance and termination payments to compensate employees and redundancy payments (including lump-sum payments) are also included in PY010G. They are not included in PY090G (unemployment benefits).

Unemployment benefits (PY090G);
PY090 includes the vocational training allowance, i.e. payment by social security funds or public agencies to targeted groups of persons in the labour force who take part in training schemes intended to develop their potential for employment. Statistics Netherlands has no information available on benefit (in-kind) related to vocational training.

3.2.2 The source or procedure used for the collection of income variables

The variables concerning income, wealth and taxes were almost entirely collected from registers. The most important source is the Tax Administration. Student grants were obtained from the student loan company. Some components were imputed on the basis of information given in the questionnaire. For example, child benefits were calculated on the basis of the information about the number and age of children in the household.

3.2.3 The form in which income variables at component level have been obtained

All income data derived from registers are recorded gross at component level. 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 housing benefits and social assistance.

3.2.4 The method used for obtaining the income target variables in the required form (i.e. gross values).

Not applicable

3.3 Tracing rules

Statistics Netherlands followed the standard EU-SILC tracing rules.

4.   Coherence

Coherence refers to the comparison of target variables with external sources. However, external data for the four-year longitudinal sample are not available. However, for EU-SILC 2006, 2007, 2008 and 2009, cross-sectional data were compared to the Dutch Income Panel survey (IPS). The main aim of the Income Panel Survey (IPS) is to provide a detailed description of the composition and distribution of income of persons and households in the Netherlands. These comparisons can be found in the intermediate quality reports.