Statistics Finland

FINAL QUALITY REPORT

relating to the EU-SILC 2006 operation

Statistics Finland

Authors Pauli Ollila, Marie Reijo, Hannele Sauli, 03.02.2009

1 Common longitudinal EU indicators

Longitudinal EU indicators are not yet available for the operation years 2004 – 2006.

2 Accuracy

2.1 Sampling design

The sampling design of the Finnish EU-SILC survey, the collection year 2006, (also parallel with the design of the Finnish Income Distribution Survey [IDS]) is a *two-phase sampling design*. The copy of the population register some weeks before the end of the study year included 4,210,524 non-institutional persons aged 16 years or over. The type of the frame was based on the *domicile code*, i.e. very exact identification of all the possible places where people can live. The first digits of this code include regional information (municipality code). Systematic sampling of persons was carried out from that frame in order to get the basis for a master sample (50,000). After various checks and combinations we get 49,150 dwelling units with all their relevant members. The loss of 850 persons is due to the difference between the register which the selector of the master sample has and the final population register of the end of the study year. This final information (coming with the tax information to be connected to the master sample in order to create the strata, for example) is available after the master sample has been selected. At this point those who have died, moved permanently abroad or placed into an institution after the time point of the copy of the register and before the end of the year are excluded from the master sample. With this processing we correct the effect of the frame imperfection (not exactly describing the right time) in the sample.

This master sample of dwelling units is used for different sampling purposes, and one of them is the Income Distribution Survey. For that the master sample is stratified by socio-economic criteria, emphasising high-earners, farmers and entrepreneurs in the allocation. The sample size of the first wave is 7,500. The second wave of the IDS (5,797) is included in the set of households to be interviewed. The final definition of the structure of the household is done during the interview. The stratum is identified for these IDS waves separately in the variable DB050.

The **first wave of the EU-SILC longitudinal component selected in 2006** of size 2500 is selected randomly within strata from the first wave of the Income Distribution Survey (of size 7500) proportionally to the size of the IDS sample.

Referring to the description of the sampling design above it can be observed that
* **the Finnish cross-sectional data 2006 are based on a nationally representative probability sample of the population residing in private households** (non-institutionalised persons, two-phase sampling in both IDS waves),
* **all private households and all persons aged 16 and over within the household are eligible for the operation** (selection of persons, creation of dwelling units around persons and definition of households during the interviews),
* **representative probability samples are achieved both for households, which are the basic units of sampling, data collection and data analysis, and for individual persons in the target population** (selection of persons aged 16 and over from the register, creation of dwelling units around persons and definition of households during the interviews), and
* **the sampling frame and methods of sample selection ensure that every individual and household in the target population is assigned a known and non-zero probability of selection** (for every non-institutionalised person the probability of selection is identified and greater than zero).

2.1.1 Type of sampling

A two-phase stratified sampling design.

2.1.2 Sampling units

The sampling unit is a person. In the first phase persons are selected (target persons), in the second phase the target persons together with their dwelling units are selected.

2.1.3 Stratification criteria

The SILC data selection follows the sampling design of the Income Distribution Survey. The IDS stratification is conducted in the first-phase master sample containing dwelling units. The strata are created by using a socio-economic categorisation based on the register information available for the members at the time of sample selection. The stratification takes the highest earning person as the categorising person, but the entrepreneur need not be the highest earning one to define the household in the class of entrepreneurs. The income class division is used to allocate the sample more to high-earners. The stratification variable is **DB050**, containing values 1-13 for the first IDS wave and 14-26 for the second IDS wave, based on the dwelling units created around the selected persons.

Table 2.1 Stratification Criteria for the IDS

IDS Wave 1 (CY 2006)			IDS Wave 2 (CY 2005)		
Socio-economic categorisation of the household	Income Class	Stratum code	Socio-economic categorisation of the target person	Income Class	Stratum code
Wage earners	Lowest	1	Wage earners	Lowest	14
	2nd lowest	2		2nd lowest	15
	3rd lowest	3		3rd lowest	16
	Highest	4		Highest	17
Entrepreneurs	Lower	5	Entrepreneurs	Lower	18
	Higher	6		Higher	19
Farmers	Lower	7	Farmers	Lower	20
	Higher	8		Higher	21
Pensioners	Lower	9	Pensioners	Lower	22
	Higher	10		Higher	23
Others	Lower	11	Others	Lower	24
	Higher	12		Higher	25
No tax information	-	13	No tax information	-	26

2.1.4 Sample size and allocation criteria

One rotational group of size 2500 for the **longitudinal component of EU-SILC** was created from the selected sample of the Income Distribution Survey in 2006. In *Regulation 1177/2003 (Annex II)* there are minimum effective sample sizes for each country participating EU-SILC. This concept describes the sample size required under the sample design *simple random sampling. Regulation 1177/2003* Article 9 (paragraph 2) states that *"sample size for the longitudinal component refers, for any pair of consecutive years, to the number of households successfully interviewed in the first year in which all or at least a majority of the household members aged 16 or over are successfully interviewed in both years"*.

Minimum effective sample size for Finland; longitudinal sample, persons aged 16 or over: **5000.**

Finland uses registers for income and other data; thus a sample of persons (instead of a sample of households) is selected. *Regulation 1177/2003* Article 9 (paragraph 3) states that *"the minimum effective sample size in terms of the number of persons aged 16 or over to be interviewed in detail shall be taken as 75 % of the figures shown in columns 3 and 4 of the table in Annex II, for the cross-sectional and longitudinal components respectively"*.

Minimum effective sample size *(sample of persons)*, longitudinal sample, persons aged 16 or over: **0.75*5000 = 3750.** This concept is later denoted by n_{eff}.

Technical document on intermediate and final quality reports (EU-SILC 132/04, abbreviation TD) provides the following concepts of sample size to be defined (TD Section 2.1.4):

The achieved sample size *"depends on the efficiency of the sample design used (i.e. on the 'design effect')"*. The design effect term ($deft^2$) is *"the ratio of variance of a certain statistics) under the actual design, to that variance under a simple random sample of the same size"*. The reference statistic to be used in the design effect calculations is *at-risk-of-poverty-rate at national level (after social transfers)* (from TD Section 2.1). This design effect term for Finland based on the calculations from the Finnish Income Distribution Survey 2001, i.e. here $deft^2$ = **1.25**.

Minimum achieved sample size: n_{ach} = $deft^2 * n_{eff}$ = **1.25*3 750 ≈ 4 688.**

Thus the waves from 2 to 4 together should provide at least the achieved sample of size 4 688.

Taking the non-response into account, the sample to be selected must be larger in order to get the minimum achieved sample size. In general, the response rate for the first wave of EU-SILC (R_1) is assumed to be 0.76, and for the second (R_2), third (R_3) and fourth (R_4) wave we expect the rate to be 0.92. According to the wave structure which began in 2004 (earlier decision: **2 500** for each wave) we have the following situation in 2005:

Three waves which began in 2004 (ending in 2005, 2006 and 2007) should each have **0.76*2 500 = 1900** as a gross sample.

Actual longitudinal sample to be selected: n_{act} = **3*1 900 = 5 700**.

When taking the attrition of 8 % into account, we get the expected achieved longitudinal sample size.

Achieved sample size estimate: 0.92*5 700 = 5 244.

This sample size exceeds the minimum achieved sample size. Note that in the subsequent one year the attrition reduces still the achieved sample size of the fourth wave when compared with this sample size of year 2006.

Table 2.2 presents the relations between the longitudinal Income Distribution Survey (IDS) (areas with bold lines) and the wave structure of SILC (shaded). The assumptions are 76 % response for the first wave and 92 % response for other waves. Thus the sample sizes in Table 2.2 are anticipated. Table 2.3 includes the realised situation of the year 2006 SILC survey. The new sample in every stratum is distributed equally for three rotational groups.

Table 2.2   Structure of the longitudinal sample

	2004 1. year	2005 2. year	2006 3. year	2007 4. year	2008 5. year
Gross sample	5 700				
	2 500	1 900			
	2 500	1 900	1 748		
	2 500	1 900	1 748	1 608	
		5 000	3 800		
		2 500	1 900	1 748	1 608
			5 000	3 800	
			2 500	1 900	1 748
				5 000	3 800
				2 500	1 900
					5 000
					2 500
Total gross sample (IDS + SILC)	13 200	13 200	16 696	16 556	16 556
IDS gross sample	13 200	13 200	13 200	13 200	13 200
Achieved IDS sample	10 944	10 944	10 944	10 944	10 944
SILC waves 3 &4: gross sample		3 800	3 496	3 356	3 356
Achieved 3. & 4. sample		3 496	3 216	3 087	3 087
Longitudinal IDS + SILC gross sample (2, 3 &4)		5 700	9 196	9 056	9 056
Longitudinal SILC gross sample		5 700	5 472	5 365	5 365
Achieved SILC sample (longitudinal)		5 244	5 034	4 935	4 935
Minimum achieved sample size requirement		4 688	4 688	4 688	4 688

Table 2.3   Information concerning the longitudinal sample in 2006

Beginning of the panel (duration)	Sample		Sample excluding over-coverage		Accepted respondents	
	frequency	%	frequency	%	frequency	%
all	5 513	100.00	5 171	100.00	5 120	100.00
since 2004 (3 years)	1 818	32.98	1 733	33.51	1 719	33.57
since 2004 (4 years)	1 783	32.34	1 680	32.49	1 663	32.48
since 2005 (4 years)	1 912	34.68	1 758	34.00	1 738	33.95
New SILC wave	2 500	-	2 471	-	1 855	-

2.1.5 Sample selection schemes

The master sample of persons (1st phase) is selected with **systematic sampling** from the population *sorted by the domicile code*. The SILC/IDS sample of the first wave with dwelling units constructed around the target persons is selected from the **stratified** master sample with **simple random sampling without replacement** within every stratum and using *non-proportional allocation*. The IDS second wave respondents from the previous year were selected at that time in the same way. The **first wave of the EU-SILC longitudinal component selected in 2006** of size 2500 is selected randomly within strata from the first wave of the Income Distribution Survey (of size 7500) proportionally to the size of the IDS sample within strata.

2.1.6 Sample distribution over time

The income reference period is constant for all households and persons: the calendar year preceding the survey year. The reference population is defined as the population registered as resident in Finland on 31 December the year preceding the data collection year. Household composition is also dated on the same day.

The field work is usually started as early as possible in January. The interviewers start with the old panels. Quite exceptionally, field work for the 2006 EU-SILC operation was started already in December 2005 with the 3. wave interviews. The motivation for this was an attempt to relieve the exceptionally heavy work load anticipated in the Field Organisation in spring 2006. Table 2.4 reveals that the time pattern for the data collection has been brought forward as the panels have 'matured': in 2006 more than half of the longitudinal interviews had already been collected at the end of January. Households that are interviewed for the first time are contacted in February.

Table 2.4   Distribution of interviews over time, 2004, 2005 and 2006

Month	DB075 2 n	3 n	4 n	Total n	DB075 2 %	3 %	4 %	Total %
2006								
Dec.2005	160	162	0	322	9.3	9.7	0.0	6.3
I	1 226	1 183	266	2 675	71.3	71.1	15.3	52.2
II	333	318	1 015	1 666	19.4	19.1	58.4	32.5
III	0	0	457	457	0.0	0.0	26.3	9.0
Total	1 719	1 663	1 738	5 120	100.0	100.0	100.0	100.0
2005								
I	897	858	0	1 755	49.3	48.1	0.0	31.8
II	870	872	236	1 964	47.9	48.9	12.3	35.6
III	51	53	607	711	2.8	3.0	31.7	12.9
IV	0	0	527	527	0.0	0.0	27.6	9.6
V-VI	0	0	542	542	0.0	0.0	28.4	9.8
Total	1 818	1 783	1 912	5 513	100.0	100.0	100.0	100.0
2004								
II	214	207	0	421	10.9	10.7	0.0	10.8
III	948	901	0	1 849	48.2	46.4	0.0	47.3
IV	489	521	0	1 010	24.9	26.8	0.0	25.8
V	315	313	0	628	16.0	16.1	0.0	16.1
Total	1 966	1 942	0	3 908	100.0	100.0	0.0	100.0

*Rotational groups (DB075) no 2 and 3 started in 2004. Rotational group no 4 started in 2005.

2.1.7 Renewal of the sample: rotational groups

The Finnish cross-sectional SILC data collection year 2006 contains two groups based on the Income Distribution Survey: one is a new rotation group (1st IDS wave) and another is a set of responded households of the IDS of the previous year (2nd IDS wave). Note that the Finnish SILC design is not purely integrative from the SILC 2006 on, only the SILC waves began in 2005 (DB075=4) and in 2006 (DB075=1) are included in the cross-sectional SILC data together with non-SILC IDS collection from 2005 (DB075=6) and from 2006 (DB075=5). The two SILC waves began in 2004 (three and four years duration, DB075=2 and DB075=3) are not included in the cross-sectional SILC data, and they are conducted separately.

2.1.8 Weightings

2.1.8.1 Design factor

Deft= $\sqrt{1.25}$, see section 2.1.4.

2.1.8.2 Master sample

Separately calculated from the master samples CY 2006 (of size **50,000**) and 2005 (of size **100,000**) we got the population figures for the person selection, e.g., where $\pi_{a,person\,k}$ is the **inclusion probability of the selected person k** in the master sample. The **inclusion probabilities of the dwelling units** created around the selected persons in the master sample were $\pi_{ak} = \pi_{a,person\,k} n_{16+,dwelling\,of\,k}$. *Note that in this year and subsequent years concerning the EU-SILC in Finland the principles of weighting at this stage are parallel to the principles which are recommended by Eurostat, i.e. the first phase weight includes the master sample information in full.*

2.1.8.3 Income Distribution Survey sample and the new SILC wave sample

The **inclusion probabilities of two-phase sampling** (the effect of selecting the master sample and the IDS sample) were calculated, at the second phase based on the stratification (13 strata) of the master sample and the allocation used. *Note that the over-coverage is now included.* For those waves we separately calculated the inclusion probabilities $\pi_k^* = \pi_{ak}\pi_{k|s_a}$, where

$$\pi_{ak} = \pi_{a,person\,k} n_{16+,HH\,of\,k} = \frac{n_{s_a} n_{16+,HH\,of\,k}}{N}$$

and $\pi_{k|s_a} = n_h / N_{h,s_a}$ is the conditional inclusion probability at the second phase taking the stratification of the master sample into account. The **sample for the new SILC wave** is selected randomly within strata from the first wave of the Income Distribution Survey proportionally to the size of the IDS sample within strata. Thus the conditional inclusion probability $\pi_{k|s_a}$ is corrected with the term $n_{SILC,h} / n_h$. The **base weights** for the new wave were constructed as follows.

As the basis of calibration **the unit non-response was corrected** by $n_{SILC,sample,h} / n_{SILC,respondents,h}$ *in every stratum h* (interpreted as the inverse of the response probability in every stratum). The sum of these corrected weights calculated separately in the data of accepted 16+ persons in the HHs coincides with N_{16+}.

These weights containing a simple correction were used in **calibration (the raking method)** conducted with the macro CALMAR (applicable in SAS) for the accepted households (for the new SILC wave 1855). The calibration could be interpreted as **integrative**, i.e. both the household and the person levels were included in the process. The percentual marginal distributions and the statistics used in calibration are the following:

1) <u>Households:</u> province; type of municipality; HH size; sums of 15 different income variables. *The first three distributions of the households were obtained from the master sample, using weights for which a primary calibration (population register: 16+ persons and persons under 16 by region; gender*age class) was conducted. The income information comes from different registers.*

2) <u>Persons:</u> gender and age classes (0-4, 5-9, … , 80-84, 85+)

Table 2.5 Description of the Calibration Variables

Variable name	Description
Alue	Region (NUTS 3 level), Capital region separated
Ask8	Size of dwelling unit
Haastkur	Degree of urbanisation
Mibs01-Mibs18	Men 0-4, 5-9, 10-14, … , 80-84, 85-
Nibs01-Nibs18	Women 0-4, 5-9, 10-14, … , 80-84, 85-
Trplopti	Income 1: Cash or near cash employee income
Saipalk	Income 2: Income 1 > 0
Lelake	Income 3: Pensions
Tyotts	Income 4: Unemployment benefits 1
Perustur	Income 5: Unemployment benefits 2
Saityott	Income 6: Income 4 > 0
Elintul3	Income 7: Income from self-employment
Yhtytulo	Income 8: Capital income 1
Maattulo	Income 9: Income from agriculture
Omaitul2	Income 10: Income from property and forestry 1
Muupaao2	Income 11: Other capital income
Metstulo	Income 12: Income from forestry 2
Myvo	Income 13: Capital gains
Saielake	Income 14: Pensions > 0
Askorot	Mortgage interests

In addition, **2,435,000** was used as the **fixed number of households** in the process. The result of this calibration was the weight that produced exactly these margins when used in the summation of these variables in the data set containing accepted observations.

2.1.8.4 Weighting of the longitudinal SILC waves

The master sample and inclusion probabilities of the three longitudinal SILC waves (durations 3 and 4 years, at this moment three years) follow the same principles as presented in the previous section for the new SILC sample of 2006, but in that case concerning the collection year 2005. The **base weights of the waves** are also calculated in the same manner as described in the previous section, but applying the response data, frequencies and calibration marginals from 2005. The fixed number of households was then **2,435,000**.

The longitudinal weights require adjustments due to the changes appearing in time at the frame, household and person levels. Following the instructions of the Eurostat report *"EU-SILC Weighting Procedures - An Outline"* the weights were constructed for the longitudinal two-year SILC data as follows.

DB080: Household Design Weight. The original design weight from cross-sectional data is not applicable as such, because that weight includes the impact of the old wave of the Income Distribution Survey. This effect is removed by multiplying the cross-sectional design weight by 2 (see the intermediate quality report of 2006 for more details). The result is such that when adding up the weights within each panel of the longitudinal D file we get an approximation of the number of households in Finland.

DB090: Household Cross-sectional Weight, RB060: Personal Base Weight and **PB050: Personal Base Weight.** Here the principles of weighting are explained for RB060, but DB090 and PB050 are strictly applied in the same way (note that as a register country Finland has the same weight for all the members of the household in the first year). In the first year of the panel the base weights in the

panel are calibrated according to the same principles as for the cross-sectional weights. In the subsequent year the first year weight serves as the basis for further weighting taking into account the changes in the population in time. Instead of logit modelling we calibrate the current base weight to the exact population in sex & age groups, which has existed from the previous year as the technical report of weighting defines. This is possible by using the Total Income Distribution Statistics results in time, based on collected registers and linked by the identification codes common for the registers. After this the exact sex and age group distributions which have "survived" in time can be produced with the weights. For the SILC 2006 data including the previous years as well all the weights are adjusted in the same manner.

PB070: Personal Design Weight for Selected Respondent. The weight is calculated by multiplying longitudinal DB070 with the number of persons aged 16 or over in the household.

PB080: Personal Base Weight for Selected Respondent. The base weight for selected respondent from the first year is adjusted with the ratio between the current RB060 and base weight for RB060 from the first year, i.e.

$$\omega_2 = \omega_1^{(SB)} \left\{ \frac{\omega_2^{(RB)}}{\omega_1^{(RB)}} \right\}.$$

Then the weights are calibrated on gender and age (in single years) according to the distribution of the total sample aged 16+ weighted <u>differently</u>, namely by $\omega^{(RB)}$. The resulting weights for the completed individual interview sample are these post-calibration weights:

$$\omega_t \xrightarrow{\text{calibrated}} \omega_t^{(RB)}.$$

New persons not included in the first year are dealt with as follows. Children born to sample women receive the weight of the mother. Persons moving into sample households from other non-sample households in the population (co-residents) are given zero base weight.

The structure of the two-year longitudinal data requires weights also for the results of the previous years. These weights (DB080, DB090, RB060, PB050, PB070, PB080) come from that year, adjusted (when necessary) so that the sum of the weights describes the target population at that time.

2.2 Sampling errors

The Framework Regulation 1177/2003 states that

"The precision requirements concerning publication of the data collected in EU-SILC shall be expressed in terms of the <u>number of sample observations</u> on which the statistic is based and the <u>level of item non-response</u> (additional to total non-response at unit level). The Commission shall not publish an estimate if it is based on fewer than 20 sample observations, or if non-response for the item concerned exceeds 50 %. The data shall be published by the Commission with a flag if the estimate is based on 20-49 sample observations, or if non-response for the item concerned exceeds 20 % and is lower than or equal to 50 %. The data shall be published by the Commission in the normal way when based on 50 or more sample observations and the item non-response does not exceed 20 %.
All data publications shall include technical information for each Member State on the <u>effective sample size</u> as well as a <u>general indication of standard error of at least the main estimates</u>."

That is, the MSs have to calculate the **effective sample size** and the **standard errors of at least the main estimates**, which are defined as follows:

Effective sample size
The effective sample size used in the construction of <u>each common cross-sectional EU indicators based on the cross-sectional component</u> of EU-SILC, for the <u>equivalised disposable income</u> and for the <u>unadjusted gender pay gap</u>, will be provided.
Standard errors
The standard errors for the <u>common cross-sectional EU indicators based on the cross-sectional component</u> of EU-SILC, for the <u>equivalised disposable income</u> and for the <u>unadjusted gender pay gap</u>, will be provided.

Table 2.6 Effective Sample Sizes, Item Non-responses and Standard Errors of the Main Estimators for the Cross-sectional Data

Estimator	Accepted observations in general	Item non-response	Effective sample size	Standard error
Equivalised disposable income	28 039	0	28 039	60.51
At-risk-of-poverty rate after social transfers	28 039	0	28 039	0.413
Inequality of income distribution S80/S20 income quintile share ratio	28 039	0	28 039	0.057
Relative median at-risk-of-poverty gap	28 039	0	28 039	0.540
Dispersion around the risk-of-poverty threshold	28 039	0	28 039	0.144
At-risk-of-poverty rate before social transfers except old-age and survivors' benefits	28 039	0	28 039	0.432
At-risk-of-poverty rate before transfers including old-age and survivors' benefits	28 039	0	28 039	0.411
Inequality of income distribution: Gini coefficient	28 039	0	28 039	0.362

The *"gender pay gap"* comes from another source, not utilising the SILC data. Note that this table contains the calculations in general; when these indicators are classified with some variables (e.g. *main activity status* and *work intensity*), some item non-response may appear due to the classification variables.

The sampling design of the Finnish EU-SILC and the Finnish Income Distribution Survey is a two-phase design, with simple random sampling without replacement (1st phase) and stratified simple random sampling with unequal allocation emphasising some groups (2nd phase). The standard error calculations are conducted with the bootstrap method (10,000 replications). The idea is to estimate the standard error of the second phase by separately carrying out simple random sampling <u>with replacement</u> in every stratum with the original sample size of the stratum. *The calibration has been conducted in every replication, and the weights are an outcome of this process.* The variance to be used is simply the variance of the bootstrap estimator. In addition, in order to take the non-negligible sampling fraction into account the variance is multiplied by the finite population correction at the whole sample level, i.e. approximately 0.77. The standard error is the square root of the variance. The standard error of the equivalised disposable income is calculated with the software CLAN.

The variance estimation process includes some aspects of uncertainty. The non-response effect is not taken into account in variance estimation. The with-replacement nature of selection differs from the original selection, and the use of the finite population correction at the general level does not take the non-proportional allocation into account. This may yield obtaining a bit conservative standard error estimates.

According to *"Technical document on intermediate and final quality reports"* the final quality report should include means, numbers of observations and standard errors for some income components at both cross-sectional and longitudinal levels. The calculations are made with the software CLAN, and they take both the sampling design and the calibration into account. Note that the results of the rotational group breakdown are based on a separate calibration of each rotational group.

Table 2.7 Mean, number of observations and standard errors for components of income, cross-sectional data

Components of income	Variable name	Mean	Number of observations		Standard error
			Before imp.	After imp.	
Total household gross income	HY010	40 046.87		10 868	161.00
Total disposable household income	HY020	29 803.44		10 868	116.51
Total disposable household income, before social transfers other than old-age and survivors' benefits	HY022	25 617.72		10 868	119.84
Total disposable household income, before social transfers including old-age and survivors' benefits	HY023	21 097.40		10 868	112.67
Imputed rent	HY030G	3 257.16		10 868	32.73
Income from rental or property or land	HY040G	408.80		10 868	18.61
Family/children-related allowances	HY050G	1 023.51		10 868	17.64
Social exclusion payments not elsewhere classified	HY060G	151.62		10 868	7.58
Housing allowances	HY070G	365.85		10 868	9.83
Regular inter-household cash transfers received	HY080G	125.42		10 868	8.76
Interest, dividends, profit from capital investments in unincorporated businesses	HY090G	1 815.31		10 868	90.71
Interest paid on mortgages	HY100G	543.45		10 868	3.03
Income received by people aged under 16	HY110G	59.60		10 868	6.67
Regular taxes on wealth	HY120G	101.88		10 868	27.09
Regular inter-household transfers paid	HY130G	206.90		10 868	12.67
Tax on income and social insurance contributions	HY140G	9 934.65		10 868	71.48
Repayments/receipts for tax adjustments	HY135G				
Cash or near-cash employee income	PY010G	14 285.19		22 134	131.57
Non-cash employee income	PY020G	107.81		22 134	8.65
Employers' social insurance contributions	PY030G	.			.
Contributions to individual private plans	PY035G	134.48		22 134	6.45
Gross cash profits or losses from self -employment (incl. royalties)	PY050G	1 336.78		22 134	40.36
Value of goods produced for own consumption	PY070G	.			.
Pensions from individual private plans other than those covered under ESSPROS	PY080G	.			.
Unemployment benefits	PY090G	851.06		22 134	21.60
Old-age benefits	PY100G	3 154.86		22 134	55.55
Survivors' benefits	PY110G	92.11		22 134	11.57
Sickness benefits	PY120G	118.45		22 134	9.96
Disability benefits	PY130G	805.05		22 134	35.21
Education-related allowances	PY140G	135.24		22 134	8.50
Gross monthly earnings for employees	PY200G	.			.

* Households which have negative values or 0 values in the variable are counted as the households which have not received the income. Negative values of the certain gross income components in which they exist are counted in the variable HY010 on the total household gross income.

Table 2.8 Mean, number of observations and standard errors for equalised disposable income in different population groups, cross-sectional data

Equivalised disposable income	Mean	Number of observations		Standard error
		Before imp.	After imp.	
All	20 234.68		28 039	60.506
1 household member	15 531.47		2 398	451.166
2 household members	22 029.09		8 030	342.807
3 household members	22 062.75		5 547	437.310
4 household members or more	20 147.57		12 064	239.770
Age group <25 years	19 067.07		9 339	331.056
Age group 25-34 years	20 583.66		2 693	808.614
Age group 35-44 years	21 877.76		3 735	668.369
Age group 45-54 years	22 721.99		4 737	465.587
Age group 55-64 years	23 014.63		4 124	441.246
Age group 65- years	16 248.31		3 411	261.091
Male	20 638.45		14 279	253.556
Female	19 848.24		13 760	238.594

Note that the longitudinal data sets cannot be used for calculating the equalised disposable income within each longitudinal panel, when using the Eurostat indicator programs. The data structure is different and does not serve the programs.

Table 2.9 Mean, number of observations and standard errors for components of income, longitudinal DB075=2

Components of income	Variable name	Mean	Number of observations		Standard error
			Before imp.	After imp.	
Total household gross income	HY010	38 380.58		1 719	292.573
Total disposable household income	HY020	28 588.66		1 719	202.226
Total disposable household income, before social transfers other than old-age and survivors' benefits	HY022	24 772.61		1 719	222.498
Total disposable household income, before social transfers including old-age and survivors' benefits	HY023	19 950.85		1 719	208.421
Imputed rent	HY030G	3 249.24		1 719	67.057
Income from rental or property or land	HY040G	391.50		1 719	42.575
Family/children-related allowances	HY050G	944.69		1 719	37.291
Social exclusion payments not elsewhere classified	HY060G	151.90		1 719	20.626
Housing allowances	HY070G	306.01		1 719	24.178
Regular inter-household cash transfers received	HY080G	115.53		1 719	14.628
Interest, dividends, profit from capital investments in unincorporated businesses	HY090G	1 999.18		1 719	102.900
Interest paid on mortgages	HY100G	510.28		1 719	14.693
Income received by people aged under 16	HY110G	42.08		1 719	12.772
Regular taxes on wealth	HY120G	88.06		1 719	4.122
Regular inter-household transfers paid	HY130G	221.74		1 719	18.822
Tax on income and social insurance contributions	HY140G	9 482.12		1 719	110.552
Repayments/receipts for tax adjustments	HY135G				
Cash or near-cash employee income	PY010G	13 483.99		3 481	239.543
Non-cash employee income	PY020G	118.95		3 481	13.272
Employers' social insurance contributions	PY030G	.			.
Contributions to individual private plans	PY035G	122.15		3 481	11.557
Gross cash profits or losses from self -employment (incl. royalties)	PY050G	1 108.50		3 481	67.452
Value of goods produced for own consumption	PY070G	.			.
Pensions from individual private plans other than those covered under ESSPROS	PY080G	.			.
Unemployment benefits	PY090G	800.99		3 481	64.854
Old-age benefits	PY100G	3 406.84		3 481	107.335
Survivors' benefits	PY110G	63.84		3 481	16.777
Sickness benefits	PY120G	111.62		3 481	21.419
Disability benefits	PY130G	717.86		3 481	75.422
Education-related allowances	PY140G	109.36		3 481	12.174

Gross monthly earnings for employees	PY200G	.			.

* Households which have negative values or 0 values in the variable are counted as the households which have not received the income. Negative values of the certain gross income components in which they exist are counted in the variable HY010 on the total household gross income.

Table 2.10 Mean, number of observations and standard errors for components of income, longitudinal DB075=3

Components of income	Variable name	Mean	Number of observations		Standard error
			Before imp.	After imp.	
Total household gross income	HY010	37 874.00		1 663	346.828
Total disposable household income	HY020	28 124.09		1 663	245.138
Total disposable household income, before social transfers other than old-age and survivors' benefits	HY022	24 111.60		1 663	271.824
Total disposable household income, before social transfers including old-age and survivors' benefits	HY023	19 340.89		1 663	261.425
Imputed rent	HY030G	3 281.84		1 663	68.220
Income from rental or property or land	HY040G	282.41		1 663	30.385
Family/children-related allowances	HY050G	982.58		1 663	37.091
Social exclusion payments not elsewhere classified	HY060G	95.17		1 663	15.026
Housing allowances	HY070G	320.35		1 663	24.414
Regular inter-household cash transfers received	HY080G	149.01		1 663	18.042
Interest, dividends, profit from capital investments in unincorporated businesses	HY090G	1 534.35		1 663	149.014
Interest paid on mortgages	HY100G	525.15		1 663	19.077
Income received by people aged under 16	HY110G	37.00		1 663	10.317
Regular taxes on wealth	HY120G	92.34		1 663	8.233
Regular inter-household transfers paid	HY130G	232.53		1 663	30.686
Tax on income and social insurance contributions	HY140G	9 425.04		1 663	133.046
Repayments/receipts for tax adjustments	HY135G				
Cash or near-cash employee income	PY010G	13 243.46		3 318	249.804
Non-cash employee income	PY020G	119.36		3 318	12.578
Employers' social insurance contributions	PY030G	.			0.000
Contributions to individual private plans	PY035G	107.16		3 318	7.446
Gross cash profits or losses from self -employment (incl. royalties)	PY050G	1 250.06		3 318	98.803
Value of goods produced for own consumption	PY070G	.			.
Pensions from individual private plans other than those covered under ESSPROS	PY080G	.			.
Unemployment benefits	PY090G	837.67		3 318	61.036
Old-age benefits	PY100G	3 329.55		3 318	191.400
Survivors' benefits	PY110G	130.69		3 318	24.535
Sickness benefits	PY120G	111.28		3 318	20.544
Disability benefits	PY130G	792.56		3 318	86.746
Education-related allowances	PY140G	132.43		3 318	12.013
Gross monthly earnings for employees	PY200G	.			.

* Households which have negative values or 0 values in the variable are counted as the households which have not received the income. Negative values of the certain gross income components in which they exist are counted in the variable HY010 on the total household gross income.

Table 2.11 Mean, number of observations and standard errors for components of income, longitudinal DB075=4

Components of income	Variable name	Mean	Number of observations		Standard error
			Before imp.	After imp.	
Total household gross income	HY010	39 816.40		1 738	457.719
Total disposable household income	HY020	29 545.42		1 738	316.689
Total disposable household income, before social transfers other than old-age and survivors' benefits	HY022	25 514.42		1 738	335.814
Total disposable household income, before social transfers including old-age and survivors' benefits	HY023	20 880.47		1 738	328.537
Imputed rent	HY030G	3 352.46		1 738	66.817
Income from rental or property or land	HY040G	393.95		1 738	40.384
Family/children-related allowances	HY050G	1 049.13		1 738	42.517
Social exclusion payments not elsewhere classified	HY060G	132.26		1 738	17.392
Housing allowances	HY070G	308.83		1 738	21.669
Regular inter-household cash transfers received	HY080G	109.94		1 738	13.194
Interest, dividends, profit from capital investments in unincorporated businesses	HY090G	2 097.74		1 738	285.271
Interest paid on mortgages	HY100G	510.05		1 738	17.061
Income received by people aged under 16	HY110G	35.62		1 738	8.313
Regular taxes on wealth	HY120G	117.18		1 738	8.748
Regular inter-household transfers paid	HY130G	189.60		1 738	22.951
Tax on income and social insurance contributions	HY140G	9 964.20		1 738	161.125
Repayments/receipts for tax adjustments	HY135G				
Cash or near-cash employee income	PY010G	14 118.99		3 488	306.798
Non-cash employee income	PY020G	99.19		3 488	14..572
Employers' social insurance contributions	PY030G	.			.
Contributions to individual private plans	PY035G	147.04		3 488	22.767
Gross cash profits or losses from self -employment (incl. royalties)	PY050G	1 249.04		3 488	78.223
Value of goods produced for own consumption	PY070G	.			.
Pensions from individual private plans other than those covered under ESSPROS	PY080G	.			.
Unemployment benefits	PY090G	828.28		3 488	68.006
Old-age benefits	PY100G	3 266.45		3 488	164.678
Survivors' benefits	PY110G	79.70		3 488	40.262
Sickness benefits	PY120G	117.34		3 488	34.711
Disability benefits	PY130G	771.94		3 488	92.784
Education-related allowances	PY140G	119.65		3 488	14.000
Gross monthly earnings for employees	PY200G	.			.

* Households which have negative values or 0 values in the variable are counted as the households which have not received the income. Negative values of the certain gross income components in which they exist are counted in the variable HY010 on the total household gross income.

2.3 Non-sampling errors

2.3.1 Sampling frame and coverage errors

The target population is the set of elements about which information is wanted and parameter estimates required. The Commission Regulation on sampling and tracing rules states that "*The target population of EU-SILC shall be all private households and their current members residing in the territory of the Member State at the time of data collection. Persons living in collective households and in institutions are generally excluded from the target population. Small parts of the national territory amounting to no more than 2 % of the national population and the national territories listed in the Regulation may be excluded from EU-SILC, after agreement between the Member States concerned and the Commission (Eurostat).*" There is no register of households in Finland, so the selection is based on the population register and the creation of the households begins with the dwelling unit information available in the register.

2.3.1.1 Description of the sampling frame

The sample is drawn from the Population Information System maintained by Population Register Centre of Finland. The register is a continuously updated population register based on domicile. It is updated daily with information on population changes: births, deaths, migration, immigration and emigration, marriages, divorces, adoptions and changes of names. The Population Information System is a compilation of local registers kept up by population register districts.

The Population Information System (PIS) includes information on Finnish citizens and aliens permanently resident in Finland. It includes persons living in households, institutions, persons living temporarily abroad, and also homeless persons. Persons living in institutions, collective households or residential homes do not belong to the target population, but they are included in the PIS household population and have to be excluded from the master sample (see below).

Every person residing in Finland has a unique identification code and each dwelling has a domicile code Each person is registered in the municipality where he/she has a permanent place of residence. The domicile code is the link between a person and his/her permanent dwelling. Even the homeless are registered in municipal registers but without information of an address. The linkage between identification and domicile codes enables the pre-entry into the IDS-SILC questionnaire of all persons permanently registered in the dwelling unit-households before the interviewer contacts the household.

The copy of the population register some weeks before the end of the study year was the **sampling frame** for the selection of the new Income Distribution Survey (IDS) sample. After the separation of the persons placed in institutions and the homeless (a specific code identifies both cases), this frame included 4,210,524 persons aged 16 years or over. The sort of the frame was based on the domicile code, i.e. a very exact identification of all the possible places where persons can live. This code includes regional information at the beginning (municipality code). That frame is used for the **construction of the dwelling units for the master sample** as well. After various checks and combinations (e.g. excluding collective households, e.g. members of the same hall of residence as the target person) we get the dwelling units with all their relevant members for the selected master sample. Before the fieldwork begins the information of the second panel of the IDS and the changes after the selection of the sample are updated based on the register of the end of the year (then already available).

2.3.1.2 Information about the frame: reference period, updating actions, quality review actions

In general, the Population Information System of the Population Register Centre can be considered exhaustive and up-to-date as regards persons. Updating activities occur constantly. The Population Register Centre updates the 5. - 8. day of every month the official population figures in all municipalities in Finland.

The system is maintained by notifications of changes made by authorities. Maternity hospitals immediately report new-born children to local register offices. Deaths have to be reported at once either to a physician or to the police. They have to report the death to the Population Information System. The inhabitants are themselves responsible only of notification of changes of residence. Those who move or immigrate are expected to report to the local register office of the new place of residence on the change of address within one week of the move, specifying all the members of the family or household involved in the move.

Those emigrating should supply a notice of change of address in the country of entry. According to an agreement between the Nordic countries - which are the main destinations of migrants – the population register authorities of the country of entry inform the population register authorities of the country of exit. In the years when municipal elections are arranged (every 4th year), the population is corrected by around 1,000 persons, when emigrants whose emigration have been left unnoticed return notifications of voting.

A quality survey on the Population Information System is conducted yearly by means of a sample interview of 10,000 persons. From the EU-SILC point of view, reliability of its address information is of special relevance[1]. In the 2005 survey, assuming that all the unverifiable addresses were incorrect the final proportion of the correct addresses was 98.6 per cent.

The Population Information System has <u>no under-coverage</u> in any population groups. Asylum seekers and refugees are not included in the resident population until their permit of residence has been processed. The small <u>over-coverage</u> is a consequence of the necessity to draw the sample in good time before the actual date of defining the sample households (31 Dec.) and may also be related to register updates - delays in the notifications of emigration, moving to reside permanently in institutions or deaths.

The presence of the members of the households are checked in the interview. Persons who recently changed place of residence and/or household, new-borns, recently moved to institutions or died are the usual sources of non-correct register-based pre-entries in the IDS-SILC questionnaire.

2.3.2 Measurement and processing errors

Finland's SILC data is a combination of interviews and register information. In this chapter, the focus is mainly on description of collection and processing of the interviewed data. A short description of the register data processing is provided in chapter 2.3.2.3.

The interviews were carried out mostly by CATI (table 2.12). Of those interviewed by CATI for the first time, 99 per cent were interviewed by CATI also the second time. Of the 97 interviews conducted by CAPI in 2005, 64 per cent had switched into CATI mode in 2006.

---

[1] The EU-SILC collects variables PB130, PB140, PB150, PB190, PB210, PB220A and PB220B directly from the PIS. None of these information, however, have been checked in the PIS quality survey.

Table 2.12 Type of interview (%), longitudinal EU-SILC, 2004, 2005, 2006

	DB075		
	2	3	4
	%	%	%
2006			
CATI	99	99	98
CAPI	1	1	2
2005			
CATI	98	98	99
CAPI	2	2	1
2004			
CATI	96	96	-
CAPI	4	4	-

2.3.2.1 Questionnaire build-up, the testing procedures, interviewer training

Processing fieldwork tools

Feedback of the field work taken into consideration in the questionnaire build-up process
In 2005, a continuous yearly systematic feedback system was created: *the interviewers' feedback survey* is collected from all interviewers at the end of the project through a standard questionnaire. The interviewers are asked about the technical and substantial functioning of the questionnaires, how the letters and brochures motivate the respondents, whether the instructions are adequate, and specific remarks on each detail on the questionnaire. This feedback is utilised in the planning of the next year's tools.

The 2005 questionnaire was functioning badly according to the opinion of 20 per cent of the interviewers. The assesment of the 2006 questionnaire improved and was 7 per cent in the 2006 survey. Percentage of interviewers who felt that the questionnaire functioned badly as to the substance, fell from 26 in 2005 to 16 in 2006.

Questionnaire build-up and testing process in SILC2006

Finland's longitudinal SILC sample responds to the questionnaire that is identical for the first and second wave. The questionnaire includes questions needed to achieve both the cross-sectional and longitudinal target variables. On the third year, the questionnaire is changed into a shorter one which focuses only on the target variables requested for the longitudinal component.

Questionnaire build-up has its starting point in the previous year's questionnaire, feedback from the Cognitive Laboratory and field interviewers and feedback from the data editing process and users. At first the questionnaire for the 1.-2. interview is built up. After that, the 3.-4. interview questionnaire is built up. It contains only the questions needed to construct the longitudinal target variables. The general principle in the questionnaire build-up is a gradual integration process of the SILC to the IDS, and to avoid too many changes in the national IDS.

During the process of BLAISE programming (fall 2005), the questionnaire was table-tested by the team responsible for the IDS and EU-SILC. Eight persons were involved. In weekly meetings details of the questions were discussed, the focus being the parts of the questionnaire undergoing some change. In the end, a group of professional interviewers checked the questionnaire against their experience. Finally, the technical functioning of the questionnaire was tested in the interviewer organisation before they were sent to the field.

The testing procedure makes use of the BLAISE-programmed questionnaire. The real field situation is simulated by a test sample, actual households from the preceding year's data base. Thus the test questionnaire is prefilled with the information about the household composition and dates of birth. As in real field situation, the second and consequent panels have more information from previous interview entered into the questionnaires. The testers fill in the questionnaire, again and again, trying all combinations of imagined situations, and likely errors (to disclose signalling), too. They are asked to pay attention to
- spelling, language, formulations and conceptual correctness of the questions,
- proper functioning of the routings and
- adequacy of logical checks, signals and interviewing instructions on the screen.

Interviewer training

Statistics Finland's interviewer organisation employs about 160 field interviewers on a permanent work contract. They work mostly part-time. They are given basic training on interviewing and questionnaire standards and codes of practices when they start working. They collect most of Statistics Finland's survey data, for the Labour Force Survey, Household Budget Survey, Time Use Survey and Adult Literacy Survey, for example. In other words, they are experienced. Of them 127 were involved with EU-SILC interviews in spring 2006.

In 2004, all interviewers had a two days training exclusively for the EU-SILC survey. In 2005, all the interviewers participated the interviewers' training courses that took place in January. One-and-a-half-hour training on the changes of the EU-SILC 2005 questionnaire was given to all interviewers in connection with a general training day for each of five different areas of Finland. Before this, they had training material in the form of the CATI questionnaire and interviewer instructions and they were paid to study the material. The training was based on written material in 2006.

The changes on the questionnaire are introduced each year to the interviewers in a separate written report and, of course, in the instructions book. The instructions book is rewritten every year and it is also under constant development. The interviewers are paid to get acquainted with the material and practice with it.

Besides that, <u>newly recruited interviewers</u> are trained separately. In 2006, they had two day's training about the SILC. The training programme includes a lecture on the planning of the survey, including a description of Eurostat's process, legislation and future uses of the data, and Eurostat guidelines on data protection. Concern over international comparability is underlined. Instructions on the fundamental rules of central data collection are given and discussed, such as the definition of target population, household definition and its implementation in practice, different concepts and classifications of activity, especially labour market activities, child care questions, housing costs and mortgages. A major part of the training time is used on going through the videoed BLAISE questionnaire with the aid of three lecturers. The panel design and the future modules are described. Data transferring, data protection and other practicalities are also tutored.

During the whole fieldwork period, interviewers' information desk is open for them. They can ask for support from the IDS-SILC team. The interviewers, who are distributed all over the country, also have organised district meetings with each other to discuss professional matters.

2.3.2.2 Possible sources of measurement errors

Measurement errors stemming from
- difficulties in understanding complex questions on the telephone,
- difficulties in remembering complex life course events like the year's activities, day care changes, payments of many sorts, and

- difficulties in knowing/reporting another household member's activities have not been systematically surveyed. The 2004 questionnaire was evaluated – in principle, not empirically - in the Cognitive Laboratory from the above-mentioned points of view. The observations from this process are still paid attention to in the questionnaire build-up.

The potentials for error prevention are used extensively in BLAISE programming.
- Most relevant question-specific instructions are on the screen with the questions.
- Routings to avoid repetitive or irrelevant questions.
- Prefillings from the Population Register are used to help household construction.
- Prefillings from previous wave (occupation, NACE)
- Coherence is maintained by introducing logical checks to interconnected questions.
- Questions presuming numerical answers are given upper and lower limits where possible.
- Signals are pre-programmed to possible incoherent answers, to violations of numerical limits or to missing answers.
- The questionnaire is programmed to accommodate the mode of addressing the respondent depending on whether the selected person him/herself or another member of the household is responding (interviewing the selected respondent about himself: Did you… ; interviewing through a proxy respondent: Did N.N. …). This helps the interviewer and respondent to keep control of the member-specific data collection.

Of the many possible sources of measurement errors, the focus in this section is on errors due to *integration problems*, *questionnaire techniques* and *fieldwork problems*. The problems are presented as *possible sources of error*. The exact nature and size of error, if any, can only rarely be detected. The quality of register data is described in the chapters on comparability and coherence.

The use of proxy respondents

In Finland, the EU-SILC is designed on the selected respondent -model. Typically, only one person is interviewed. He/she gives all the information: household questionnaire and the personal questionnaires of the selected person and the other members of the household. The interviewers have been instructed to prefer interviewing the selected respondent if he is able to give information about the household economy, housing and the other household members' activity. Otherwise, a proxy respondent is interviewed. The proxy respondent is chosen by the interviewer.

The use of proxy respondents is a problematic choice because of the integration of the IDS and EU-SILC. In the EU-SILC, it is important to interview persons about their subjective evaluations (especially about health). Person-specific facts are also collected in the IDS, but these facts are of objective nature and can easily be reported by a household representative.

The consequences of the use of proxy are twofold: on the one hand, missing and/or poorer quality answers in questions focused on personal labour market variables or personal assesments of health; on the other hand, proxies are chosen to give informed and/or better quality answers in questions focusing on household income, mortgage and other loans, arrears in payments, housing variables, child care, and report uniformly about other household members' activities and incomes.

Interviewing more than one household member - both the selected person and a household respondent – is supported, but it rarely happens. Other members are allowed to be consulted during the interview if they are available. This option is often used.

The interviewers have traditionally been trained to find a household respondent in the earlier years when collecting the IDS data and they have been continuing this procedure. According to an estimate of the interviewers, about 85 per cent of their informants are those who have the best knowledge of the household's affairs. In case the selected person is aged less than 18 years, the

contact letter is also sent to his/her parents or guardians. The proxy use is slowly decreasing from 26 per cent of the selected respondents in 2004 to 22 per cent in 2006. (Table 2.13).

Table 2.13   Proxy interviews in the longitudinal component, 2004, 2005, 2006

Rotational group	Collection year 2004			Collection year 2005			Collection year 2006		
DB075 Start year / duration	Total	proxy	%	Total	proxy	%	Total	proxy	%
2 2004 / 3 years	1 966	545	27.7	1 818	446	24.5	1 719	404	23.5
3 2004 / 4 years	1 942	501	25.8	1 783	398	22.3	1 663	360	21.6
4 2005 / 4 years	-	-	-	1 912	441	23.6	1 738	341	19.6

Household figures. In this table, a proxy is a person other than the selected person who responded in the interview on behalf of the selected person.

Proxies are mostly persons responsible for the accommodation. The youngest selected persons under the age of 18 have most often been represented by a proxy respondent. Most of the proxy respondents are parents or spouses of the selected respondent. (More details in the intermediate reports 2004 - 2006.)

Fieldwork problems

<u>Mode of data collection.</u> According to interviewers' estimate, about half of the interviews are conducted through mobile phones and about 6 per cent of them outside home. The interviewers are allowed to change the mode into CAPI, in case the respondent has no phone or has an exceptionally large household. See chapter 2.4.

Telephone interviews in general are afflicted by a sense of rush. In large households, the interview is too long for telephone. Although the interview takes approximately half an hour on the first and second wave, and 15 minutes on the subsequent waves, it is a long time on a phone. According to feed-back from the interviewers, the questionnaires are hard to manage cognitively. Many questions require reminiscence and retrospection. This may have an effect on attrition, but to what extent is unknown. We have noticed a sudden increase of declines at the close of the first interview. The respondents ask the interviewer not to call again.

According to feed-back from the interviewers, the 2004 and 2005 questionnaires were rather complex and hard to manage, but they were more satisfied to the 2006 questionnaire.

<u>Refusals</u>. The share of sampled households who refuse co-operation with the interviewer slowly rises each year. See chapter 2.3.3.4.

Integration of the questionnaires of the national IDS and EU-SILC

The questionnaire for the first EU-SILC operation was built up using the national Income Distribution Survey 2002 BLAISE questionnaire that has been in use in its present form (with only slight modifications from year to year) since 1994. A major part of the questionnaire contents was shared with the national IDS and EU-SILC, but there were differences, too. A serious concern in the integration process is to preserve the national time-series without violating demands made to EU-SILC comparability. A stepwise integration strategy aims to achieve full integration in 2007.

Labour information in IDS and EU-SILC

Labour information is the most problematic area of integration. The basic concepts of main and second job differ in the IDS and EU-SILC. The reference periods for the activities and job-taking in the IDS and EU-SILC are not easily reconciled. The solution was <u>to reduce the number of reference</u>

<u>periods</u>. That was achieved in defining "current" to be included in the IRP. The consequent time gaps are reported in Section 5.

On the 3rd and 4th waves, national questions are deleted from the questionnaire. For continuity, we must, of course, use similar reference periods.

Table 2.14 Examples of labour information with different requirements in the IDS and EU-SILC

Concepts / Variables	Requirements		Solution
	IDS	**EU-SILC**	**Integrated**
			Current = Last December
Main job	Longest period of employment during the year or highest income	Current	If main job is different from current job, both are collected
Second job	The second longest period of employment during the year or second highest income	Current	If second job is different from current second job, both are collected
PL020	---	Current - 4 weeks	December
PL025	---	Current + 2 weeks	December
PL030	---	Current	December
PL040	Status in main job	Current	If main job is different from current job, both are collected
PL050	Occup. in main job	Current	If main job is different from current job, both are collected
PL070, PL072, PL080, PL085, PL087, PL090	Number of months for each activity - 12 categories - overlaps allowed	Number of months for each main activity - no overlaps allowed	Number of months and calendar of activities collected for all members 16+
PL110	NACE in main job	Current	If main job is different from current job, both are collected
PL140	Contract in main job	Current	If main job is different from current job, both are collected
PL180	---	since last interview	constructed through calendar of activities

Changes in the 2006 questionnaire

The CATI questionnaire is almost identical on the 1st and 2nd interviews for the cross-sectional and longitudinal components. On the 3rd and 4th interviews, the questionnaire only consists of the questions needed to construct the SILC target variables for the longitudinal component. In 2006 the questionnaire for the 3rd interview was needed for the first time. It was BLAISE-programmed using the questionnaire for the cross-section as a base. Basically it takes only omitting a large number of questions, but since the omissions disturb routings the programming had to be largely reconstructed and tested. In the end, we have an autonomous BLAISE questionnaire for longitudinal data collection.

In 2006, this questionnaire had to be built up twice, because the longitudinal data collection started in December. All the time references had to be suited for the time of interview and then changed after the turn of the year.

The interview duration on the 3rd interview fell into a half from what it was on the 2nd time, partly because of the omission of the national questions, partly because the ad hoc modules are not collected from 3rd and 4th interview.

Measurement failures due to questionnaire techniques: variable-specific problems

Income information or income-related information collected by interviews. Since the income data are mainly collected from registers, the questionnaire covers only those types of income that are not registered at all or on individual level (transfers between households, income from abroad, interest received, grants, non-taxed insurance compensations, strike pay). These variables cover a very small part of the total household income. The nature of questionnaire-collected income data is supplementary, important for some special groups such as single parents or students. These income items usually form only small fractions of target variables. These income items are collected similarly for the IDS and EU-SILC as the income definitions, reference periods and units of observation. Problems of the respondent's willingness to answer, perception and remembering constitute the source of measurement errors and, as a consequence, possible under-coverage of non-registered income data.

HB100, PB120 - Household and personal interview duration - separate measurement failed in 2004, 2005 and 2006. Only total duration of the interview is available (but not as target variables in the UDB). In the selected respondent -model there is less sense in collecting the duration of personal interview since actually there is no personal interview, one person - the selected respondent or the proxy respondent gives all information.

HS130 Lowest monthly income to make ends meet. The difficulty of this question for the respondent is well illustrated by the high item non-response. In the longitudinal data, the number of **missing answers** varies between 15 and 17 per cent of the cases.

PE030 Year when the highest level of education was attained. To collect the highest ISCED level attained (PE040), the register on examinations was utilised. In case the person had an examination In this register, the coverage of the examinations is good, but the year of passing the examination is often missing in the register - hence a **high number of missing values**. The missing values are concentrated on examinations passed before the year 1970 - the register was established in 1970. This is an example of fieldwork problems resulting from the intertwined use of register and interview data.

PH010 - PH030 Health questions: **item response rate is somewhat lower** than the overall response rate since the health questions were not allowed to be answered by a proxy respondent. In addition, in 2006 the item nonresponse was even lower due to a flaw in interviewer training.

We became aware only in 2006 in connection with a cognitive laboratory study that the questions are not formulated according to the regulation. **The scale used in PH010 is not in accordance with the regulation**, and the **formulation of PH030** differs slightly from the regulation. All the three years are in harmony with each other, but not the regulation.

On the 2004-2006 questionnaire the modalities of health questions PH010 and PH020 deviate from the formulation given in Doc65 in the following way:

Questionnaire:	Doc65:
PH010	
Do you find your present state of health as 1. good 2. rather good 3. fair 4. rather bad 5. bad	How is your health in general, is it 1. very good 2. good 3. fair 4. bad 5. very bad

PH030	
Has an illness, complaint or disability limited your working or daily activity in the past six months: 1. a lot 2. somewhat 3. not at all?	For at least the past six months (and for the present moment) , have you been limited by health problems: 1. limited to a great extent 2. limited to some extent 3. not at all limited?

2.3.2.3 Processing errors

The data collection and files construction process is shortly described in this chapter. No uncorrected errors have been detected in the processing.

Description of the data processing procedures

Fieldwork management and data reception. The interviewers collect the data and transmit them to the central unit. At Statistics Finland, there is a separate organisation, the Interviewers' Central Unit, to control, monitor and supervise the field work. The central unit transmits the fieldwork tools to the field and organises interviewer training at the beginning of the project, follows the fieldwork progress, and receives the output from the field, checks that all the sampled units are adequately processed and transmits the data to the IDS-SILC team. It also collects feed-back from the interviewers with a standardised questionnaire. All data contents processing takes place in the IDS-SILC team, either using the BLAISE system or SAS. Mainly the IDS and SILC data processing is integrated.

Checking and editing of the interview data. The BLAISE programming system already described above (Section 2.3.2.1) is a major data entry controller. However, there is still much processing to be done in the central unit. *Missing identification codes* are found out with the help of the Population Information System and added to the database. The checking process starts with the *interviewers' remarks* saved on the questionnaires. They comment whenever they feel that the coded answer does not reflect the individual real world. All comments are read and the need to edit the data is evaluated, and when necessary, entries are edited before transferring the data to the database. This work starts during the fieldwork period, usually in mid-February. All comments were processed before the end of June.

After the fieldwork period, the IDS-SILC team looks through *incomplete interviews* and makes a decision on the acceptance. Some of the received incomplete interviews are rejected. Since the register income data are nearly perfect, the acceptance decision is based on the sufficiency of the labour activities and housing information. In the later process, the discarded cases are treated as non-response, since a they are typically cases, where the interviewer finds that the respondent is unable to answer or the respondent refuses during the interview.

Next, checking against the register data is started as soon as the relevant register information is available. Occupation and NACE are processed through automatic coding. Some of the cases will remain open, and they are processed manually.

Activity months, occupation, NACE, housing costs and child care are checked against other information with special intensity. The checks include error lists generated by comparisons of interview and register data. Statistics Finland has access to administrative data on an individual level, which makes this data process especially useful. Great differences between different sources of information, if detected, are processed one by one. All variables, except variables where opinions are expressed, are checked: missing answers, denials and don't knows are checked

against other information. Clear mistakes are corrected. Missing values are completed whenever possible (e.g. missing dwelling rents are corrected with average rents per m^2 in the area, other missing housing cost information is completed with supporting information collected on the questionnaire). Illogical answers are straightened if possible. Outliers (considerably small or high values in numerical variables, e.g. inter-household transfers, housing costs) are detected and checked against other information.

Processing inconsistency in the integrated project. The 12 **IDS** variables on months of activity are heavily edited to comply with register data, especially with income data. That can be done, since in the IDS there are not too many connections between months of activity and other interviewed variables. As a result, some of the respondents' own answers are rejected and replaced with answers in coherence with their earnings. **Corresponding editing is not executed on the SILC variables** concerning categories of activity or inactivity, since that would destroy the coherence of the large set of other interviewed variables interconnected with activities. In other words, as a result of different editing, activity information in the IDS and SILC differs from each other. Months of activity (PL070, PL072, PL080, PL085, PL087, PL090) in the EU-SILC are, thus, subjective responses given by respondents, as defined in the EU-SILC document 065/04.

Database construction. Simultaneously with the checking process, a database is opened and variable formation begins. Interview-based and register-based variables construction is started. Interview-based variables are transferred from the questionnaires to the database. Variables that need constructing - ie. combined interview- and register information and complex questionnaire items – are added one by one into the database after all the checks have been made. In 2006 operation, the SILC data files for EUROSTAT were compiled from the data base by SAS after the IDS data were completed. The cross-sectional and the longitudinal target variables for year t are mostly programmed together and stored in the database. The longitudinal files of year t are compiled into SAS-files after the cross-sectional component of the year t+1 is completed.

Processing register data. Register data - that have been subscribed from the register authorities with a special procedure - arrive in electronic form to the Statistics Finland's data processing unit. In 2006, use was made of eleven registers. The incoming data are checked technically and contentually. Possible defects are notified to the authority in charge. They then transmit the corrected data. The registers cover all units - population, dwelling units, income receivers, etc. The data are linked to the sample persons and transmitted into the database of the IDS-SILC. The data are compared with available external data, i.e. those of the tax authority, pensions authority and other statistics. In this phase, the data are in their elementary form. Imputations are made using the hot-deck method (interest income) or the modelling/sratification method (imputed rent). Outliers are handled. Final weights are calculated. The SILC target variables are constructed only after all their elements have been checked in the IDS process.

Comparison of aggregates. Routines have been developed to compare the results on variable level with external sources such as the Labour Force Survey, National Accounts, wage statistics and statistics on different social transfers and taxation produced by the National Pensions Institute, National Board of Taxes and National Research and Development Centre for Welfare and Health. Standard comparisons ase routinely made each year. These comparisons also have an effect on error detection.

2.3.3 Non-response errors

2.3.3.1 Rotational groups

This section concentrates on non-response errors in the cross-sectional SILC data. **Many of the subsequent tables include the *rotational breakdown* as a requirement.** The second Finnish SILC data include the rotational group variable **DB075**, which is coded as follows:

1: Households included in the first wave of the Income Distribution Survey **and** in the longitudinal SILC panel beginning in the Collection Year 2006 (Survey Year 2005).
6: Households included in the first wave of the Income Distribution Survey **but not** in the longitudinal SILC panel beginning in the CY 2006 (SY 2005).
5: Households included in the second wave of the Income Distribution Survey **but not** in the longitudinal SILC panel beginning in the CY 2005 (SY 2004).
4: Households included in the second wave of the Income Distribution Survey **and** in the longitudinal SILC panel beginning in the CY 2005 (SY 2004).

The two-year longitudinal data in 2006:
2: Households included in the three-year panel of SILC (longitudinal DB075 = 2 , CY 2004, SY 2003).
3: Households included in the four-year panel of SILC (longitudinal DB075 = 3 , CY 2004, SY 2003).
4: Households included in the second wave of the Income Distribution Survey **and** in the the four year panel of SILC (longitudinal DB075 = 4).

For the three-year longitudinal data in 2006 only 2 and 3 apply.

The Finnish SILC design can be interpreted as *semi-rotational*, i.e. only a part of longitudinal rotational groups are included in the cross-sectional data. However, the forthcoming tables include both the cross-sectional rotational breakdown as well as the longitudinal study groups.

Table 2.15 Anticipated Sample Sizes and Wave Structure in Subsequent Years

Rotational group	2004 1. year	2005 2. year	2006 3. year	2007 4. year	2008 5. year	2009 6. year
1	5 700					
2	2 500	1 900				
3	2 500	1 900	1 748			
4	2 500	1 900	1 748	1 608		
5		5 000	3 800			
6		2 500	1 900	1 748	1 608	
1			5 000	3 800		
2			2 500	1 900	1 748	1 608
. . .						

Shaded area = longitudinal study, box with a bold line in a column = cross-sectional study.

2.3.3.2 Achieved sample size

Table 2.16 Interview Information

Rotational group	Number of households for which an interview is accepted for the database (DB135 = 1) .	Number of persons aged 16 or older who are members of the households for which the interview is accepted for the database (DB135 = 1) and who completed a personal interview (RB250 = 11 to 13).	Number of selected respondents who are members of the households for which the interview is accepted for the database (DB135 = 1) and who completed a personal interview (RB250=11 to 13).
Total cross-sectional	10 868	22 134	10 868
Non-longitudinal from 2005	3 566	7 196	3 566
Non-longitudinal from 2006 (new)	3 709	7 654	3 709
SILC longitudinal panels			
from 2004 for 3 years (not in cross-sectional)	1 719	3 481	1 719
from 2004 for 4 years (not in cross-sectional)	1 663	3 318	1 663
from 2005 for 4 years	1 738	3 488	1 738
from 2006 for 4 years (new)	1 855	3 796	1 855

2.3.3.3 Unit non-response

For Member States using a rotational design, information on unit non-response will be provided for the new replications in accordance with the formulas described below. For the total sample, the unit non-response will be calculated by removing, from the numerator and the denominator of the formulas described below, those units that according to the tracing rules are out of scope.

* Household non-response rates $NRh = (1-(Ra * Rh)) *100$

Ra (address contact rate) = addresses successfully contacted / valid addresses selected =
sum(DB120=11) / [sum(DB120=all) - sum(DB120=23)]

Rh (proportion of complete household interviews accepted for the database) = number of HH interviews completed and accepted for the database / number of eligible households at contacted addresses) =
sum(DB135=1) / sum(DB130=all)

DB120 is the record of contact at the addresses
DB130 is the household questionnaire result
DB135 is the household interview acceptance result

* Individual non-response rates $NRp = (1-Rp)*100$

Rp (proportion of complete personal interviews within the households accepted for the database) = Number of personal interviews completed / number of eligible individuals in the household whose interviews were completed and accepted for the database = sum(RB250=11+12+13) / sum(RB245 = 1+2+3)

RB245 is the respondent status
RB250 is the data status

* Overall individual non-response rates **NRp = (1-(Ra*Rh*Rp))*100*

For those MSs where a <u>sample of persons</u> rather than a sample of HHs (addresses) was selected, the individual non-response rates will be calculated for the selected respondent (RB245=2), for all individuals aged 16 or older (RB245=2+3) and for the non-selected respondent (RB245=3).

Table 2.17 Non-response Rates

Rotational group	DB075	Household non-response rate	Individual non-response rate			Overall individual non-response rate		
			Selected respondent	All individuals 16 or older	Non-selected respondent	Selected respondent	All individuals 16 or older	Non-selected respondent
Total cross-sectional		17.1	0	0	0	17.1	17.1	17.1
Non-longitudinal from 2005	**5**	24.7	0	0	0	24.7	24.7	24.7
Non-longitudinal from 2006 (new)	**6**	7.2	0	0	0	7.2	7.2	7.2
SILC Longitudinal Panels								
from 2004 for 3 years (not in cross-sectional)	**2**	4.6	0	0	0	4.6	4.6	4.6
from 2004 for 4 years (not in cross-sectional)	**3**	5.7	0	0	0	5.7	5.7	5.7
from 2005 for 4 years	**4**	8.0	0	0	0	8.0	8.0	8.0
from 2006 for 4 years (new)	**1**	24.5	0	0	0	24.5	24.5	24.5

The following longitudinal tables do not include new households, because that practice does not occur in the Finnish practice of EU-SILC (tracing only selected respondents).

Table 2.18 Household response rates: comparison of result codes between wave 2 and wave 1 (began 2005, DB075=4)

		Household questionnaire completed		Address unable to access	Entire HH temporarily away for duration of fieldwork	Household unable to respond (illness, incapacity …)	Other reasons	Refusal to co-operate	Address cannot be located	Not contacted	Fusion	Address does not exist etc.	Total (col. %)
		Interview accepted for database	Interview rejected										
Household questionnaire completed	Interview accepted for database (row %)	1 738 (90.90)	0	0	24 (1.26)	4 (0.21)	47 (2.46)	77 (4.03)	0	22 (1.15)	0	0	**1 912 (78.31)**
	Interview rejected												**0**
Address cannot be located													**0**
Address unable to access													**0**
Address does not exist or is non-residential address or is unoccupied or not principal residence													**42 (1.68)**
Refusal to co-operate													**338 (13.52)**
Entire household temporarily away for duration of fieldwork													**54 (2.16)**
Household unable to respond (illness, incapacity …)													**41 (1.64)**
Other reasons													**113 (4.52)**
Total for the first wave													**2500**
Total for the second wave (row %)		**1 738 (90.90)**	**0**	**0**	**24 (1.26)**	**4 (0.21)**	**47 (2.46)**	**77 (4.03)**	**0**	**22 (1.15)**	**0**	**0**	**1 912 (78.31)**

Table 2.19 Household response rates: comparison of result codes between wave 3 and wave 2 (began 2004, duration three years, DB075=2)

Wave 1	Wave 2	Household questionnaire completed		Address unable to access	Entire HH temporarily away for duration of fieldwork	Household unable to respond (illness, incapacity …)	Other reasons	Refusal to co-operate	Address cannot be located	Not contacted	Fusion	Address does not exist etc.	Total (col. %)
		Interview accepted for database	Interview rejected										
Household questionnaire completed	Interview accepted for database (row %)	1 719 (87.44)	0	0	19 (0.97)	5 (0.25)	25 (1.27)	33 (1.68)	0	17 (0.86)	0	0	1 818 (92.47)
	Interview rejected												0
Address cannot be located													0
Address unable to access													0
Address does not exist or is non-residential address or is unoccupied or not principal residence													19 (0.97)
Refusal to co-operate													60 (3.05)
Entire household temporarily away for duration of fieldwork													35 (1.78)
Household unable to respond (illness, incapacity …)													16 (0.81)
Other reasons													18 (0.92)
Total for the first wave													1 966
Total for the second wave (row %)		1 719 (87.44)	0	0	19 (0.97)	5 (0.25)	25 (1.27)	33 (1.68)	0	17 (0.86)	0	0	1 818 (92.47)

Table 2.20 Household response rates: comparison of result codes between wave 3 and wave 2 (began 2004, duration four years, DB075=3)

Wave 1	Wave 2	Household questionnaire completed		Address unable to access	Entire HH temporarily away for duration of fieldwork	Household unable to respond (illness, incapacity …)	Other reasons	Refusal to co-operate	Address cannot be located	Not contacted	Fusion	Address does not exist etc.	Total (col. %)
		Interview accepted for database	Interview rejected										
Household questionnaire completed	Interview accepted for database (row %)	1 663 (85.63)	0	0	11 (0.57)	5 (0.26)	36 (1.85)	49 (2.52)	0	19 (0.98)	0	0	1 783 (91.81)
	Interview rejected												0
Address cannot be located													0
Address unable to access													0
Address does not exist or is non-residential address or is unoccupied or not principal residence													30 (1.54)
Refusal to co-operate													63 (3.24)
Entire household temporarily away for duration of fieldwork													28 (1.44)
Household unable to respond (illness, incapacity …)													13 (0.67)
Other reasons													25 (1.29)
Total for the first wave													1 942
Total for the second wave (row %)		1 663 (85.63)	0	0	11 (0.57)	5 (0.26)	36 (1.85)	49 (2.52)	0	19 (0.98)	0	0	1 783 (91.81)

Table 2.21 Longitudinal statistics for households

Rotational group	Response rates (%)				Longitudinal follow-up rate (%)	Achieved sample size rate (%)
	Wave response rate	Refusal rate	No contacted & others	Total		
Total	93.10	2.89	4.00	100.00	96.06	92.87
DB075=2	94.82	1.82	3.36	100.00	97.25	94.55
DB075=3	93.53	2.76	3.71	100.00	96.19	93.27
DB075=4	91.09	4.04	4.87	100.00	94.82	90.90

The follow-up ratio is the same as the follow-up rate due to the non-existent new households.

Table 2.22 Personal response rates (began 2005, DB075=4)

		PERSONAL INTERVIEW OUTCOME IN WAVE 2										total
		RB250 = 11, 12, 13	Not completed because of									
			RB250 = 21	RB250=22	RB250=23	RB250=31	RB250=32	RB250=33	HHnc	Pn	Pl	
SAMPLE PERSONS (RB100=1 and RB245 = 1,2,3) FROM THE SAMPLE FORWARDED FROM LAST WAVE (1)												
[1] RB110 = 1-2		1 738	0	0	0	0	0	0				**1 738**
[2] RB110 = 6												**0**
[3] RB110 = -1												**0**
[4] RB120 = 2												**0**
[5] RB120 = 3												**0**
[6] RB120 = 4												**0**
[7] DB135=2 or -1, or DB110=7 or DB120=21-23, or DB130 = 21-24 or -1												**174**
[8] DB110 = 3-6												**0**
NON-SAMPLE PERSONS 16+												
[11] This wave	from wave 1	1 699	0	0	0	0	0	0	0	0	0	**1 699**
	not in wave 1	51	0	0	0	0	0	0	0	0	0	**51**
[12] Earlier wave		1 750	0	0	0	0	0	0	0	0	0	**1 750**
SAMPLE PERSONS FROM SAMPLE NOT FORWARDED FROM LAST WAVE (1) (excluded died or not eligible according to the tracing rules)												
[13]												**0**
Sum of rows:												
[1], [3], [6], [7], [9], [10]		1 738	0	0	0	0	0	0	0	0	0	**1 738**
[1], [3], [6], [7], [9], [10], [13]		1 738	0	0	0	0	0	0	0	0	0	**1 738**
[1], [3], [6], [7], [9], [10], [11]		3 437	0	0	0	0	0	0	0	0	0	**3 437**

Table 2.23 Personal response rates (began 2004, DB075=2)

		RB250 = 11, 12, 13	PERSONAL INTERVIEW OUTCOME IN WAVE 3									total
			Not completed because of									
			RB250 = 21	RB250=22	RB250=23	RB250=31	RB250=32	RB250=33	HHnc	Pn	Pl	
SAMPLE PERSONS (RB100=1 and RB245 = 1,2,3) FROM THE SAMPLE FORWARDED FROM LAST WAVE (2)												
[1] RB110 = 1-2		1 719	0	0	0	0	0	0				**1 719**
[2] RB110 = 6												**0**
[3] RB110 = -1												**0**
[4] RB120 = 2												**0**
[5] RB120 = 3												**0**
[6] RB120 = 4												**0**
[7] DB135=2 or -1, or DB110=7 or DB120=21-23, or DB130 = 21-24 or -1												**99**
[8] DB110 = 3-6												**0**
NON-SAMPLE PERSONS 16+												
[11] This wave	from wave 1	**1 708**	0	0	0	0	0	0	0	0	0	**1 708**
	not in wave 1	**54**	0	0	0	0	0	0	0	0	0	**54**
[12] Earlier wave		**1 762**	0	0	0	0	0	0	0	0	0	**1 762**
SAMPLE PERSONS FROM SAMPLE NOT FORWARDED FROM LAST WAVE (2)												
(excluded died or not eligible according to the tracing rules)												
[13]												**0**
Sum of rows:												
[1], [3], [6], [7], [9], [10]		1 719	0	0	0	0	0	0	0	0	0	**1 719**
[1], [3], [6], [7], [9], [10], [13]		1 719	0	0	0	0	0	0	0	0	0	**1 719**
[1], [3], [6], [7], [9], [10], [11]		3 481	0	0	0	0	0	0	0	0	0	**3 481**

Table 2.24 Personal response rates (began 2004, DB075=3)

		RB250 = 11, 12, 13	PERSONAL INTERVIEW OUTCOME IN WAVE 3									total
			Not completed because of									
			RB250 = 21	RB250=22	RB250=23	RB250=31	RB250=32	RB250=33	HHnc	Pn	Pl	
SAMPLE PERSONS (RB100=1 and RB245 = 1,2,3) FROM THE SAMPLE FORWARDED FROM LAST WAVE (2)												
[1] RB110 = 1-2		1 663	0	0	0	0	0	0				**1 663**
[2] RB110 = 6												**0**
[3] RB110 = -1												**0**
[4] RB120 = 2												**0**
[5] RB120 = 3												**0**
[6] RB120 = 4												**0**
[7] DB135=2 or -1, or DB110=7 or DB120=21-23, or DB130 = 21-24 or -1												**120**
[8] DB110 = 3-6												**0**
NON-SAMPLE PERSONS 16+												
[11] This wave	from wave 1	1 603	0	0	0	0	0	0	0	0	0	**1 603**
	not in wave 1	52	0	0	0	0	0	0	0	0	0	**52**
[12] Earlier wave		1 655	0	0	0	0	0	0	0	0	0	**1 655**
SAMPLE PERSONS FROM SAMPLE NOT FORWARDED FROM LAST WAVE (2)												
(excluded died or not eligible according to the tracing rules)												
[13]												**0**
Sum of rows:												
[1], [3], [6], [7], [9], [10]		1 663	0	0	0	0	0	0	0	0	0	**1 663**
[1], [3], [6], [7], [9], [10], [13]		1 663	0	0	0	0	0	0	0	0	0	**1 663**
[1], [3], [6], [7], [9], [10], [11]		3 318	0	0	0	0	0	0	0	0	0	**3 318**

Note that only one person of the household is a selected respondent which yields the situation that the household is not included in the R file, if there is no response for the selected respondent (due to various reasons). Thus the reasons for the non-existing response come from the D file. The persons reaching age 16 are not included in the set of sample persons [9]. There are no new sample persons added because of the exceptional structure of the original cross-sectional data (only two waves included, only one selected respondent per household) [10]. When identified, all the information for non-sample persons 16+ of the household comes from the registers.

Table 2.25 Longitudinal statistics for households, personal level

Rotational group	Wave response rate of sample persons	Wave response rate of co-residents	Longitudinal follow-up rate (%)	For all causes* non-response rate (%)	Achieved sample size ratio for sample persons (%)	Ach. s. s. ratio for sample persons and co-residents (%)	Ach. s. s. ratio for co-residents in the first wave (%)	Response rate for non-sample persons (%)
Total	100.00	100.00	100.00	0.00	92.87	87.10	83.59	100.00
from 2004 for 3 years	100.00	100.00	100.00	0.00	94.55	88.79	85.30	100.00
from 2004 for 4 years	100.00	100.00	100.00	0.00	93.27	86.82	82.92	100.00
from 2005 for 4 years	100.00	100.00	100.00	0.00	90.90	85.75	82.59	100.00

* Causes presented in Table 2.26.

Table 2.26 Distribution of households by household status (DB110), by record of contact at address (DB120), by household questionnaire result (DB130) and by household interview acceptance (DB135)

Household status

	Total	DB110= 1	DB110= 2	DB110= 3	DB110= 4	DB110= 5	DB110= 6	DB110= 7	DB110= 8	DB110= 9	DB110= 10
Total (row %)	5 513 (100.00)	5 455 (98.95)	0	14 (0.25)	7 (0.13)	30 (0.54)	0	7 (0.13)	0	0	0
from 2004 for 3 years	1 818 (100.00)	1 801 (99.09)	0	4 (0.22)	1 (0.06)	9 (0.50)	0	3 (0.17)	0	0	0
from 2004 for 4 years	1 783 (100.00)	1 764 (98.98)	0	1 (0.06)	2 (0.11)	14 (0.79)	0	2 (0.11)	0	0	0
from 2005 for 4 years	1 912 (100.00)	1 890 (98.85)	0	9 (0.47)	4 (0.21)	7 (0.37)	0	2 (0.10)	0	0	0

There are no observations for the table *Record of contact at address (when DB110 = 2, 8, 10).*

Household questionnaire result (when DB120 = 11 or DB110 = 1)

	Total	DB130= 11	DB130= 21	DB130= 22	DB130= 23	DB130= 24	missing
Total (row %)	5 455 (100.00)	5 120 (93.86)	159 (2.91)	54 (0.99)	14 (0.24)	108 (1.98)	0
from 2004 for 3 years	1 801 (100.00)	1 719 (95.45)	33 (1.83)	19 (1.05)	5 (0.28)	25 (1.39)	0
from 2004 for 4 years	1 764 (100.00)	1 663 (94.27)	49 (2.78)	11 (0.62)	5 (0.28)	36 (2.04)	0
from 2005 for 4 years	1 890 (100.00)	1 738 (91.96)	77 (4.07)	24 (1.27)	4 (0.21)	47 (2.49)	0

No observations for the table **Household interview acceptance (when DB130 = 1).**

Table 2.27 Distribution of persons for membership status (RB110)

	Total	RB110= 1	RB110= 2	RB110= 3	RB110= 4	missing
Total (row %)	12 995 (100.00)	12 700 (97.73)	0	172 (1.32)	123 (0.95)	0
from 2004 for 3 years	4 377 (100.00)	4 277 (97.72)	0	58 (1.33)	42 (0.96)	0
from 2004 for 4 years	4 216 (100.00)	4 109 (97.46)	0	60 (1.42)	47 (1.11)	0
from 2005 for 4 years	4 402 (100.00)	4 314 (98.00)	0	54 (1.23)	34 (0.77)	0

The category "no current household members" is not applicable in Finland because of the person approach. No observations for the table Distribution of persons moving out by variable RB120.

2.3.3.4. Distribution of households (original units) by 'record of contact at address' (DB120), by 'household questionnaire result' (DB130), and by 'household interview acceptance' (DB135), for each rotational group (if applicable) and for the total

Table 2.28 Distribution of Households by DB120, DB130 and DB135

Description	Cross-sectional		Non-longitudinal from 2005		Non-longitudinal from 2006 (new)		from 2004 for 3 years (not in cross-sectional)		from 2004 for 4 years (not in cross-sectional)		from 2005 for 4 years		from 2006 for 4 years (new)	
	number	%	number	%	number	%	number	%	number	%	number	%	number	%
Total	13 297	100	3 885	100	5 000	100	1 818	100	1 783	100	1 912	100	2 500	100
Address contacted	13 111	98.60	3 843	98.91	4 913	98.26	1 801	99.06	1 764	98.93	1 890	98.84	2 465	98.6
Address non-contacted	186	1.40	42	1.09	87	1.74	17	0.94	19	1.06	22	1.15	35	1.4
Total address non-contacted	186	100	42	100	87	100	17	100	19	100	22	100	35	100
Address cannot be located	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Address unable to access	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Address does not exist, etc.	186	100	42	100	87	100	17	100	19	100	30	100	35	100

Description	Cross-sectional		Non-longitudinal from 2005		Non-longitudinal from 2006 (new)		from 2004 for 3 years (not in cross-sectional)		from 2004 for 4 years (not in cross-sectional)		from 2005 for 4 years		from 2006 for 4 years (new)	
	number	%	number	%	number	%	number	%	number	%	number	%	number	%
Total	13 111	100.0	3 843	100.0	4 913	100.0	1 801	100.0	1 764	100.0	1 890	100.0	2 465	100.0
Household questionnaire completed	10 868	82.89	3 566	92.79	3 709	75.49	1 719	95.45	1 663	94.27	1 738	91.96	1 855	75.25
Interview not completed	2 243	17.11	277	7.21	1 204	24.51	82	4.55	101	5.73	152	8.14	610	24.75
Total interview not completed	2 243	100.0	277	100.0	1 204	100.0	82	100.0	101	100.0	152	100.0	610	100.0
Refusal to co-operate	1 289	57.47	117	42.24	732	60.80	33	40.24	49	48.51	77	50.66	363	59.51
Entire household temporarily away for duration of fieldwork	240	10.70	63	22.74	95	7.89	19	23.17	11	10.89	24	15.79	58	9.51
Household unable to respond	153	6.82	17	6.14	87	7.23	5	6.10	5	4.95	4	2.63	45	7.38
Other reasons	561	25.01	80	28.88	290	29.08	25	30.49	36	35.64	47	30.92	144	23.60
Household questionnaire completed	10 868	100.0	3 566	100.0	3 709	100.0	1 719	100.0	1 663	100.0	1 738	100.0	1 855	100.0

Interview accepted for database	10 868	100.0	3 566	100.0	3 709	100.0	1 719	100.0	1 663	100.0	1 738	100.0	1 855	100.0
Interview rejected	0	0	0	0	0	0	0	0	0	0	0	0	0	0

2.3.3.5 Item non-response

Item non-responses before imputing exist for the interviewed item of the variable HY090G (interest income taxed at source) and register income components HY030G, HY100N, HY022 and HY023 for which imputation procedures have been used. For the total household income variables HY010 and HY020, the number of households with item non-responses have been counted from that part they have effect on the total income formation. The counting method equals to the income flag values.

The data collection of interests income taxed at source has changed from the 2005 survey year. The former question on precise income value has been divided to two questions on precise and range income values. The additional range value question produces an increase in response rate, especially with regards to small values. As a result, the number of the households that received HY090G, but also the number of ones with impartial information in that variable is now higher than in the previous survey years. Both question types include missing values, which have been imputed. (See Chapter 2.5.)

There are differences in the HY090G impartial information rates between rotational groups of the longitudinal survey (DB075=1,2 vs. DB075=3), which are caused by changes in the data collection and processing (editing/imputing) of interest income taxed at source.

Table 2.29 Distribution of item non-response in the EU-SILC 2006 survey sample selected for the cross-sectional survey only (X) or the longitudinal survey according to rotational group (L, DB075 = 2, 3, 4), all households and persons 16+

Income component	(A) * % of households having received an amount (<0, >0)					(B) % of households with missing values (before imputation)					(C) % of households with partial information (before imputation) of all households				
	X	L	DB075			X	L	DB075			X	L	DB075		
	All	All	2	3	4	All	All	2	3	4	All	All	2	3	4
HY010	100.0	99.9	99.9	99.9	99.9	0.0	0.0	0.0	0.0	0.0	24.2	29.7	32.5	33.1	24.0
HY020	100.0	100.0	100.0	100.0	100.0	0.0	0.0	0.0	0.0	0.0	26.4	32.4	36.0	36.1	25.9
HY022	98.5	98.8	99.0	99.0	98.5	0.0	0.0	0.0	0.0	0.0	27.5	34.0	37.8	38.6	26.7
HY023	97.7	98.1	98.3	98.4	97.8	0.0	0.0	0.0	0.0	0.0	28.7	36.4	41.3	41.2	27.8
HY030G	77.4	77.4	78.1	76.5	77.5	77.4	77.4	78.1	76.5	77.5	77.4	77.4	78.1	76.5	77.5
HY040G	10.5	10.0	9.9	10.1	10.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
HY050G	33.6	32.5	32.3	32.4	32.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
HY060G	6.5	6.2	6.0	6.0	6.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
HY070G	16.6	17.3	16.8	18.6	16.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
HY080G	8.1	9.1	9.0	10.4	8.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
HY090G	85.8	87.3	88.8	87.9	85.3	0.0	0.0	0.0	0.0	0.0	37.6	45.3	49.8	52.4	35.3
HY100G	36.7	35.6	35.7	35.2	36.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
HY110G	3.3	3.0	2.7	3.2	3.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
HY120G	55.9	55.1	54.8	53.6	56.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
HY130G	15.3	17.7	20.5	18.2	14.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
HY140G	98.8	99.1	99.4	99.2	98.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
HY135G															
HY100N	36.7	35.6	35.7	35.2	36.0	36.7	35.6	35.7	35.2	36.0	36.7	35.6	35.7	35.2	36.0

Income component	% of persons 16+ having received an amount (<0, >0)					% of persons 16+ with missing values (before imputation)					% of persons 16+ with partial information (before imputation) of all persons				
	X	L	DB075			X	L	DB075			X	L	DB075		
	All	All	2	3	4	All	All	2	3	4	All	All	2	3	4
PY010G	63.7	63.3	64.2	62.7	62.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PY020G	2.1	2.2	2.7	2.3	1.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PY030G	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
PY035G	11.7	11.0	11.1	9.9	11.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PY050G	21.6	21.1	21.1	21.0	21.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PY070G	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
PY080G	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
PY090G	13.5	13.3	12.8	13.1	14.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PY100G	18.0	19.0	18.8	19.0	19.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PY110G	1.1	1.3	1.4	1.5	1.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PY120G	6.5	6.4	6.2	7.1	6.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PY130G	8.1	8.6	8.3	9.1	8.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PY140G	10.1	10.4	10.4	10.4	10.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
PY200G	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..

.. information is not available
The computation of the impartial information rates have been done on the basis of the target variable flag values.

38

2.4 Mode of data collection

Table 2.30 Distribution of Household Members aged 16 and over by 'RB250' and 'RB245'

Rotational group	Total	RB250 =11	RB250 =12	RB250 =13	RB250 =21	RB250 =22	RB250 =23	RB250 =31	RB250 =32	RB250 =33
Household members 16+ and RB245 = 1 to 3										
Cross-sectional total	22 134	0	0	22 134	0	0	0	0	0	0
	100	0	0	100	0	0	0	0	0	0
Non-longitudinal from 2005	7 196	0	0	7 196	0	0	0	0	0	0
	100	0	0	100	0	0	0	0	0	0
Non-longitudinal from 2006 (new)	7 654	0	0	7 654	0	0	0	0	0	0
	100	0	0	100	0	0	0	0	0	0
from 2004 for 3 years (not in cross-sectional)	3 481	0	0	3 481	0	0	0	0	0	0
	100	0	0	100	0	0	0	0	0	0
from 2004 for 4 years (not in cross-sectional)	3 318	0	0	3 318	0	0	0	0	0	0
	100	0	0	100	0	0	0	0	0	0
from 2005 for 4 years	3 488	0	0	3 488	0	0	0	0	0	0
	100	0	0	100	0	0	0	0	0	0
from 2006 for 4 years (new)	3 796	0	0	3 796	0	0	0	0	0	0
	100	0	0	100	0	0	0	0	0	0
Household members 16+ and RB245 = 2										
Cross-sectional total	10 868	0	0	10 868	0	0	0	0	0	0
	100	0	0	100	0	0	0	0	0	0
Non-longitudinal from 2005	3 566	0	0	3 566	0	0	0	0	0	0
	100	0	0	100	0	0	0	0	0	0
Non-longitudinal from 2006 (new)	3 709	0	0	3 709	0	0	0	0	0	0
	100	0	0	100	0	0	0	0	0	0
from 2004 for 3 years (not in cross-sectional)	1 719	0	0	1 719	0	0	0	0	0	0
	100	0	0	100	0	0	0	0	0	0
from 2004 for 4 years (not in cross-sectional)	1 663	0	0	1 663	0	0	0	0	0	0
	100	0	0	100	0	0	0	0	0	0
from 2005 for 4 years	1 738	0	0	1 738	0	0	0	0	0	0
	100	0	0	100	0	0	0	0	0	0
from 2006 for 4 years (new)	1 855	0	0	1 855	0	0	0	0	0	0
	100	0	0	100	0	0	0	0	0	0
Household members 16+ and RB245 = 3										
Cross-sectional total	11 266	0	0	11 266	0	0	0	0	0	0
.	100	0	0	100	0	0	0	0	0	0
Non-longitudinal from 2005	3 630	0	0	3 630	0	0	0	0	0	0
	100	0	0	100	0	0	0	0	0	0
Non-longitudinal from 2006 (new)	3 945	0	0	3 945	0	0	0	0	0	0
	100	0	0	100	0	0	0	0	0	0
from 2004 for 3 years (not in cross-sectional)	1 762	0	0	1 762	0	0	0	0	0	0
	100	0	0	100	0	0	0	0	0	0
from 2004 for 4 years (not in cross-sectional)	1 655	0	0	1 655	0	0	0	0	0	0
	100	0	0	100	0	0	0	0	0	0
from 2005 for 4 years	1 750	0	0	1 750	0	0	0	0	0	0
	100	0	0	100	0	0	0	0	0	0
from 2006 for 4 years (new)	1 941	0	0	1 941	0	0	0	0	0	0
	100	0	0	100	0	0	0	0	0	0

Table 2.31   Distribution of Household Members aged 16 and over by 'RB260' and 'RB245'

Rotational group	Total	RB260=1	RB260=2	RB260=3	RB260=4	RB260=5	RB260= missing
Household members 16+ and RB245 = 1 to 3							
Cross-sectional total .	22 134	0	325	10 562	0	11 247	0
	100	0	1.47	47.72	0	50.81	0
Non-longitudinal from 2005	7 196	0	61	3 508	0	3 627	0
	100	0	0.85	48.75	0	50.40	0
Non-longitudinal from 2006 (new)	7 654	0	156	3 567	0	3 931	0
	100	0	2.04	46.60	0	51.36	0
from 2004 for 3 years (not in cross-sectional)	3 481	0	19	1 700	0	1 762	0
	100	0	0.55	48.84	0	50.62	0
from 2004 for 4 years (not in cross-sectional)	3 318	0	18	1 645	0	1 655	0
	100	0	0.54	49.58	0	49.88	0
from 2005 for 4 years	3 488	0	40	1 698	0	1 750	0
	100	0	1.15	48.68	0	50.17	0
from 2006 for 4 years (new)	3 796	0	68	1 789	0	1 560	0
	100	0	1.79	47.13	0	41.10	0
Household members 16+ and RB245 = 2							
Cross-sectional total .	10 868	0	276	8 284	0	2 308	0
	100	0	2.54	76.22	0	21.24	0
Non-longitudinal from 2005	3 566	0	55	2 746	0	765	0
	100	0	1.54	77.01	0	21.45	0
Non-longitudinal from 2006 (new)	3 709	0	123	2 763	0	823	0
	100	0	3.32	74.49	0	22.19	0
from 2004 for 3 years (not in cross-sectional)	1 719	0	17	1 298	0	404	0
	100	0	0.99	75.51	0	23.50	0
from 2004 for 4 years (not in cross-sectional)	1 663	0	14	1 289	0	360	0
	100	0	0.84	77.51	0	21.65	0
from 2005 for 4 years	1 738	0	36	1 361	0	341	0
	100	0	2.07	78.31	0	19.62	0
from 2006 for 4 years (new)	1 855	0	62	1414	0	379	0
	100	0	3.34	76.23	0	20.43	0
Household members 16+ and RB245 = 3							
Cross-sectional total .	11 266	0	49	2 278	0	8 939	0
	100	0	0.43	20.22	0	79.35	0
Non-longitudinal from 2005	3 630	0	6	762	0	2 862	0
	100	0	0.17	20.99	0	78.84	0
Non-longitudinal from 2006 (new)	3 945	0	33	804	0	3 108	0
	100	0	0.84	20.39	0	78.78	0
from 2004 for 3 years (not in cross-sectional)	1 762	0	2	402	0	1 358	0
	100	0	0.11	22.81	0	77.07	0
from 2004 for 4 years (not in cross-sectional)	1 655	0	4	356	0	1 295	0
	100	0	0.24	21.51	0	78.25	0
from 2005 for 4 years	1 750	0	4	337	0	1 409	0
	100	0	0.23	19.26	0	80.51	0
from 2006 for 4 years (new)	1 941	0	6	375	0	1 560	0
	100	0	0.31	19.32	0	80.37	0

2.5 Imputation procedure

Imputation procedures were used for an interviewed item of the income variables HY090G, and for the other income variables HY030G, HY100N and HY022 and HY023.

Interests income taxed at source, which is counted in HY090G interest, dividends, profit from capital investments in unincorporated business was imputed by the hot-deck method (a stochastic method). Imputing was done automatically by the SAS/EG-supporting program as follows:

1. Defining the sample to the household units received the interests income taxed at source during the reference year (yes), if a precise monetary amount value was given or not.
2. Detecting record outliers of the responded monetary values, and dropping the units out from the donors.
3. Grouping the units by domicile code (indicates the location of the household's dwelling) and range value given in the interview.
4. Checking the criterion for the proportion of responded records of all records in the groups.
5. Filling the item non-responses by selecting randomly from the responded records of the nearest donors in the range value groups. Automatic imputing.
6. Grouping the units by domicile code, the socio-economic status of the household reference person and the number of the household members.
7. Checking the criterion for the proportion of value records of all records.
8. Filling the rest of the item non-responses by selecting randomly from the records (responded values and imputed values in the groups) of the nearest donors. Automatic imputing.

Compared to the 2005 cross-sectional survey, the change in the imputation procedure is due to the changed interview in the 2006 survey. Information in interests income taxed at source is interviewed in two phases (1. a precise value; 2. if doesn't know, a range value), instead of one phase interview (a precise value) earlier. As a result, there is an increase of HY090G receivers and small values in the sample.

The processing and imputing of the second wave units (DB070=4) in the 2006 longitudinal survey is consistent with the 2006 cross-sectional survey. The data on interest income taxed at source was not collected to the units in the third wave (DB075=2,3) of the 2006 longitudinal survey. For them, responded values in the 2006 cross-sectional survey (second wave of the cross sectional survey) were used as donors for imputing. Item was imputed by the hot-deck method in the groups by sample strata (a proxy measure for socio-economic group), the number of household members and information on interest income taxed at source received during the previous year/wave.

For HY030G, the stratification method as a deterministic method was used to impute market rents values to households' equivalent dwellings from an external data source (See Table 3.3). For HY100N, HY022 and HY023, deductive imputation was used to convert taxable social transfers in gross amount for net amount (See Table 3.3).

2.6 Imputed rent

Imputed rent was provided for the target variable HY030G in Finland's SILC 2006. The method is described in Table 3.3, Section 3.2.3.

2.7 Company cars

Information on a company car was collected from the Personal Tax Register of National Board of Taxes.

3 Comparability

3.1 Basic concepts and definitions

Basic concepts and their definitions are in accordance with the Commission Regulation (EC) No 1980/2003 provided for the community statistics on income and living conditions as regards definitions and updated definitions. To some extent, adaptation of the definitions used in the national statistical system is allowed for the EU-SILC. In Finland, private household and household membership in particular are the ones that have been defined nationally (e.g. IDS) with less detailed information (i.e. time duration for temporarily absence in private accommodation) than stated in the regulations, but within the framework.

The reference population consists of the members of the private households permanently resident (usually resident: the census definition) in Finland on 31 December 2005. For migrants in particular, permanently residence means that they have resided or intend to reside for at least 12 months and they have not permanent residence elsewhere. Persons living in institutions, in collective households or in residential homes[2] have been excluded.

The private household was constructed to include a person residing alone, or all the persons, related or not, who reside and have their meals together or otherwise use their income together. The definition equals with the obliged EU-SILC definition on shares in household expenses, but uses other words "use income together" in the interview.

If a person was temporarily absent from the household's main dwelling and from home, no specific time duration was set for the absence provided that the above-mentioned criteria of household formation and membership (shares in household expenses) were fulfilled. Such persons have close family ties to the household and they do not form a household of their own. Therefore, the following persons are also counted in household members:
- Persons conducting military service or conscript service
- Persons residing and working in another locality or abroad if they are involved
  in the acquisition and use of household income
- Persons residing and studying in another locality if they use income received mostly from
  their parents
- Persons temporarily in institutions, on holiday or travelling.

The following persons form a household of their own:
- Subtenants
- Domestic staff
- Students living on their own if they live mostly on their own income or on a student loan
- Students residing in dormitories, unless they are married or officially cohabiting.

In the longitudinal survey, the following persons except the sample persons, were not household members any more:
- Persons moved out from households permanently or died during the year 2005
- or persons who otherwise were not permanently living in the household containing a sample person on 31 December 2005

---

[2] Residential homes are situated either in residential or institutional care buildings and do not meet the definition of dwelling. They do not include a kitchen or cooking facilities, a water closet or cleaning facilities (shower, bathroom or sauna). Students dormitories which are counted in the private household definition above include these facilities.

The permanently resident population that has not included in private households refers to the difference between the number of all persons and the private household persons permanently resident in Finland on 31 December. The number of total population is 5,255,580, from which about 1.5 per cent was not in the private households, but was permanently institutionalised or living in collective households or residential homes. The number of estimated household population was 5,179,228 in the cross-sectional survey. The estimated household population number was 5,137,344 (DB075=1), 5,098,883 (DB075=2) and 5,113,347 (DB075=3) in the rotational groups of the longitudinal survey.

Other definitional solutions done are due to the collection of the information both from registers and by interviews. These are related to **reference times**. First, current as a reference time refers to several calendar time points. Information collected solely by interviews (e.g. non-monetary deprivation indicators, education, health) refers to the interview time point in the survey year (2006). Information collected by interviews, but used for the target variables as combined with the information from registers and other information interviewed on themes close to income is related to the income reference period, which is the fixed 12-month period before the survey year, i.e. the whole calendar year (2005). Then, the current is either the last day (dwelling, characteristics of dwelling for the imputed rent, housing environment) or the last month (economic activity, housing costs) of the income reference year. In particular, information on housing arrears is consistent with information on housing costs from the income reference period, not from the last twelve months preceding the interview time point as provided.

Table 3.1 Finland's definitions for the reference periods in the EU-SILC 2006 survey.

Current, time point of interview for the respondent in the survey year 2006:
- Non-monetary household deprivation indicators
- Housing (amenities in the dwelling)
- Education
- Health

Current, last day (31 Dec.) of the income reference period (2005):
- Basic data
- Physical and social environment
- Housing (dwelling type, tenure status and housing conditions)

 Current, last month (December) of the income reference period (2005):
- Child care
- Labour information on current activity status and current main job,
  including information on last main job for unemployed
- Detailed labour information
- Housing costs (a part of housing costs)

Last 12 months preceding the time point of interview:
- Health (access to health care)

Income reference period (a fixed 12-month period), i.e. 2005:
- Changes in household composition (longitudinal survey)
- Income
- Labour information (basic information on activity status, calendar of activities)
- Housing and non-housing related arrears.
- Housing costs (a part of housing costs, e.g. income related items)

The income reference period is the preceding calendar year of the survey year, i.e. a fixed 12-month period. Income taxed by the Bookkeeping Act received from the completed accounting periods in the income reference period is included. These are business income, income from dividends and interest.

The reference period for taxes on income and social contributions is the years when taxes are paid from the income received during the income reference period. The taxes are paid in the income reference period (t) and the following years (t+1, t+2). The social contributions are mostly paid in the income reference period (t).[3]

The reference period of taxes on wealth is the years when taxes are paid from the wealth owned in the income reference period. Most of the payments are done during the income reference year.

The time lag between the income reference period and current variables is in its maximum when current information is from the interview time point. The last interview was conducted on 12 June in the survey year **to the cross-sectional part of the survey**. The time lag **is then 5.4 months**. However, most of the current information is from the end of the income reference period and then the time lag does not exist.

Interviews were conducted from 17 January to 12 June in the survey year 2006. **The duration of interviewed data collection was 4.8 months to the cross-sectional part of the survey**. Of all household interviews, 25 per cent were collected by 14 February, 50 per cent by 8 March, 75 per cent were collected by 5 April, and 90 per cent by 17 May.

The interview data collection was started earlier on 12 December in 2005 to **the rotational groups selected for the longitudinal survey** than in the ones selected for the cross-sectional survey. About 6 per cent of the interviews were done during the end of the year 2005. All the interviews were done in longitudinal groups by 27 March. This means, that the time lag of the current information in relation to income information and the **duration of interviewed data collection** were shorter **to the longitudinal** than to the cross-sectional **part of the survey, about 3.9 months in it's maximum.**

For the register database, the last information was collected on 20 November in the survey year 2006. When data collection from registers is included in the measurement, **the duration of the whole data collection both from interviews and registers was 10.5 months.**

The basic information on activity status during the income reference period was derived from information on a person's main activity in each month by summing the activities over the months (twelve in total, see interviewed groups below). The information on a person's main activity was interviewed from the household respondent. For answering to a question, the respondent was instructed to give priority to employment over non-economic activity and inactivity if that person had had several activities during the month. Full-time and part-time work was separated by working hours. Work was full-time if a person worked at least 30 hours per week. Otherwise, work was part-time if a person worked under 30 hours per week. In economically inactive statuses, the answer is based on the respondent's assessment about his/her main activity during the month.

---

[3] Most of the taxes (incl. taxes on net wealth owned) and social contributions are actually done during the income reference year (t) as withholdings by a payer or advance payments by a person, i.e. 89 per cent of enforced taxes in 2005 (National Board of Taxes 2007). Some of these payments can be done up till March of the year after the income reference period (t+1). As a result of the enforced taxation by tax authorities, 6.7 per cent of the enforced taxes were received as tax refunds in the year after the income reference period (t+1), 4.2 per cent of the enforced taxes were paid as residual taxes in the year after the income reference period (t+1) and further in the beginning of the following year (t+2). If demands of rectification and petition of appeals were proceeded, in a few cases, taxes are paid later (t+3,…,n).

The consistency is highest among employees and pensioners. 90.5 per cent of the total withholdings and advance payments for employees and 94 per cent of the total withholdings and advance payments for pensioners were in accordance with the enforced taxes in 2005. The consistency was not as high among self-employed persons, 78.7 per cent of total advance payments done by farmers and 75.6 per cent done by other self-employed persons were in accordance with the enforced taxes. (National Board of Taxes 2007).

The target variables on a person's activity status during the income reference period and the detailed subgroups interviewed are as follows :

PL070, Number of months at full-time work:
- Employee working full-time (at least 30 hours per week)
- Entrepreneur or unpaid worker in family enterprise working full-time (at least 30 hours per week)
PL072, Number of months at part-time work:
- Employee working part-time (under 30 hours per week)
- Entrepreneur or unpaid worker in family enterprise working part-time (under 30 hours per week)
PL080, Number of months in unemployment:
- Unemployed
PL085, Number of months in retirement:
- Retiree
PL087, Number of months in studying:
- Pupil, student
PL090, Number of months in inactivity:
- On unpaid sickness leave, etc.
- Others
- In military service or conscript service

3.2 Components of income

3.2.1 Differences between the national definitions and standard EU-SILC definitions

Total household gross income and disposable household income

The target variables on gross income components, on **gross total household income, HY010**, and on **total disposable household income HY020 and total disposable household income before social transfers other than old-age and survivors' benefits HY022 and including old-age and survivors' benefits HY023** are well in accordance with the requirements.

HY010 is the sum of gross income components at the household level. HY020 is HY010 after current transfers paid are deducted. HY010 is a positive value (incl. 0 values). Negative values of the net income variables HY020, HY022 and HY023 on total disposable household income are due to such current transfers paid which are not related to the total household gross income HY010. These are regular taxes on wealth HY120G, which may exceed the amount of the total household gross income by the EU-SILC definition. The number of the sample households with negative values was six in HY020, 82 in HY022 and 494 in HY023 in the cross-sectional part of the survey. For calculating the common European Union indicators, the negative values were set for zero values. The conversion has an effect on the mean and variance estimates (HY020 mean equivalised income and Gini coefficient estimates in the indicators as well).

Tax on income and social insurance contributions HY140G and regular inter-household transfers paid HY130G were subtracted from the total household gross income received during the income reference year. They did not cause negative values to total household income components.

The target variable **HY030G on imputed rent** is required as a compulsory variable from 2007 onwards. Information on the variable is according to the national Income Distribution Statistics in the 2006 EU-SILC survey. The stratification method used follows the EU-SILC regulation (Table 3.3). Imputed rent was not counted in the total household income components (HY010, HY020, HY022, HY023).

The target variable HY030G covers imputed rent for all households occupying own dwellings and the households renting a dwelling from other households at a lower price than the market price, or getting it free. Differing from what the regulation states about the coverage, imputed rent was not yet calculated for dwellings rented from other sector (e.g. public, municipal, voluntary or non-profit agency) in the 2006 survey.

Income received

The variables on gross income components were obtained by summing the detailed gross items to the person and household unit level. Especially when register income is available as very detailed items, the aggregating of the items for the target variables is closely in accordance with the regulations. **Compared with the Regulation definitions on the EU-SILC gross income components, the following differences,** however, appear due to using register information within the Personal Tax Register frame:

- Payments (PY080G) received by persons from their voluntary personal insurance plans, which

in the same register item as income received from statutory, voluntary insurance taken by

ployers for their employees. This income could not be excluded exactly. It was counted in

rests, dividends, profits from capital investment in an unincorporated business (HY090G).
- Income received from statutory, voluntary insurance taken by employers or employed persons (incl. collective voluntary schemes) themselves to supplement compulsory social security was counted in interests, dividends, profits from capital investment in an unincorporated business (HY090G).
- Earned income from agriculture, forestry and business activities was counted in self-employment income (PY050G). In addition, income from business activities was included for shareholders in corporations. Income is in gross amounts after expenses except interests on loans for acquisition of income. They were counted as deductions for taxable income and result as lower taxes paid (HY140G).
- Loan interests diminish the taxable income as deductions, and result as lower taxes paid (HY140G).
- Self-employed income (PY050G) is positive (incl. 0 income). Losses were considered for lower taxes paid from other type of income in the income reference period, or in the spouse's taxes paid. If no taxable income was received at all, the confirmed losses are considered in taxes that will be paid from the income received in the following years. Therefore, confirmed losses both from the income of the income reference period and from previous periods as well can both have an effect on taxes paid from the reference period's income (HY140G).[4]
- Both received social benefits and social benefits obliged to be returned to payers were included in the certain target variables on social benefits (PY090G, PY100G, PY110G, PY120G, PY140G, HY050G, HY060G, HY070G). The statistical units have negative values on these variables if the social benefits were solely returned back, or the returned amount exceeded the amount received during the income reference period. The social benefits are

---

[4] In the cross-sectional sample, 19.2 per cent of self-employed persons (PL030 =1,2, & PL040 = 1,2) had income on PY050G (n = 670 / 3 492). Most of them had other income sources, from which personal income on PY010G and on HY090G were the highest income sources. 93 per cent of the persons got income on either PY010G or HY090G at personal level, 3.7 of the persons had only other type of income and 3.7 of the persons had not income at all during the reference year. Persons who were temporarily away from work are counted in the numbers. Losses were in 6.2 per cent of all self-employed-persons (n=217), of whom 36.4 per cent had losses (incl. losses from HY090G) which were considered as deductions for taxes paid (incl. all income to which deductions focused), and of whom 72.4 per cent had confirmed losses (the rest of the losses or all) which can be considered in the taxes paid from income received after the income reference year. In addition, a small number of losses were counted in the spouse's taxation.

obliged to be returned if the income or living conditions have changed and they are not valid in relation to the allowed criteria any more.

- Income received personally by people aged under 16 was counted in the target variable HY110G. The variable consists of the following type of income: employee income, self-employment income, property income, education related allowances, survivors' benefits, disability benefits and part of family/children-related allowances. Other social benefits within the ESSPROS system are paid for children's carers, and were counted in family benefits (HY050G).

Current transfers paid

The target variable on **tax on income and social insurance contributions (HY140G)** includes taxes paid for the state taxation and for the municipal taxation. For the state taxation, taxes from earned income (incl. social benefits) are paid progressively by the person's income level, taxes from capital income are paid uniformly (28 per cent of capital income in 2005). For municipal taxation, taxes from earned income are paid by the tax rate of the place of domicile that a person hold at the end (31 Dec.) of the year preceding the income reference year.

The social contributions include the following items: compulsory sickness contributions, unemployment contributions and pension contributions.

The target variable on **regular taxes on wealth (HY120G)** includes taxes on real property owned and taxes paid on net wealth owned at the end (31 Dec.) of the income reference period. Besides, taxes on real property owned are paid indirectly in utility costs of dwellings by shareholders in housing corporations. The tax was not included in HY120G, but it was counted in housing costs (HH070) and consequently, as a part of the housing costs component it diminishes the gross rent value in the imputed rent (HY030G).

Table 3.2 Components of income. Finland's definitions and assessed consequences resulting from differences compared with the EU-SILC definition in the 2006 survey.

Components of income	Variable name	The definition	Consequences to comparability 1 = comparable 0 = not comparable
	HY010		1 See notes below
Total disposable household income	HY020		1 See notes below
Total disposable household income, before social transfers other than old-age and survivors' benefits	HY022		1 See notes below
Total disposable household income, before social transfers including old-age and survivors' benefits	HY023		1 See notes below
Imputed rent	HY030G	Imputed gross rentals for all households that do not report paying full rent, either because they are owner-occupiers or they live in accommodation rented from another household at a lower price than the market price or rent-free minus housing costs actually paid (incl. subsidies received from government). Imputed for the dwellings which are used as the main residence of the household.	0 Note: Consistent with the Finnish IDS. **Imputed rents for households living in dwellings rented at a lower price than the market price from a public, municipal, voluntary or non-profit agency have not been included.** The information on imputed gross rentals (equivalent to market rent) is based on the rent including other utility costs done besides the "space rent". After deducting consistent housing costs paid by the household, the definition is comparable. Imputed rent has not been included in the gross

			household income variable (HY010) or the total disposable household income variables (HY020, HY022, HY023).
Income from rental of property or land	HY040G	Income received, during the income reference period, from renting a property less expenses except interest payments.	1 Note: Interest payments on loans for acquisition of income are considered as deductions from taxable income in taxation, and thus diminish the amount of taxes paid on the income (HY140G).
Family/children-related allowances	HY050G	Financial support to households for bringing up children and financial assistance to people who support relatives other than children: income maintenance benefit in the event of childbirth, birth grant, parental leave benefit, family or child allowance, other cash benefits.	1
Social exclusion payments not elsewhere classified	HY060G	Social benefits to the socially excluded or to those at risk of social exclusion: income support to people with insufficient resources, and other cash benefits as support for destitute and vulnerable persons to help alleviate poverty or assist in difficult situations.	1 Note: A register-based item on income support also includes a minor part of means-tested housing allowance. Parts can not be separated.
Housing allowances	HY070G	Rent benefit or benefit to owner-occupiers, means-tested	1
Regular inter-household cash transfers received	HY080G	Regular monetary amounts or monetary amounts over the certain minimum amount (EUR 100) received during the income reference period, from other households or persons: compulsory child support, voluntary support to education, voluntary payments for housing costs and utility bills.	1
Interest, dividends, profit from capital investments in unincorporated businesses	HY090G	The amount of interest from assets, dividends and profits from capital investment in an unincorporated business in which the person does not work, received during the income reference period, less expenses incurred. Interests on loans for acquisition of income are considered as expenses for certain income items, but not for all income items.	1 Note: Interest payments on loans for acquisition of income are subtracted as deductions from taxable income in taxation, and thus diminish the taxes paid on income. (HY140G). The component includes income from statutory benefits (incl. pensions) undertaken voluntarily by an employer, an employed person (entrepreneur), or a person individually in addition to the compulsory scheme of social benefits. These are a few register items which cannot be separated.
Interest paid on mortgages	HY100G	Total gross amount, before deducting any tax credit or tax allowance, of mortgage interest on the main residence of the household during the income reference period.	1
Interest paid on mortgages	HY100N	Total gross amount, after deducting any tax credit or tax allowance, of mortgage interest on the main residence of the household during the income reference period. Tax allowance from mortgage interest expenses is considered as deductions from taxable capital and earned income in taxation, and thus diminish taxes paid on the income (HY140G).	1
Income received by people aged under 16	HY110G	Gross income received by all household members aged under 16 during the income reference period.	1
Regular taxes on wealth	HY120G	Taxes that are payable annually on the ownership or use of land and buildings paid during the income reference period (t). Taxes that are payable on the net wealth owned at the end (31 Dec.) of the income reference period (t).	1 Taxes paid on the ownership and use of land and buildings by shareholders in housing companies have been included in housing costs and imputed rent.
Regular inter-household transfers paid	HY130G	Regular monetary amounts or monetary amounts over the certain minimum amount (EUR 100) paid during the income reference period to other households or persons: compulsory child support, voluntary support to education, voluntary payments for housing costs and utility bills.	1
Tax on income and social insurance contributions	HY140G	Taxes on income, profits and capital gains: taxes on individual, household or tax-unit income (income from employment, property, entrepreneurship, pensions, etc.) including taxes deducted by employers (i.e. withholdings), other taxes at source and taxes on the income of owners of unincorporated enterprises paid **from the income received in the income reference year**. Social insurance contributions paid during the income reference period.	1 Note: Interests charged on arrears of taxes due and any fines imposed by tax authorities have not been included.
Repayments/receipts for tax adjustments	HY145G	-	-
Cash or near-cash employee income	PY010G	Monetary component of the compensation of employees in cash payable by an employer to an employee: value of any social contributions and income taxes payable by an employee or by the	1 Note: Tips and bonuses, and benefits based on profit

		employer on behalf of the employee to social insurance schemes or tax authorities.	sharing from stock options (excl. the ones converted into cash) have been included in this component according to the Finnish IDS.
Non-cash employee income	PY020G	Non-monetary income components which may be provided free or at a reduced price to an employee as part of the employment package by an employer: company car and associated costs, free or subsidised meals, luncheon vouchers, reimbursement or payment of housing-related expenses, other goods and services provided free or at a reduced price by their employer to their employees.	1 Note: Company car has been included.
Employers' social insurance contributions	PY030G	-	-
Cash profits or losses from self-employment (including royalties)	PY050G	The income received, during the income reference period, by individuals, for themselves or in respect of their family members, as a result of their current or former involvement in self-employment jobs: operating profit accruing to working owners or partners of an unincorporated enterprise, royalties earned on writing, inventions and so on, not included in the profit/loss of unincorporated enterprises, rentals from business buildings, vehicles, equipment, etc., not included in the profit/loss of unincorporated enterprises, after deduction of related costs. Interests on loans for acquisition of income are considered as costs for a few income items, but not for all income items.	1 Note: Interest payments on loans for acquisition of income are subtracted as deductions for taxable income in taxation, and thus diminish the taxes paid on income (HY120G). Positive values (incl. 0 values). Losses are considered as deductions for taxable income and diminished taxes paid from other type of income in the income reference year, or in the spouse's taxes paid. If such taxable income that deductions concern has not been received at all, losses will be considered as taxes paid from the income received in the following years. Comparability over time: Abolishment of avoir fiscal credit reduces PY050G compared to the 2005 survey. The effect is assessed to be slight. 0.3 per cent of persons received the income and the item consisted of 0.1 per cent of the estimated PY050G amount in the 2005 survey.
Value of goods produced for own consumption	PY070G	-	Note: Value is not significant at national level, or to particular groups of households. The information is not collected.
Unemployment benefits	PY090G	Benefits that replace income lost by a worker due to the loss of gainful employment, provide subsistence income to persons entering or re-entering the labour market, provide subsistence income to unemployed persons not members in unemployment funds, provide subsistence income to persons in long-term unemployment, and to elderly persons who retire after long-term unemployment before the legal retirement age, contribute to the cost of training or re-training people looking for employment.	1
Old-age benefits	PY100G	Benefits that provide replacement income when an aged person retires from the labour market, or guarantee certain income when a person has reached the prescribed age. Old-age pensions, early old-age pensions, deferred old-age pensions and part-time pensions are counted in old-age benefits. The statutory retirement age for old-age pension under the national scheme is 65. The Pension reform came into force at the beginning of the year 2005. The pension entitlement age criterion under the employment scheme has changed from 65 to 63 - 68. Persons secured under the employment scheme are in certain professions entitled to start old-age pensions earlier. In addition, early old-age pensions are awarded after the age of 60 in earliest in public sector contracts and in private sector contracts under the employment scheme. Part-time pensions are awarded to persons after the age 56 in the public sector and after the age of 58 in private sector contracts under the employment scheme. Income on PY110G and PY130G has been reclassified to PY100G according to person's actual retirement to old-age pension (excl. part-time pensions) or last, by using statutory retirement age (65) under the national scheme.	1
Survivors' benefits	PY110G	Benefits that provide temporary or permanent income to people below the retirement age after the death of their spouse, partner or	1

		next-of-kin, usually when the latter represented the main breadwinner for the beneficiary. Survivors' pension to the deceased person's children, to a surviving spouse and under the employment pension scheme to a former spouse have been counted in survivors' benefits.	
Sickness benefits	PY120G	Benefits that replace in whole or in part loss of earnings during temporary inability to work due to sickness or injury.	1
Disability benefits	PY130G	Benefits that provide an income to persons below the standard retirement age whose ability to work and earn is impaired beyond the minimum level laid down by legislation by physical or mental disability. Income for the disabled persons entering or returning to work.	1
Education-related allowances	PY140G	Grants, scholarships and other education assistance received by students.	1
Gross monthly earnings for employees	PY200G		Note: The gender pay gap is calculated by the Wages and Salaries Statistics unit, Statistics Finland

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

Income information is primarily register information, which was linked to the EU-SILC sample persons from the register database, i.e. the Total Income Database (TIDB) maintained by Statistics Finland. TIDB is compiled from register sources maintained by several administrative authorities[5], who are also in charge of the data quality. The sources cover the whole population of Finland. For TIDB, information is further checked in order to ensure the consistency of the data from several sources.

Items which were not available from registers were collected by interviews (1.4 per cent from all gross income and 1.9 per cent from all paid transfers weighted at total households were interviewed). Interviewed items on income were as follows:

- Wages and salaries for persons who have no taxable income in Finland (incl. in PY010G)
- Income from agriculture received by a party to an estate (incl. in PY050G)
- Income from forestry after expenses (incl. in PY050G)
- Interest income taxed at source (incl. in HY090G)
- Pensions from abroad to persons who have no taxable income in Finland (incl. in PY100G)
- Tax-free care allowances and convalescent's grants, unspecified tax-free pensions (incl. in PY130G)
- Small subsidies for studying (incl. in PY140G)
- Maintenance support for children (incl. in HY050G)
- Strike assistance (incl. in HY060G)
- Regular inter-household transfers received (HY080G)
- Regular inter-household transfers paid (HY130G)

Interviewed items were automatically checked and corrected in relation to acceptable values in the Blaise questionnaire on the basis of information received in the course of the interview and further, after the information collection, the checking was continued in order to detect and correct erroneous values (Section 2.3.3 Processing errors). The hot-deck method was used as a stochastic method to

---

[5] Administrative registers are the Personal Tax Register of National Board of Taxes, the Pension Register of the Finnish Centre for Pension, the Pension Register, Social Insurance Register, Rehabilitation Register, Study Aid Register, Housing Allowance Register of the Social Insurance Institution; the Registers of the Education Fund, the Farm Register of the Information Service Centre of the Ministry of Agriculture and Forestry, the Social Assistance Register of the National Research and Development Centre for Welfare and Health (STAKES), the Tax Database of the military injury benefits system of the State Treasury. The main frame for income information is the Personal Tax Register to which other registers give more detailed information, or supplement it by tax-free income information.

impute item non-responses of interest income taxed at source in the component HY090G interest, dividends, profit from capital investments in unincorporated business to the households (Section 2.5. Imputation procedure).

Except small differences due to interviewed data collection (Section 2.5 Imputation procedure), the sources and procedures for producing income target variables were consistent to the statistical units selected for the cross-sectional and longitudinal surveys. Consistency also means that income information of the longitudinal sample was processed equally irrespective of the rotational group.

3.2.3 The form in which income variables at component level have been obtained (Table 3.3)

Except for the target variables HY020, HY022 and HY023, the target variables on income are in gross amounts.

Table 3.3 Components of income. Finland's sources or procedures used for collection of income components, the form and the methods used for obtaining the target variables in the 2006 survey.

	Variable name	Source or procedure used for collection	The form	The method used for obtaining the target variable
Total household gross income	HY010	The register database, the IDS/EU-SILC interview	Gross value	The sum for all household members of gross personal income components (PY010G, PY020G (company cars), PY030G, PY50G, PY070G, PY090G, PY100G, PY110G, PY120G, PY130G, PY140G) plus gross income components at household level (HY040G, HY050G, HY060G, HY070G, HY080G, HY090G, HY110G)
Total disposable household income	HY020	The register database, the IDS/EU-SILC interview	Net value	The sum for all household members of gross personal income components (PY010G, PY020G (company cars), PY030G, PY50G, PY070G, PY090G, PY100G, PY110G, PY120G, PY130G, PY140G) plus gross income components at household level (HY040G, HY050G, HY060G, HY070G, HY080G, HY090G, HY110G) minus regular taxes on wealth (HY120G), regular inter-household cash transfers paid (HY130G), tax on income and social insurance contributions (HY140G)
Total disposable household income, before social transfers other than old-age and survivors' benefits	HY022	The register database, the IDS/EU-SILC interview	Net value	The total disposable income (HY020) minus **total gross to net converted transfers** of unemployment benefits (PY090G), sickness benefits (PY120G), disability benefits (PY130G), education-related allowances (PY140G), family/children-related allowances (HY050G), social exclusion not elsewhere classified (HY060G) and housing allowances (HY070G) For net conversion of the social transfers, detailed information from the Personal Tax Register was used. The phases in deriving HY022 and HY023 were as follows: 1. Deductions which are focused on social transfers subject to taxation were counted in order to obtain taxable social transfers. Deductions of the state and municipal taxation were done separately. 2. Taxes paid on taxable social transfers in state and municipal taxation were deducted. These are the actual taxes paid defined by the rate of the taxed social transfers and taxed earned income (incl. social transfers in the Finnish taxation). 3. The gross to net converted social transfers subject to taxation and social transfers not subject to taxation excluding and including old-age benefits and survivors' benefits were deducted from HY020, resulting in HY022 and HY023.

Total disposable household income, before social transfers including old-age and survivors' benefits	HY023	The register database, the IDS/EU-SILC interview	Net value	The total disposable income (HY020) minus **total gross to net converted transfers** of unemployment benefits (PY090G), old-age benefits (PY100G), survivors' benefits (PY110G), sickness benefits (PY120G), disability benefits (PY130G), education-related allowances (PY140G), family/children-related allowances (HY050G), social exclusion not elsewhere classified (HY060G) and housing allowances (HY070G) See the method of HY022.
Imputed rent	HY030G	**The stratification method** has been used for imputing equivalent gross rent values to the EU-SILC sample dwellings from the external data source. The external data source is Rent statistics for which information is collected by monthly Labour Force Survey interviews (the whole sample size is 12,000), and from register sources maintained by Statistics Finland. Rent statistics are compiled by a **conventional method based on classification and regression analysis (hedonic method).** The available data from the statistics include mean rents/m2 for dwellings in different sizes, types, and areas. Source for repurchase prices: Federation of Finnish Insurance Companies, Finland's Tax Act The IDS/EU-SILC interviewed data on sample household dwellings. The HBS interviewed data (for imputing insurance for detached houses)	Gross value	**Stratification method:** Information about mean rent / m2 (incl. utility costs which is not separable from "the space rent", incl. new and old contracts) of privately financed rented dwellings was imputed from the Rent statistics for the floor area of the sample households' main dwelling by using the following groups: - HH010: Type of building (detached houses with 1-2 dwellings and other type of buildings, semi-detached or terraced house, block of flats) - HH030: Number of rooms (1, 2, 3, 4+) - Construction or renovation year (-60, 61-70, 71-80, 81-90, 91-) - Statistical grouping of municipalities (urban / other) Since the base year, the mean rent (i.e. a price index) of the Rent Statistics by statistical grouping of municipalities has been annually extrapolated to the base year rents by the strata, and imputed to the equivalent sample dwellings. To obtain the value of imputed rent, costs on housing the household actually paid (rents, maintenance and repair of the dwelling, electricity, gas and other fuels, incl. subsidies received for them) were subtracted from the value. Further, depreciation of detached houses was imputed for the equivalent dwellings by stratifying, and subtracted from the value. Depreciation was imputed to detached houses according to the following strata: - Floor area m2 available to households (<60, 60-89, 90-124, 124-) - Construction or renovation year (-50, 50-64, 64-) - Construction material (wood, other) - Statistical grouping of municipalities (urban / other)
Income from rental of property or land	HY040G	Register database	Gross value	
Family/children-related allowances	HY050G	Items either from the Register database or from the IDS/EU-SILC interview	Gross value	
Social exclusion payments not elsewhere classified	HY060G	Items either from the Register database or from the IDS/EU-SILC interview	Gross value	
Housing allowances	HY070G	Items either from the Register database or from the IDS/EU-SILC interview	Gross value	
Regular inter-household cash transfers received	HY080G	The IDS/FI-SILC interview	Gross value	
Interest, dividends, profit from capital investments in unincorporated businesses	HY090G	Items either from the Register database or from the IDS/EU-SILC interview	Gross value	Item non-responses of interest income taxed at source were imputed for the households that responded in the interview that they had received the income during the income reference year, but did not specify the exact amount. Imputing was done in two phases: first, to the households with the answered range value and second, to ones with missing value. Hot-deck method was used for this. Grouping variables were domicile code (dwelling location) and range value in the first, and domicile code, socio-economic status and the number of household members in the second phase imputation. (Chapter 2.5.)

				Imputation method for the rotational groups (DB075) 2 and 3 differs from the group 4 of the longitudinal survey and the cross-sectional survey. (Chapter 2.5.) Comparability over time: Change for the two phase imputation procedure was due to the change in data collection which was done in order to improve interviewed data completeness. As a result, **the number of small values on interest income taxed at source and the households received the income increased markedly from the 2005 cross-sectional survey. The effect on the total income amount estimate of this income item was however slight. The item consisted of about 10 per cent of the whole HY090G component. The increase of the item amount indicating both the actual change and change due to measurement errors (e.g. change in data collection) was 21 per cent from the previous year.**
Interest paid on mortgages	HY100G	Register database	Gross value	
Interest paid on mortgages	HY100N	Register database Deductive imputation by using Personal Tax Register information of National Board Taxes.	Gross value	
Income received by people aged under 16	HY110G	Register database	Gross value	
Regular taxes on wealth	HY120G	Register database	Gross value	The item of tax on real property was derived by using data on the taxes paid on the real property owned in the year (t-1) before the income reference period (t) and data on change of taxable value of the real property from the year t-1 to the income reference period t.
Regular inter-household transfers paid	HY130G	The IDS/EU-SILC interview	Gross value	
Tax on income and social insurance contributions	HY140G	Register database	Gross value	
Repayments/receipts for tax adjustments	HY135G	-	-	-
Cash or near-cash employee income	PY010G	Register database	Gross value	
Non-cash employee income	PY020G	Register database	Gross value	
Employers' social insurance contributions	PY030G	-	-	-
Cash profits or losses from self-employment (including royalties)	PY050G	Register database, the IDS/EU-SILC interview.	Gross value	The component includes the gross item of timber selling as earned and capital forestry income, for which expenses were imputed by using the parameter estimates of the regression model of the expenses based on the IDS data on the earlier year. For imputing the parameters of the expenses values, the following variables were used in a linear scale: - Forestry income from timber selling - Subsidies for forest improvement - Forestry levy - Forest area
Value of goods produced for own consumption	PY070G	-	-	-
Unemployment benefits	PY090G	Register database	Gross value	
Old-age benefits	PY100G	Register database and the IDS/EU-SILC interview data	Gross value	Survivors' benefits and disability benefits which were received simultaneously with old-age benefits were regrouped into old-age benefits by using the retirement age of the national pension scheme, which is 65.
Survivors' benefits	PY110G	Register database	Gross value	
Sickness benefits	PY120G	Register database	Gross value	

Disability benefits	PY130G	Register database	Gross value	
Education-related allowances	PY140G	Register database and the IDS/EU-SILC interview	Gross value	
Gross monthly earnings for employees	PY200G	-	-	-

3.2.4 The method used for obtaining the income target variables

See the previous chapters 3.2.1, 3.2.2, 3.2.3 and Table 3.3, the column on the method used for obtaining the target variables.

3.3 Tracing rules

The tracing rules for the follow-up of sample persons, sample households and co-residents have been followed in the longitudinal survey according to the EU-SILC requirements framework. Because of the sampling design and the sampling unit definition used (the selected individuals), only the initial sample persons of the first wave are followed over the survey years. Households are constructed and household members are defined (mostly co-residents, see the household membership definition) around these sample persons. Household members include the ones who were currently living in the households containing the initial sample person or who were temporarily absent from that household at the end of the income reference period (31 December). Membership status is checked in the each wave.

4 Coherence

4.1 Comparison of income target variables and number of persons who receive income from each income component with external sources

Tables 4.1 - 4.3 show results from income comparisons with relevant data sources. They are the Income Distribution Statistics (IDS), Total Statistics on Income Distribution (TSID) and National Accounts (NA) by Statistics Finland. IDS is the primary national source for the household income statistics. TSID is compiled from the Total Income Database (TIDB) which is used as a register income source both for IDS and EU-SILC. The TSID household definition is based on the household dwelling unit, not the housekeeping unit like in the sample statistics IDS and EU-SILC.

Social transfers received have been compared with the social expenditure on cash benefits by main group from the European System of Integrated Social Protection Statistics (ESSPROS) compiled by the National Research and Development Centre for Welfare and Health (STAKES), Finland. Social transfers of ESSPROS cover also those ones paid to the persons in institutional care (incl. pensions) and the persons permanently resident abroad, but entitled to benefits (e.g. employees and their family members). Benefits in kind (e.g. institutional care for children, young people and elderly) are not in the figures except housing allowances.

The differences in total income amounts between the statistics are mostly due to differences in items defined to the components. Almost all of the income information is collected to the EU-SILC sample units from TIDB. Further, the EU-SILC data is estimated to the private households by using information on crucial demographic and income variables from TIDB in the sampling and the weightings (Section 2.1). Therefore, inconsistencies between the estimated EU-SILC and TSID income are primarily resulting from the unit-non responses among the units having received certain type of register-based income not used in the weightings (see below). Interviewed information again completes the register information on income, and as a result from this part, the income is slightly more complete in EU-SILC than in TSID.

The EU-SILC and IDS income data is processed equally in the integrated statistical survey. The sample and the frame households are same. Small differences between these two statistics are caused by income definitions. IDS includes profits from sales. All items of gross non-cash employee income are included in IDS, not only income from company cars. Inter-household transfers paid except a compulsory child support are not in IDS.

Compared with the ESSPROS (Table 4.1) and with the TSID social benefits in more detail (Table 4.2), definitions and used classifications have an effect on the figures.

The definitions cause differences between EU-SILC and ESSPROS statistics in the following income components, PY110G, PY120G, PY130G, HY070G. Sick pay which is included in EU-SILC PY010G, not in PY120G, consists of 53 per cent of all sickness cash benefits in ESSPROS. PY110G survivors' benefits and PY130G sickness benefits have not been grouped to PY100G old age benefits after statutory retirement age in ESSPROS like in EU-SILC. From housing allowances which have been counted in HY070G, pensioner's housing allowances are as a part of old age benefits in kind, whereas students' housing supplements have not been included in ESSPROS.

In addition to estimation, under-coverage in relation to ESSPROS in particular is also due to the reference population (See above). The effect of the benefits received in resident in collective households and institutions included in ESSPROS can be supposed to be relatively small on the basis of the estimated number of these persons (Chapter 3.1). Information on these and social benefits paid abroad is not available as a separate statistical data from ESSPROS.

The differences from comparing income recipients by main income components in Table 4.3 are caused by the same factors as in total income sums. Further, the household definition used in the sample statistics and TSID has also an effect on the figures.

Table 4.4. presents the number of income receivers according to the gross income components in the cross-sectional and longitudinal EU-SILC surveys. Mean income and standard errors have been reported in the Chapter 2.2. Compared to the general differences resulting from initial wave non-response and attrition, the differences appear as relatively more marked in income components for which non-unit response rates were higher and income calibrating could not be used for correcting. Variance is higher in the longitudinal than cross-sectional income components in general which also has an effect on the differences.

Table 4.1a   The total gross income of private households in the income reference year 2005 according to different data sources (cross-sectional surveys: EU-SILC; IDS; TSID; NA; ESSPROS)

	EU-SILC (X)	IDS	Difference	Difference	Notes
Income components	**Sum (EUR 1 000)**	**Sum (EUR 1 000)**	**Sum (EUR 1 000)**	**%**	
2.1. Gross employee income (py010g, py020g)	60 569 174	60 855 686	-286 512	-0.5	IDS: Employee income received by persons aged under 16, and other non-cash employee income than company car is included.
2.2. Self-employment income	5 625 490	5 581 608	43 882	0.8	IDS: Employee income received by persons aged under 16 is included.
2.3. Imputed rent	..	..			IDS: Not a separate income component. Imputed rent is included in other income components.
2.4. Property income (hy040g, hy090g, py080g)	5 415 728	14 390 845	.	.	IDS: Imputed rent (incl. interest payments) to own-occupied dwellings and profits from sales are in property income.
excl. imputed rent	5 415 728	7 965 883	-2 550 115	-32.0	Profits from sales cause mostly the difference.
2.5. Current transfers received	25 758 606	26 470 930	-712 324	-2.7	IDS: Imputed rent to dwellings rented from other household is in current transfers received. Inter-household transfers received have been defined as more widely in IDS than in EU-SILC.
excl. imputed rent	25 758 606	26 256 129	-497 523	-1.9	
2.6. Other income received	145 126	..	.	.	The income is included in other IDS income components.
2.7. Interest payments	1 323 295	..	.	.	
2.8. Current transfers paid	24 942 752	25 381 486	-438 734	-1.7	IDS: Inter-household transfers paid except compulsory child support is not included. Taxes paid on profits from sales are included.
Total disposable household income (incl. imputed rent)	..	81 917 800	.	.	
Total disposable household income (excl. imputed rent, positive values)	72 573 104	75 278 259	-2 705 155	-3.6	The difference is due to other non-cash employee income (-), profits from sales (-) and inter-household transfers paid except a compulsory child support (-).
	EU-SILC (X)	TSID	Difference	Difference	
Income components	**Sum (EUR 1 000)**	**Sum (EUR 1 000)**	**Sum (EUR 1 000)**	**%**	
2.1. Gross employee income (py010g, py020g)	60 569 174	60 768 472	-199 298	-0.3	TSID: Employee income received by persons aged under 16, and other non-cash employee income than company car is included.
2.2. Self-employment income	5 625 490	5 472 257	153 232	2.8	TSID: Employee income received by those aged under 16 is included.
2.3. Imputed rent	..	..			
2.4. Property income	5 415 728	7 318 265	-1 902 536	-26.0	TSID: Profits from sales are included.
2.5. Current transfers received	25 758 606	24 449 113	1 309 493	5.4	TSID: All inter-household transfers received are not included.
2.6. Other income received	145 126	..	.	.	The income is included in other TSID income components.
2.7. Interest payments	1 323 295	..	.	.	
2.8. Current transfers paid	24 942 752	24 803 273	139 479	0.6	TSID: Inter-household transfers paid are not included. Taxes paid on profits from sales are included.
Total disposable household income (incl. imputed rent)	..	..			
Total disposable household income (excl. imputed rent, positive values)	72 573 104	73 205 306	-632 202	-0.9	In addition to estimation of EU-SILC, the difference is mostly due to profits from sales and inter-household transfers not included in TSID, and other non-cash employee income

	EU-SILC (X)	NA	Difference	Difference	included in TSID.
Income components	**Sum (EUR 1 000)**	**Sum (EUR 1 000)**	**Sum (EUR 1 000)**	**%**	
2.1. Gross employee income (py010g, py020g)	60 569 174	61 936 000	-1 366 826	-2.2	
2.2. Self-employment income	5 625 490	6 002 000	-376 510	-6.3	NA: B4N-B43N, B3N. Self-employment income, excl. imputed rent, mixed income.
2.3. Imputed rent	7 931 182	3 714 000			NA: Imputed rent of owner occupied dwellings (net), depreciation, mortgage interests, fiscal services included. Owner-occupied main and free-time residences.
Property income					
2.4. Property income excl. imputed rent	5 415 728	8 563 000			NA: D4R
2.5. Current transfers received	25 758 606	27 562 000	-1 803 394	-6.5	
2.6. Other income received	145 126				NA: Item is included in other income components
2.7. Interest payments	1 323 295	1 602 000	-278 705	-17.4	NA: Includes fiscal services
2.8. Current transfers paid	24 942 752	28 114 000	-3 171 248	-11.3	
Total disposable household income	72 573 104	72 948 000	-374 896	-0.5	

	EU-SILC (X)	ESSPROS	Difference	Difference	
Income components	**Sum (EUR 1 000)**	**Sum (EUR 1 000)**	**Sum (EUR 1 000)**	**%**	
PY090G. Unemployment benefits	3 581 000	3 388 000	193 000	5.7	
PY100G. Old-age benefits	13 276 000	12 192 000	1 084 000	8.9	ESSPROS includes pensioners' housing allowances (benefits in kind), it does not include income received from PY110G and PY130G to the persons after the standard retirement age.
PY110G. Survivors' benefits	388 000	1 466 000	-1 078 000	-73.5	See PY100G
PY120G. Sickness benefits	498 000	1 857 000	-1 359 000	-73.2	ESSPROS includes sick pay which has been counted in PY010G employee income
PY130G. Disability benefits	3 388 000	3 821 000	-433 000	-11.3	See PY100G.
PY140G.Education-related allowances	569 000	..	.	.	
HY050G. Family/children-related allowances	2 492 000	2 568 000	-76 000	-3.0	ESSPROS includes the income maintenance benefits paid in the event of child birth and the parental leave benefits which are in PY010G employee income.
HY060G. Social exclusion payments not elsewhere classified	369 000	439 000	-70 000	-15.9	ESSPROS includes wage quarantee which is in PY010G employee income.
HY070G. Housing allowances	891 000	437 000	454 000	103.9	ESSPROS does not include students' housing supplements or pensioner's housing allowances. Housing consists of benefits in kind only.
Total, excl. education-related allowances	24 883 000	26 168 000	-1 285 000	-4.9	
Same definitions in accordance with ESSPROS:					
HY070G. Housing allowances	411 000	437 000	-26 000	-5.9	
PY100G,PY110G,PY130G	17 052 000	17 479 000	-427 000	-2.4	

.. Information is not available

. Information is not logical

Table 4.1b   The total gross income of private households in the income reference year 2005 according to the EU-SILC (L) longitudinal survey and TSID

	EU-SILC (L)	TSID	Difference	Difference	
Income components	Sum (EUR 1 000)	Sum (EUR 1 000)	Sum (EUR 1 000)	%	
2.1. Gross employee income	62 380 596	60 768 472	1 612 124	2.7	TSID: Employee income received by persons aged under 16, and other non-cash employee income than company car is included.
2.2. Self-employment income	5 481 695	5 472 257	9 437	0.2	TSID: Employee income received by those aged under 16 is included.
2.3. Imputed rent	8 224 342		..		
2.4. Property income	5 438 800	7 318 265	-1 879 465	-25.7	TSID: Profits from sales are included.
2.5. Current transfers received	25 186 996	24 449 113	737 882	3.0	TSID: All inter-household transfers received are not included.
2.6. Other income received	89 236		..		The income is included in other TSID income components.
2.7. Interest payments	1 339 945		..		
2.8. Current transfers paid	25 423 856	24 803 273	620 583	2.5	TSID: Inter-household transfers paid are not included. Taxes paid on profits from sales are included.
Total disposable household income (incl. imputed rent)	..	..			
Total disposable household income (excl. imputed rent, positive values)	73 153 810	73 205 306	-51 495	-0.1	In addition to estimation of EU-SILC, the difference is mostly due to profits from sales and inter-household transfers not included in TSID, and suchlike non-cash employee income included in TSID.

Table 4.2 Income items of social benefits in the income reference year 2005 according to the EU-SILC(X) cross-sectional survey and TSID

	EU-SILC	TSID	Difference	Difference	
Income components	Sum (EUR 1 000)	Sum (EUR 1 000)	Sum (EUR 1 000)	%	
PY090G. Unemployment benefits	3 581 451	3 439 143	142 308	4.1	
PY100G. Old-age benefits	13 276 405	11 674 848	1 601 557	13.7	TSID includes pensioners' housing allowances (benefits in kind), it does not include income received from PY110G and PY130G to the persons who are on old-age pensions after the standard age.
PY110G. Survivors' benefits	387 620	1 312 680	-925 060	-70.5	See PY100G.
PY120G. Sickness benefits	498 484	459 871	38 614	8.4	
PY130G. Disability benefits	3 387 843	3 379 391	8 451	0.3	See PY100G
PY140G. Education-related allowances	569 126	422 671	146 456	34.7	TIDS does not include interviewed items. Certain differences in classification.
HY050G. Family/children-related allowances	2 492 255	2 412 837	79 418	3.3	
HY060G. Social exclusion payments not elsewhere classified	369 192	425 742	-56 549	-13.3	
HY070G. Housing allowances	890 837	921 931	-31 093	-3.4	

Table 4.3 The number of income recipients in the income reference period 2005 in EU-SILC (X: cross-sectional survey, L: longitudinal survey), IDS and TSID

	EU-SILC (X)	IDS	Difference	EU-SILC (X)	IDS	Difference
Income components	Households (1 000)	Households (1 000)	%	Persons (1 000)	Persons (1 000)	%
2.1. Gross employee income	1 684	1 694	-0.6	2 681	2 736	-2.0
2.2. Self-employment income	390	370*	5.6	466*	434	7.4
2.3. Imputed rent	..	..	.	..	..	..
2.4. Property income	1 991	2 136	-6.8	.	.	.
excl. imputed rent	1 991	2 000	-0.4	.	.	.
2.5. Current transfers received	2 075	2 155	-3.7	.	.	.
excl. imputed rent	2 075	2 149	-3.4	.	.	.
2.6. Other income received	59	..	.	.	.	.
2.7. Interest payments	774	..	.	.	.	.

2.8. Current transfers paid	2 394	2 393	0.1	.	.	.

	EU-SILC (X)	TSID	Difference	EU-SILC (X)	TSID	Difference
Income components	Households (1 000)	Households (1 000)	%	Persons (1 000)	Persons (1 000)	%
2.1. Gross employee income	1684	1678	0.3	2 681	2 694	-0.5
2.2. Self-employment income	390	362	7.8	466	439	6.4
2.3. Imputed rent	..	..		..	..	..
2.4. Property income	1991	1581	26.0	.	.	.
2.5. Current transfers received	2 075	1 999	3.8	.	.	.
2.6. Other income received	59	..	.	.	.	.
2.7. Interest payments	774	..	.	.	.	.
2.8. Current transfers paid	2 394	2 348	2.0	.	.	.

	EU-SILC (L)	IDS	Difference	EU-SILC (L)	IDS	Difference
Income components	Households (1 000)	Households (1 000)	%	Persons (1 000)	Persons (1 000)	%
2.1. Gross employee income	1 705	1 694	0.7	2 676	2 736	-2.2
2.2. Self-employment income	389	370*	5.1	480	434	10.5
2.3. Imputed rent	..	..	.	..	..	..
2.4. Property income	2 022	2 136	-5.3	.	.	.
excl. imputed rent	2 022	2 000	1.1	.	.	.
2.5. Current transfers received	2 050	2 155	-4.9	.	.	.
excl. imputed rent	2 050	2 149	-4.6	.	.	.
2.6. Other income received	50	..	.	.	.	.
2.7. Interest payments	787	..	.	.	.	.
2.8. Current transfers paid	2 383	2 393	-0.4	.	.	.

	EU-SILC (L)	TSID	Difference	EU-SILC (L)	TSID	Difference
Income components	Households (1 000)	Households (1 000)	%	Persons (1 000)	Persons (1 000)	%
2.1. Gross employee income	1 705	1 678	1.6	2 676	2 694	-0.7
2.2. Self-employment income	389	362	7.4	480	439	9.3
2.3. Imputed rent	..	..	.	..	..	..
2.4. Property income	2 022	1 581	27.9	.	.	.
2.5. Current transfers received	2 050	1 999	2.6	.	.	.
2.6. Other income received	50	..	.	.	.	.
2.7. Interest payments	787	..	.	.	.	.
2.8. Current transfers paid	2 383	2 348	1.5	.	.	.

* positive values
.. Information is not available
. Information is not logical

Table 4.4 The number of income receivers by the total gross income components in the income reference period 2005 according to the EU-SILC cross-sectional (X) and longitudinal (L) surveys

	EU-SILC (X)	EU-SILC (L)	Difference %	EU-SILC (L)	EU-SILC (L)	EU-SILC (L)
				DB075=2	DB075=3	DB075=4
Households (N)	2 434 999	2 409 758	1.0	2 388 410	2 402 359	2 438 505
Persons aged 16+ (N)	4 208 238	4 183 754	0.6	4 200 076	4 193 205	4 187 211
	Number of statistical units received the income (1 000)	Number of statistical units received the income (1 000)		Number of statistical units received the income (1 000)	Number of statistical units received the income (1 000)	Number of statistical units received the income (1 000)
Income components*	Households (N)	Households (N)	Difference %	Households (N)	Households (N)	Households (N)
HY010	2 435	2 409	1.1	2 388	2 402	2 438
HY020 (incl. negative values)	2 435	2 410	1.0	2 388	2 402	2 439
HY022 (incl. negative values)	2 375	2 364	0.5	2 344	2 365	2 384
HY023 (incl. negative values)	2 334	2 339	-0.2	2 325	2 334	2 358
HY030G	1 689	1 708	-1.1	1 691	1 691	1 741
HY040G	176	172	2.3	163	174	181
HY050G	600	595	0.8	572	602	611
HY060G	212	180	18.1	172	148	219
HY070G	531	505	4.9	473	539	505
HY080G	222	239	-7.0	233	275	208
HY090G	1 985	2 015	-1.5	2 029	2 035	1 981
HY100G	774	787	-1.7	780	769	811
HY110G	59	50	18.0	49	51	49
HY120G	1 024	1 014	1.0	1 010	949	1 083
HY130G	312	363	-13.9	425	379	285
HY140G	2 389	2 380	0.4	2 367	2 381	2 392
HY135G	..	..		..	..	..
	Persons (N)	Persons (N)		Persons (N)	Persons (N)	Persons (N)
PY010G	2 681	2 675	0.2	2 723	2 675	2 662
PY020G	67	72	-7.1	81	79	59
PY030G	4 208		0.6			
PY035G	392	365	7.2	356	349	395
PY050G	466	480	-2.8	467	486	487
PY070G	..	..		..	..	..
PY080G	..	..		..	..	..
PY090G	730	614	18.9	592	587	666
PY100G	990	976	1.4	983	966	982
PY110G	74	60	22.2	59	78	41
PY120G	239	242	-1.3	238	239	252
PY130G	377	362	4.3	338	382	369
PY140G	429	436	-1.5	431	449	447
PY200G	..	..	.	..	..	..

.. Information is not available

* Income receivers on HY030G and HY100G is not included in the total income components

4.2 Comparison of labour target variables with Labour Force Survey (LFS)

The differences between the EU-SILC self defined current activity status (PL030) and the LFS activity status are logical to their definitions. Compared with EU-SILC, LFS uses the ILO concept which is more detail in relation to the employment and unemployment definitions in particular. After deriving more comparable unemployment definition with LFS by using information on actively looking for a job (PL020) and availability for work (PL025) in addition to self defined current activity status, EU-SILC results less persons in labour force groups and consequently, more persons not in labour force groups (Table 4.5.) Employment is prioritised, but not as definitively as in LFS. In the interview, one hour working or temporary absence from work was not so strictly considered as working, although the latter criterion (temporary absence) had been provided in the survey question definitions and interview guidelines. Otherwise, the perception of own activity was based on the one activity from all other activities in which a person had been involved during the reference period.

There are also differences in reference time periods which may explain the differences between the variable frequencies. The whole December was the time reference period in EU-SILC, whereas it was used one week periods over the whole December as the references periods in LFS. The LFS estimates are the averages of these reference periods.

Table 4.5 Self defined current activity status (PL030) completed by information on looking for a job (PL020) and availability for a job (PL025), persons of aged 16-64 (LFS: persons aged 15-64) in December 2005 according to the EU-SILC cross-sectional (X) and longitudinal (L) surveys and LFS, %

	EU-SILC (X) (December)	EU-SILC (L) (December)		EU-SILC (X) (December)	EU-SILC (L) (December)	LFS (December)
PL030 Self defined activity status						
1. Working full time	58.9	59.2	Working full time or part time	66.7	66.2	67.8
2. Working part time	7.9	7.0				
3. Unemployed	7.3	6.5	PL020 & PL025. Without work, actively looked for a job in previous four weeks and available for work in the next two weeks	5.0	4.6	5.6
In labour force				71.8	70.8	73.4
4. Pupil, students, further training etd.	11.0	11.5				10.2
5. In retirement or in early retirement or has given up business	3.1	4.1				3.8
6. Permanently disabled or/and unfit to work	7.1	7.0				6.3
7. In compulsory military or community service	0.7	0.6				0.5
8. Fulfilling domestic tasks and care responsibilities	3.5	3.6				2.6
9. Other inactive persons	0.7	0.5				3.2
Not in labour force				28.2	29.2	26.6
Total	**100.0**	**100.0**		**100.0**	**100.0**	**100.0**
Number of persons	**3 323 815**	**3 369 343**		**3 323 815**	**3 369 343**	**3 505 400**

Table 4.6 Status in employment , employed persons of aged 16-64 (LFS: persons aged 15-64) in December 2005 according to the EU-SILC cross-sectional (X) and longitudinal (L) surveys and LFS, %

	EU-SILC (X) (December)	EU-SILC (L) (December)	LFS (December)
PL040 Status in employment			
1. Self-employed with employees	4.6	5.2	..
2. Self-employed without employees	8.3	8.5	..
Self employed in total	12.9	13.7	11.8
3. Employee	86.8	86.0	87.7
4. Family worker	0.2	0.3	0.4
Missing	0.0	0.0	0.1
Total	**100.0**	**100.0**	**100.0**
Number of persons	**2 217 503**	**2 232 089**	**2 375 400**