

Central Statistical Bureau of Latvia



**INTERMEDIATE QUALITY REPORT
EU-SILC 2005 OPERATION
IN LATVIA**

Riga 2006

CONTENTS

Background.....	4
1. Common cross-sectional European Union indicators.....	4
2. Accuracy.....	6
2.1. Sampling design.....	6
2.1.1. Type of sample design.....	6
2.1.2. Sampling units.....	6
2.1.3. Stratification criteria.....	6
2.1.4. Sample size and allocation criteria.....	7
2.1.5. Sample selection schemes.....	7
2.1.6. Sample distribution over time.....	8
2.1.7. Renewal of sample: Rotational groups.....	8
2.1.8. Weightings.....	8
2.1.8.1. Design factor.....	8
2.1.8.2. Non-response adjustments.....	9
2.1.8.3. Adjustments to external data (level, variables used and sources).....	9
2.1.8.4. Final cross-sectional weights.....	10
2.1.9. Substitutions.....	10
2.2. Sampling errors.....	10
2.2.1. Standard error and effective sample size.....	10
2.3. Non-sampling errors.....	11
2.3.1. Sampling frame and coverage errors.....	11
2.3.2. Measurement and processing errors.....	11
2.3.2.1. Measurement errors.....	11
2.3.2.2. Processing errors.....	14
2.3.3. Non-response errors.....	15
2.3.3.1. Achieved sample size.....	15
2.3.3.2. Unit non-response.....	15
2.3.3.3. Distribution of households by ‘record of contact at address’ (DB120), by ‘household questionnaire result’ (DB130) and by ‘household interview acceptance’ (DB135).....	16
2.3.3.4. Distribution of substituted units.....	18
2.3.3.5. Item non-response.....	18
2.4. Mode of data collection.....	20
2.5. Interview duration.....	20
3. Comparability.....	21
3.1. Basic concepts and definitions.....	21
3.1.1. The reference population.....	21
3.1.2. The private household definition.....	21
3.1.3. The household membership.....	21
3.1.4. The income reference period.....	22

3.1.5.	The period of taxes on income and social contributions.....	22
3.1.6.	The reference period for taxes on wealth.....	22
3.1.7.	The lag between the income reference period and current variables.....	22
3.1.8.	The total duration of the data collection of the sample.....	22
3.1.9.	Basic information on activity status during the income reference period.....	22
3.2.	Components of income.....	22
3.2.1.	Total household gross income.....	22
3.2.2.	Total disposable household income.....	23
3.2.3.	Total disposable household income before social transfers other than old-age and survivor's benefits.....	23
3.2.4.	Total disposable household income, before social transfers including old-age and survivor's benefits.....	23
3.2.5.	Imputed rent.....	23
3.2.6.	Income from rental property and land.....	23
3.2.7.	Family/children-related allowances.....	23
3.2.8.	Social exclusion payments not elsewhere classified.....	23
3.2.9.	Housing allowances.....	23
3.2.10.	Regular inter-household cash transfers received.....	23
3.2.11.	Interest, dividends, profit from capital investments in unincorporated business.....	23
3.2.12.	Interest paid on mortgages.....	23
3.2.13.	Income received by people aged under 16.....	24
3.2.14.	Regular taxes on wealth.....	24
3.2.15.	Regular inter-household transfers paid.....	24
3.2.16.	Tax on income and social contributions.....	24
3.2.17.	Repayments and tax adjustments.....	24
3.2.18.	Cash or near-cash employee income.....	24
3.2.19.	Non-cash employee income.....	24
3.2.20.	Employers' social contributions.....	25
3.2.21.	Cash profits or losses from self-employment (including royalties).....	25
3.2.22.	Value of goods produced for own consumption.....	25
3.2.23.	Unemployment benefits.....	25
3.2.24.	Old-age benefits.....	25
3.2.25.	Survivors' benefits.....	25
3.2.26.	Sickness benefits.....	26
3.2.27.	Disability benefits.....	26
3.2.28.	Education related allowances.....	26
3.2.29.	Gross monthly earnings for employees.....	26
4.	Coherence.....	27
4.1.	Comparison of income target variables and number of persons who receive income from each 'income component', with external sources.....	27
4.2.	Comparison of other target variables with external sources.....	28

Background

The first collection for the Latvian EU-SILC was conducted in 2005. The Latvian EU-SILC survey is an annual survey with a four-year rotational panel and has been carried out as independent survey, covering both cross-section and longitudinal primary target variables by the single operation.

1. Common cross-sectional European Union indicators

Table 1.1. Laeken indicators and other indicators

Indicator	Value
<i>Primary Laeken indicators of social cohesion</i>	
At-risk-of-poverty rate after social transfers: Total	19.3
At-risk-of-poverty rate after social transfers: Male	18.3
At-risk-of-poverty rate after social transfers: Female	20.0
At-risk-of-poverty rate after social transfers: 0-15 total	20.6
At-risk-of-poverty rate after social transfers: 0-64 total	18.9
At-risk-of-poverty rate after social transfers: 0-64 male	19.2
At-risk-of-poverty rate after social transfers: 0-64 female	18.5
At-risk-of-poverty rate after social transfers: 16+ total	19.0
At-risk-of-poverty rate after social transfers: 16+ male	17.5
At-risk-of-poverty rate after social transfers: 16+ female	20.2
At-risk-of-poverty rate after social transfers: 16-64 total	18.5
At-risk-of-poverty rate after social transfers: 16-64 male	18.5
At-risk-of-poverty rate after social transfers: 16-64 female	18.4
At-risk-of-poverty rate after social transfers: 16-24 total	19.5
At-risk-of-poverty rate after social transfers: 16-24 male	19.0
At-risk-of-poverty rate after social transfers: 16-24 female	20.0
At-risk-of-poverty rate after social transfers: 25-49 total	17.1
At-risk-of-poverty rate after social transfers: 25-49 male	16.8
At-risk-of-poverty rate after social transfers: 25-49 female	17.5
At-risk-of-poverty rate after social transfers: 50-64 total	20.3
At-risk-of-poverty rate after social transfers: 50-64 male	21.8
At-risk-of-poverty rate after social transfers: 50-64 female	19.2
At-risk-of-poverty rate after social transfers: 65+ total	21.2
At-risk-of-poverty rate after social transfers: 65+ male	11.6
At-risk-of-poverty rate after social transfers: 65+ female	25.7
At-risk-of-poverty rate after social transfers: At work total	9.0
At-risk-of-poverty rate after social transfers: At work male	8.9
At-risk-of-poverty rate after social transfers: At work female	9.1
At-risk-of-poverty rate after social transfers: Not at work total	30.8
At-risk-of-poverty rate after social transfers: Not at work male	31.0
At-risk-of-poverty rate after social transfers: Not at work female	30.7
At-risk-of-poverty rate after social transfers: Unemployed total	58.5
At-risk-of-poverty rate after social transfers: Unemployed male	63.6
At-risk-of-poverty rate after social transfers: Unemployed female	53.0

Indicator	Value
At-risk-of-poverty rate after social transfers: Retired total	23.8
At-risk-of-poverty rate after social transfers: Retired male	18.6
At-risk-of-poverty rate after social transfers: Retired female	26.3
At-risk-of-poverty rate after social transfers: Other inactive total	30.5
At-risk-of-poverty rate after social transfers: Other inactive male	29.1
At-risk-of-poverty rate after social transfers: Other inactive female	31.4
At-risk-of-poverty rate after social transfers: No dependent children	20.3
At-risk-of-poverty rate after social transfers: Single total	40.6
At-risk-of-poverty rate after social transfers: Single male	42.3
At-risk-of-poverty rate after social transfers: Single female	39.8
At-risk-of-poverty rate after social transfers: Single <65 years	36.6
At-risk-of-poverty rate after social transfers: Single 65+	45.2
At-risk-of-poverty rate after social transfers: 2 adults no children, <65 years	18.7
At-risk-of-poverty rate after social transfers: 2 adults no children, 65+	11.2
At-risk-of-poverty rate after social transfers: Other, no dependent children	13.5
At-risk-of-poverty rate after social transfers: All households with dependent children	18.4
At-risk-of-poverty rate after social transfers: Single parent	32.5
At-risk-of-poverty rate after social transfers: 2 adults 1 dependent child	13.9
At-risk-of-poverty rate after social transfers: 2 adults 2 dependent children	17.8
At-risk-of-poverty rate after social transfers: 2 adults 3+ dependent children	37.5
At-risk-of-poverty rate after social transfers: Other households with children	13.7
At-risk-of-poverty rate after social transfers: Owner or rent-free	18.2
At-risk-of-poverty rate after social transfers: Tenant	23.8
S80/S20 income quintile share ratio	6.7
Relative median at-risk-of-poverty gap: Total	27.2
Relative median at-risk-of-poverty gap: Male	33.3
Relative median at-risk-of-poverty gap: Female	23.4
Relative median at-risk-of-poverty gap: 0-15	31.3
Relative median at-risk-of-poverty gap: 16+ total	25.9
Relative median at-risk-of-poverty gap: 16+ male	33.3
Relative median at-risk-of-poverty gap: 16+ female	21.8
Relative median at-risk-of-poverty gap: 16-64 total	32.7
Relative median at-risk-of-poverty gap: 16-64 male	36.3
Relative median at-risk-of-poverty gap: 16-64 female	29.1
Relative median at-risk-of-poverty gap: 65+ total	10.8
Relative median at-risk-of-poverty gap: 65+ male	12.7
Relative median at-risk-of-poverty gap: 65+ female	10.0
<i>Secondary Laeken indicators of social cohesion</i>	
Dispersion around the risk-of-poverty threshold: 40%	8.4
Dispersion around the risk-of-poverty threshold: 50%	12.4
Dispersion around the risk-of-poverty threshold: 70%	26.3
At-risk-of-poverty rate before all transfers	40.3
At-risk-of-poverty rate before transfers including old-age and survivors' benefits	25.8
Gini coefficient	36.2
<i>Other indicators</i>	
Equivalised disposable income	1818.47

The calculation of gender pay gap is based on other sources than EU-SILC. Wage statistics is used for calculating gender pay gap.

2. Accuracy

2.1. Sampling Design

In Latvia stratified two-stage sampling design was used for EU-SILC survey. At the first stage systematic sampling of the primary sampling units (Population Census counting areas) had been carried out. At the second stage simple random sampling to select secondary sampling units (addresses) had been made. The stratification had been made depending on degree of urbanization of area. The code of administrative territories was used for stratifying.

Table 2.1. Sampling design information

Stratum	1st stage	2nd stage	
	PSUs	SSUs	households
1	292	2 279	2 307
2	136	967	1 008
3	148	990	1 008
4	152	1 456	1 490
All	728	5 692	5 813

2.1.1. Type of sample design

Stratified two-stage sampling was used for EU-SILC survey in Latvia. Systematic sampling with inclusion probabilities proportional to unit size had been carried out at the first stage and simple random sampling had been carried out at the second stage.

2.1.2. Sampling units

The Population Census counting areas were used as primary sampling units (PSUs) at the first stage. In general, all territory of Latvia is covered in lists of population counting areas. PSUs were selected by systematic sampling with inclusion probabilities proportional to population size (number of households) of PSUs.

Addresses were used as secondary sampling units (SSUs). Simple random sampling was used to select SSUs from PSUs selected at first sampling stage. In Latvia several households can be registered in one address. All households and individuals living in the selected address were included in EU-SILC survey.

2.1.3. Stratification criteria

The stratification was made depending on degree of urbanization of area. Riga (the capital city), six largest towns, other towns and rural areas forms four strata. The code of administrative territories was used for stratifying. The stratum is identified in the variable DB050.

2.1.4. Sample size and allocation criteria

According to the Regulation (EC) No 1553/2005 of European Parliament and of the Council of 7 September 2005 amending Regulation (EC) No 1177/2003 concerning Community statistics on income and living conditions (EU-SILC), Annex II in Latvia the minimum effective sample size was 3 750 households. The total gross sample size (number of households) has been made by analysing available resources and considering the output of the survey. The non-response rate was estimated by using the results of EU-SILC pilot survey and Household Budget Survey. To compensate the non-response it was decided to select 5 692 addresses from which 45 addresses had not been used in the survey and there is no information about them. In Latvia more than one household can live in one address. Therefore, there were 5 813 households living in the selected addresses. In case if it was not possible succeed to contact the selected address (f.e. address cannot be located, it was not possible to contact any person living in the address or the address was inaccessible) it was assumed that one household is living in selected address.

The response rates differ very much in each stratum. For this reason addresses were not included with probabilities proportional to stratum size, but the initial sample size was proportional to population size of each stratum. The initial sample size was adjusted according to response rates in each stratum to get the final sample size in each stratum. R_h is the number of persons aged 16 and over living in stratum h as at 01.01.2005., n_h is number of respondents (aged 16 and over) of the stratum h and n_h / R_h (%) is the sampling fraction in the corresponding stratum.

Table 2.2. Sampling fractions in the corresponding stratum

Stratum	R_h	n_h	n_h / R_h (%)
1	630 243	2 605	0.0041
2	337 096	1 452	0.0043
3	362 165	1 512	0.0042
4	600 364	2 702	0.0045

2.1.5. Sample selections schemes

In the first stage 728 Population Census counting areas (PSUs) were selected by systematic sampling with inclusion probabilities proportional to their population size.

Simple random sampling without replacement was used to select 5 692 addresses (SSUs) in sampled PSUs. Non-proportional allocation was used to select SSUs.

2.1.6. Sample distribution over time

Sample distribution over time was not used because EU-SILC survey is organized on annual basis. The number of households successfully interviewed in each month of fieldwork is shown below in *Table 2.3*.

Table 2.3. Sample distribution over time

Month	Number of households	% of surveyed households	Cumulative % of surveyed households
May	1 421	37.0	37.0
June	1 607	41.8	78.8
July	298	7.8	86.5
August	126	3.3	89.8
September	391	10.2	100

2.1.7. Renewal of sample: rotational groups

Rotational sampling design was used for EU-SILC survey. Initially sample consisted from four equal rotational groups (sub-samples). To provide cross-sectional component it was foreseen to drop one group and add the new one in next years of the survey. Unfortunately it was not possible to evaluate properly gross sample size for all sub-samples. The calculated gross sample size for all groups was not sufficient to provide minimum effective net sample size for longitudinal component in next years. Therefore, part of successfully interviewed households of sub-sample included only for 1st year of the survey was included into the sample also in following years. This will have great impact on structure and size of four sub-sample groups.

2.1.8. Weightings**2.1.8.1. Design factor**

The design weights (DB080) for addresses were calculated according the sample design:

$$DB080 = \frac{1}{prob_adr};$$

$$prob_adr = \frac{hhpsupop \cdot psustrat \cdot adrpsus}{hhstrpop \cdot adrp sup},$$

where **prob_adr** - inclusion probabilities of addresses;

hhpsupop - a number of households in each strata's each PSU of all population;

psustrat - a number of the PSUs in each strata of sample;

adrpsus - a number of addresses in each strata's each PSU of sample;

hhstrpop - a number of households in each strata of all population;

adrp sup - a number of addresses in each strata's each PSU of population.

The inclusion probability of the household and the individual is equal to the inclusion probability of the address. The design weights were adjusted for outliers (extremely high design weights) at the address level.

2.1.8.2. *Non-response adjustments*

The design weights adjusted for outliers *desig1_w* were adjusted for non-response (in household level) in each primary sampling unit (PSU) with correction coefficients *k2_k3* and *k4*:

$$k2_k3 = \frac{samplpsu \cdot cov_sum}{restppsu \cdot resp};$$

$$nonrespw = k2_k3 \cdot desig1_w;$$

$$k4 = \frac{m1}{m2};$$

$$nonr_w = nonrespw \cdot k4,$$

where *samplpsu* - a number of households in each PSU of sample;

cov_sum – a number of households useful for survey in each PSU of sample;

restppsu - a number of households in each PSU of sample, which belong to target population;

resp – a number of responded households in each PSU of sample;

m1 – a number of addresses in sample, which have at least one responded household;

m2 – a number of responded households in sample.

2.1.8.3. *Adjustments to external data (level, variables used and sources)*

The adjusted design weights were calibrated (in household level) on the basis of demographic data by breaking it down by degree of urbanization (four groups — strata), 12 age groups (0-15; 16-20; 21-25; 26-30; 31-35; 36-40; 41-45; 46-50; 51-55; 56-60; 61-65; 66+), sex and 6 regions of Latvia (NUTS 3). GREG calibration was used. The final household weights were used both for households and for individuals.

It was planned to calibrate cross-sectional weights on basis of demographic data by breaking it down by degree of urbanization (four groups — strata), 13 age groups (0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12) and sex to get personal cross-sectional weights for children aged 0-12 individual. It has not done temporarily. For children cross-sectional weights we used personal cross-sectional weights.

2.1.8.4. Final cross-sectional weights

The final cross-sectional weights DB090 were calculated as the product of the design factor, non-response adjustment factor and calibration factor:

$$DB090 = nonr_w \cdot g,$$

where **g** - g-weights of the regression estimator.

2.1.9. Substitutions

No substitution was used.

2.2. Sampling errors

2.2.1. Standard error and effective sample size

- **At-risk-of poverty rate and mean equivalised disposable income**

It was assumed that at-risk-of poverty rate is similar to ratio of two totals (ignoring that threshold is estimate from sample). Standard error and design effect for at-risk-of poverty rate were estimated as standard error and design effect for ratio. Standard error was estimated using *jackknife* method. The correction of finite population at PSU level was applied to variance estimate in each stratum. The same methodology was used for estimating standard error and design effect for mean equivalised disposable income.

- **Gini coefficient**

Linearization was applied for *Gini* coefficient. Standard error for *Gini* coefficient was estimated as standard error for total of linearized variable. Standard error was estimated using *jackknife* method. The correction of finite population at PSU level was applied to variance estimate in each stratum.

- **Design effect**

Design effect was calculated as ratio of the variance for sampling design used in EU-SILC and the variance for simple random sampling of households.

- **Software**

The variance estimates and design effect were computed using the software SUDAAN and SPSS.

Table 2.4. Estimates, the standard error and design effect for common cross-sectional EU indicators

Indicator	Value	Achieved sample size	Standard error	Design effect	Effective sample size
At-risk-of-poverty rate after social transfers	19.3	3 843	0.78	1.17	3 285
At-risk-of-poverty rate before all transfers including old-age and survivor's benefits	25.8	3 843	0.86	1.08	3 558
At-risk-of-poverty rate before all transfers	40.3	3 843	0.98	1.22	3 150
Gini coefficient	36.2	3 843	0.77	-	-
Mean equivalised disposable income	1 818.5	3 843	37.31	1.12	3 431

2.3. Non-sampling errors

2.3.1. Sampling frame and coverage errors

Two sampling frames are built for each sampling stage. At the first stage counting areas from the list of Population Census 2000 are used as sampling frame. All territory of Latvia was divided in small territories (smaller than NUTS4) during the Population Census 2000. The list contains information about the number of households in each counting area.

At the second stage sampling frame is built from The Population Register, statistical register of dwellings and statistical register of households.

Second stage sampling frame was built by using the copy of Population Register given at the beginning of year 2004. Both statistical register of dwellings and statistical register of households was also updated by using the Population Register. Thus the time lag between last update of the registers and the moment of actual EU-SILC survey sampling was 10 months.

The over-coverage relates either to misclassified units that are in fact out of scope, or to units that do not exist in practice (i.e. address does not exist or is non-residential address or is unoccupied or not principal residence (DB120 = 23)). Overall, over-coverage rate of total amount of addresses included in EU-SILC survey was 6 % (342 from 5 692 addresses).

The level of under-coverage is not estimated.

Table 2.5. Distribution of over coverage

Type of over-coverage	Number of addresses	Proportion of the over-coverage type, (%)
Address does not exist (DB120=231)	39	11.40
Non - residential address (DB120=232)	217	63.45
Address is unoccupied (DB120=233)	19	5.56
Address is not principal residence (DB120=234)	67	19.59
Total	342	100

2.3.2. Measurement and processing errors

2.3.2.1. Measurement errors

The basic source of measurement errors is the questionnaire established for EU-SILC survey (wording, design etc), as the data is collected only via face-to-face personal interview (PAPI). 3 types of questionnaires have been developed for EU-SILC survey: Household Register (to collect

demographic information about all household members), Household Questionnaire (to collect all information related to household – dwelling costs, housing conditions, income components received at household level etc.), Personal Questionnaire (to collect all needed information for each household member aged 16 and over) and Household List (additional document to record all necessary information about household member for tracing purposes and for linkage with data from administrative registers in the future). The household members' first, second names, contact addresses, phone numbers (fixed and mobile phone numbers) and personal identification codes were recorded in Household List). The questionnaires were developed on basis of experience gained during the EU-SILC pilot survey (i.e. the structure of questionnaire, analysis of Interviewer's reports and results of meetings with interviewers organised after fieldwork execution to discuss questionnaire of EU-SILC pilot survey). The questionnaires of EU-SILC survey were available in Latvian and in Russian (the language of the largest ethnic minority in Latvia).

The interviewers of CSB carried out the fieldwork of EU-SILC survey. For the field staff (interviewers and supervisors) was organised a 2-day intensive training session. The aims of the training were to introduce fieldwork staff with methodology of EU-SILC survey, to instruct interviewers for accurate fieldwork execution of the survey and give them information to motivate respondents for participation in the survey. Special emphasis was placed on survey questions about income. Several tests (including practical interview to fill EU-SILC questionnaires) were developed to check interviewers' knowledge after training session.

To increase response rates several steps had been made to introduce Latvian residents with EU-SILC survey before starting fieldwork. Press release had been prepared, several publications had been made in state and regional newspapers, explanatory interviews had been given in some radio and TV channels to provide publicity of EU-SILC survey. Introduction letter with EU-SILC booklet was sent to selected address to establish first contact with household before interview.

Measurement errors had been detected by analysing Interviewer's reports, by organizing discussions with interviewers after fieldwork execution and by logical checks and verification of received questionnaires. Overall, the topic of EU-SILC survey was very sensitive and important for respondents. Therefore, the respondent's attitude to the survey was very different. Part of respondents had shown distrust to governmental institutions and expressed disbelief in improvement of living conditions in Latvia. Other part of respondent was very optimistic. They saw importance of EU-SILC survey to identify socio economical situation. In many cases the respondent's strong attitude burdened the interview process because people was speaking a lot

about living conditions and quality of life in general and not answering the questions. Several problems have been identified in Interviewer's reports:

- many reference periods were confusing for respondents during interview process because they couldn't focus on particular time period;
- respondents couldn't identify themselves to any particular socio-economical status (f. e. woman in child care leave, unpaid family worker assisting in agricultural production);
- old people had difficulties to tell the year when highest level of education was attained;
- many respondents were not willing to tell truth amounts of income components;
- many respondents couldn't tell annual income amounts and housing costs;
- question related to household's ability to keep home adequately warm was not understandable;
- meaning of subjective rent (the potential monthly market price for non-tenants or for persons renting the dwelling at reduced price) was not understandable;
- old professions (during Soviet time or even before soviet occupation) did not correspond to the current ISCO-88 classification;
- some non-citizens felt offended by question related to citizenship.

Interviewers were also complaining about length of questionnaire covering too much information. And some interviewers remarked difficulties to follow the skips.

Besides, received questionnaires was logically checked and verified. Before data entry 44.9 % of all received questionnaires was completely checked, 46.3 % of questionnaires remained unchecked. In 8.8 % of received questionnaires was checked only occupation and NACE Rev.1 codes. Errors have been found in 52.6 % of received questionnaires. On average, 1.93 mistakes were found in erroneous questionnaires. The most part of mistakes had been found in Personal questionnaire (63.4 %), the least part of mistakes – in Household List (1.4 %).

Most typical errors in Household Register:

- skips were not followed;
- inconsistencies between marked answer alternatives in question about marital status (PB190) and in question about partner living in the same household (PB200);

- incorrect personal numbers of household member's relatives in primary target variables RB220, RB230 and RB240);
- incorrect person's age in primary target variables RL010, RL020, RL030, RL040, RL050, RL060 and RB245.

Most typical errors in Household Questionnaire:

- the skips were not followed;
- in several households telephone bills, cable television bills and irregular maintenance payments (capital investments) were included in total housing costs (HH070);
- incorrect answer alternatives were marked in question related to arrears on mortgage payments (HS010) and utility bills (HS020). The answer alternative "no" was marked instead of answer alternative "doesn't have to make such payments" in cases if household didn't make the payments mentioned above;
- incorrect person's age in primary target variable HY110N;
- incorrect amounts of family related benefits (HY050N) received during income reference period: in some cases respondents named amounts of benefits received after income reference period.

Most typical errors in Personal Questionnaire:

- incorrect profession coding: description of profession didn't correspond to ISCO-88 occupation code (PL050);
- incorrect coding of economical activity status: description of economical activity didn't correspond to NACE Rev.1 code (PL110);
- the skips were not followed;
- there were loads of errors in question related to most recent changes in the individual's activity status (PL180)

The errors possible to correct without respondent's assistance were corrected offhand. In cases if additional information from respondent was needed the questionnaire had been returned Interviewer Section, which contacted respondent or interviewer afterwards.

2.3.2.2. Processing errors

Data Entry Section of CSB has made the data entry. Data entry application in ACCESS program also included primary logical checks. The description including approximately 899 skips/conditions and 808 error warnings was created to establish verification system. 289 from 808 (35.8 %) error checks fixed 2 085 mistakes and warnings.

Finally the files are transformed into Eurostats' standard format and are tested by using the checking program developed by Eurostat. Errors were corrected during the checking procedure.

2.3.3. *Non-response errors*

2.3.3.1. *Achieved sample size*

3 843 households interviews were accepted for the database and used in analysis.

There are 7 913 (from 8 079) persons 16 years and older who are members of the households for which the interview is accepted for the database, and who completed a personal interview.

2.3.3.2. *Unit non-response*

The final response rates were calculated according to formulas given by Eurostat:

- Household non-response rate $NRh = 29$
- Individual non-response rate $NRp = 2$
- Overall non-response rate $*NRp = 30.4$

2.3.3.3. *Distribution of households (original units) by 'record of contact at address' (DBI20), by 'household questionnaire result' (DBI30) and by 'household interview acceptance' (DBI35)*

Table 2.6. Distribution of households by 'record of contact at address' (DBI20) for each rotational group

	Rotational group 1		Rotational group 2		Rotational group 3		Rotational group 4		Total	
	N	%	N	%	N	%	N	%	N	%
Total (DBI20 = 11 to 23)	713	100	1 454	100	1 612	100	1 989	100	5 768	100
Address contacted (DBI20 = 11)	551	77.3	1 312	90.2	1 484	92.1	1 844	92.7	5 191	90.0
Address non-contacted (DBI20 = 21 to 23)	162	22.7	142	9.8	128	7.9	145	7.3	577	10.0
Total address non-contacted (DBI20 = 21 to 23)	162	100	142	100	128	100	145	100	577	100
Address cannot be located (DBI20 = 21)	11	6.8	9	6.3	12	9.4	12	8.3	44	7.6
Address unable to access (DBI20 = 22)	51	31.5	44	31.0	46	35.9	50	34.5	191	33.1
Address does not exist or is non-residential address or is unoccupied or not principal residence (DBI20 = 23)	100	61.7	89	62.7	70	54.7	83	57.2	342	59.3

The description of rotational sampling groups has to be considered (see 2.1.7.). Besides, 45 addresses have not been used and there is no information about them.

Table 2.6. Distribution of addresses contacted by 'household questionnaire result' and by 'household interview acceptance' for each rotational group

	Rotational group 1		Rotational group 2		Rotational group 3		Rotational group 4		Total	
	N	%	N	%	N	%	N	%	N	%
Total (DB130 = 11 to 24)	551	100	1 312	100	1 484	100	1 844	100	5 191	100
Household questionnaire completed (DB130 = 11)	233	42.3	965	73.6	1 162	78.3	1 486	80.6	3 846	74.1
Interview not completed (DB130 = 21 to 24)	318	57.7	347	26.4	322	21.7	358	19.4	1 345	25.9
Total interview not completed (DB130 = 21 to 24)	318	100	347	100	322	100	358	100	1 345	100
Refusal to co-operate (DB130 = 21)	146	45.9	157	45.2	157	48.8	149	41.6	609	45.3
Entire household temporarily away for duration of fieldwork (DB130 = 22)	148	46.5	148	42.7	149	46.3	185	51.7	630	46.8
Household unable to respond (illness, incapacity, etc) (DB130 = 23)	8	2.5	17	4.9	7	2.2	12	3.4	44	3.3
Other (DB130 = 24)	16	5.0	25	7.2	9	2.8	12	3.4	62	4.6
Household questionnaire completed (DB135 = 1 to 2)	233	100	965	100	1 162	100	1 486	100	3 846	100
Interview accepted to database (DB135 = 1)	233	100	963	99.8	1 162	100	1 485	99.9	3 843	99.9
Interview rejected (DB135 = 2)	0	0	2	0.2	0	0	1	0.1	3	0.1

The description of rotational sampling groups has to be considered (see 2.1.7.). Besides, 45 addresses have not been used and there is no information about them.

2.3.3.4. Distribution of substituted units

Substitution was not used.

2.3.3.5. Item non-response

The tables below show the amount following information on each income component at personal and household level:

- percentage of persons/households having received an amount of income (other than 0),
- percentage of persons/households having received an income but no information about amount of the received income have been obtained from the questionnaire (missing value);
- percentage of persons/households providing partial information about income variable in the questionnaire (responding part of questions related to income amounts)

Table 2.7. Distribution of item non-response for income variables collected at household level

Income variable	% of households having received an amount	% of households with missing values (before imputation)	% of households with partial information (before imputation)
Total disposable household income	99.3	0.2	10.7
Total disposable household income before social transfers other than old-age and survivor's benefits	98.2	0.3	10.1
Total disposable household income before social transfers including old-age and survivor's benefits	87.4	0.4	11.2
Net income components at household level			
Income from rental of a property or land	1.7	4.5	0
Interest, dividends, profit from capital investments in unincorporated business	2.3	2.3	1.1
Family/Children related allowances	31.5	0	0
Social exclusion not elsewhere classified	8.7	1.5	0
Housing allowances	4.4	8.2	0
Regular inter-household cash transfer received	11.3	4.4	0
Income received by people aged under 16	1.1	4.8	0
Regular taxes on wealth	51.5	9.6	0
Regular inter-household cash transfer paid	10.2	4.8	0
Repayments/receipts for tax adjustment	12.0	0.7	0.4

Table 2.8. Distribution of item non-response for income variables collected at personal level

Income variable	% of persons 16+ having received an amount	% of persons 16+ with missing values (before imputation)	% of persons 16+ with partial information (before imputation)
Net income components at personal level			
Employee cash or near cash income	49.5	1.4	0.2
Non-cash employee income	1.3	100	0
Contributions to individual private pension plans	0.8	36.4	0
Cash benefits or losses from self-employment	5.5	3.0	0
Pension from individual private plans	0	0	0
Unemployment benefits	2.6	3.4	2.4
Old-age benefits	30.9	0.1	0.1
Survivor's benefits	1.3	0	0
Sickness benefits	5.0	0.5	0
Disability benefits	3.5	0	0
Education-related benefits	2.2	0	0

Missing values of income components were filled using imputation methods. Multiple imputation method in combination with Hot Deck method was chosen for imputation of missing values in EU-SILC survey. The main principle of the Hot Deck method is to use the current data (donors) to provide imputed values for records with missing values.

Before imputation data of households was divided in similar groups by type of dwelling, year the dwelling was built and number of rooms in dwelling. Data of individuals were divided in similar groups by sex, person's family status and person's social status. After this distribution we obtained all groups of households and persons with similar income level. This factor gave better imputation results.

2.4. Mode of data collection

Table 2.9. Distribution of household members aged 16 and over by Data status (RB250) and rotational group

HOUSEHOLD MEMBERS AGED 16 AND OVER (RB245 = 1)

	Total	RB250 = 11	RB250 = 12	RB250 = 13	RB250 = 21	RB250 = 22	RB250 = 23	RB250 = 31	RB250 = 32	RB250 = 33
Total	8 079	7 913	0	0	9	1	63	90	3	0
%	100	97.9	0	0	0.1	0.0	0.8	1.1	0.0	0
Rotational group 1	544	531	0	0	1	0	3	9	0	0
%	100	97.6	0	0	0.2	0	0.6	1.7	0	0
Rotational group 2	2 020	1 974	0	0	0	0	27	17	2	0
%	100	97.7	0	0	0	0	1.3	0.8	0.1	0
Rotational group 3	2 453	2 416	0	0	1	1	12	23	0	0
%	100	98.5	0	0	0.0	0.0	0.5	0.9	0	0
Rotational group 4	3 062	2 992	0	0	7	0	21	41	1	0
%	100	97.7	0	0	0.2	0	0.7	1.3	0.0	0

Table 2.10. Distribution of household members aged 16 and over by Type of interview and rotational group (RB260)

HOUSEHOLD MEMBERS AGED 16 AND OVER ((RB245 = 1) and (RB250 = 11 or 13))

	Total	RB260 = 1	RB260 = 2	RB260 = 3	RB260 = 4	RB260 = 5
Total	7 913	7 374	0	0	79	460
%	100	93.2	0	0	1.0	5.8
Rotational group 1	531	505	0	0	0	26
%	100	95.1	0	0	0	4.9
Rotational group 2	1 974	1 832	0	0	25	117
%	100	92.8	0	0	1.3	5.9
Rotational group 3	2 416	2 243	0	0	30	143
%	100	92.8	0	0	1.2	5.9
Rotational group 4	2 992	2 794	0	0	24	174
%	100	93.4	0	0	0.8	5.8

2.5. Interview duration

Mean duration of household interview: 17 minutes and 50 seconds.

Mean interview duration per household: 58 minutes and 27 seconds.

Thus, mean interview duration per household is lower than the one-hour limit set in Regulation 1177/2003.

3. Comparability

3.1. Basic concepts and definitions

Overall, there are no differences between national interpretations of EU-SILC basic definitions and concepts and common standards set up in Commission regulations and doc. EU-SILC 065/04. Special attention has been paid on definition of household member, which has been described more comprehensively according to the most typical cases faced by interviewers during the EU-SILC pilot survey (see 3.1.3.).

3.1.1. The reference population

There were no divergences from common definition. Persons living in private households within national territory were the reference population of EU-SILC survey.

3.1.2. The private household definition

There were no divergences from common definition.

3.1.3. The household membership

There were no divergences from common definition. Due to the complexity of household membership several practical and comprehensive explanations based on concrete cases (examples) were given. After entering European Union many Latvian residents goes to work abroad but at the same time they are keeping ties with family and plan to return home after some time period (which very often is unspecified). The experience of EU-SILC pilot survey has shown that this situation occurs in many Latvian households. Therefore, it was very important do provide more detailed and strict explanations in which cases person working abroad should be considered a household member.

Other typical case is youngsters who are moving from rural areas to cities for studying. Mostly they are living in cities for study time and go back home in holidays. In this situation proper identification of household membership is very important. Thus person is identified to proper household and he/she is neither excluded from EU-SILC survey nor double counted. In the situation described above for absent persons (who moved out abroad or study in another location) it was prescribed to follow 2 criteria: 1) if person has another private address in another country and 2) how long is person's actual and planed length of absence.

3.1.4. The income reference period

There were no divergences from common definition. In Latvia the income reference period is previous calendar year (2004).

3.1.5. The period of taxes on income and social insurance contributions

According to the Commission Regulation (EC) No 676/2006 of 2 May 2006 amending Regulation (EC) No 1980/2003 implementing Regulation (EC) No 1177/2003 of the European Parliament and of the Council concerning Community statistics on income and living conditions (EU-SILC) as regards definitions and updated definitions Latvia is authorized to not deliver any gross income data before 2007. Thus, no data on income tax and on social contributions was collected.

3.1.6. The reference period for taxes on wealth

See 3.1.4.

3.1.7. The lag between the income reference period and current variables

The lag between end of income reference period and current variables is from 4 to 9 months.

3.1.8. The total duration of the data collection of the sample

Fieldwork (data collection) started in May 2005 and lasted till September 2005. The gross sample size has been increased (additional sampling has been made) during the fieldwork and Interviewer Service has to survey more addresses than it was planned before (so exceeding the planned length of data collection).

3.1.9. Basic information on activity status during the income reference period

There were no divergences from common definitions.

3.2. Components of income

Classification of net income components in national EU-SILC survey is made according to description of doc. EU-SILC 065/04 with exception of income from self-employment (see 3.2.21). As Latvia has derogation to collect gross income components from 2007, there are only net income components collected in 2005.

3.2.1. Total household gross income

As Latvia has derogation to collect gross income components from 2007, the values are not recorded.

3.2.2. Total disposable household income

There are no divergences from common standards. Total disposable household income was calculated by using only net income components.

3.2.3. Total disposable household income, before social transfers other than old-age and survivor's benefits

See 3.2.2.

3.2.4. Total disposable household income, before social transfers including old-age and survivor's benefits

See 3.2.2.

3.2.5. Imputed rent

The variable is not filled. Latvia has not developed model to calculate imputed rent. Only subjective rent was collected in 2005.

3.2.6. Income from rental property and land

There are no divergences from common standards. Only net income component was collected.

3.2.7. Family/children-related allowances

There are no divergences from common standards. Only net income component is recorded although it's not taxable income in Latvia.

3.2.8. Social exclusion payments not elsewhere classified

See 3.2.7.

3.2.9. Housing allowances

See 3.2.7.

3.2.10. Regular inter-household cash transfers received

See 3.2.7.

3.2.11. Interest, dividends, profit from capital investments in unincorporated business

See 3.2.6.

3.2.12. Interest paid on mortgages

The value was not recorded, as it's mandatory to collect this variable from 2007.

3.2.13. *Income received by people aged under 16*

There are no divergences from common standards. Basically there are included wages and salaries received during holidays or out of school time. Only net income component was collected.

3.2.14. *Regular taxes on wealth*

There are no divergences from common standards. Taxes on land and real estate are included in this variable.

3.2.15. *Regular inter-household transfers paid*

See 3.2.7.

3.2.16. *Tax on income and social contributions*

This variable is not recorded as Latvia has derogation to collect it from 2007.

3.2.17. *Repayments and tax adjustments*

See 3.2.6.

3.2.18. *Cash or near-cash employee income*

See 3.2.6.

3.2.19. *Non-cash employee income*

Only non-cash employee income from use of company car for personal purposes was collected in 2005. According to Latvian situation method based on system analyses model has been chosen for calculating employee income from company car. Part of the calculation components has been included in questionnaires and collected directly from respondents: class of the car, year of the car make, annual amount of kilometres driven by the vehicle for private use, company car user's occupation, type of fuel, coverage of fuel costs made by employer (i.e. did the employer pay bills for fuel purchasing).

As the model has been developed also after beginning of fieldwork, several components has not been included questionnaire and they have been collected by making assumptions:

- 1) total annual amount of kilometres driven by vehicle: it was assumed that total annual amount of kilometres driven by vehicle is twice higher than annual amount of kilometres driven for private use;
- 2) restrictions for use of company car: it was assumed that there are no restrictions for personnel of superior management, for other employees it was assumed that there are no restrictions for private use of company car only if employer covers fuel costs;

- 3) coverage of costs related to technical repairs and purchasing car tyres: it was assumed that all costs were covered by employer.

3.2.20. Employers' social contributions

The information is not collected and no feasibility studies are made.

3.2.21. Cash profits or losses from self-employment (including royalties)

The income (or losses) from self-employment are collected in 2 components: 1) income from agricultural production and 2) income of the rest self-employment (except income from agricultural production).

The income from agricultural self-employment was collected in the same way as in Household Budget Survey (HBS). Household member responsible for agricultural production was asked to calculate all income components and expenditures the household had during income reference period. Thus, all self-employment income from agricultural production was counted to responsible household member and amount self-employment income was agricultural profit minus expenditures related to production. There were cases when expenditures were greater than profit and this resulted in minus values.

Second income component (from rest self-employment except agricultural production) was asked to each household member aged 16 years and more in Personal Questionnaire. Respondents were asked to tell amount they gained from self-employment for their own use during the income reference period.

3.2.22. Value of goods produced for own consumption

The value is not recorded

3.2.23. Unemployment benefits

See 3.2.7.

3.2.24. Old-age benefits

See 3.2.6.

3.2.25. Survivors' benefits

See 3.2.6.

3.2.26. *Sickness benefits*

See 3.2.6.

3.2.27. *Disability benefits*

See 3.2.6.

3.2.28. *Education related allowances*

See 3.2.6.

3.2.29. *Gross monthly earnings for employees*

Value is not recorded as Latvia uses wage statistics for calculating gender pay gap.

4. Coherence

This section will compare the EU-SILC data to various external sources: the Household Budget Survey (HBS), the Labour Force Survey (LFS), wage statistics and social protection statistics.

The HBS is a continuous survey of households, which has been carried out since 1995. The annual net sample size is approximately 4 thousand households. The HBS is designed to collect information on income and expenditure of households. The HBS was the source of Laeken indicators up until EU-SILC.

The LFS is a continuous survey, which has been carried out according to a common EU methodology since 1995. The yearly sample size is about 17.7 thousand persons aged 15 - 74. The LFS is the main source for labour market information.

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

The average employee cash or near cash income (PY010) was 168 LVL per month in EU-SILC. The respective figure from wage statistics is lower – 150 LVL per month. Data of EU-SILC survey has been calculated for respondent, who received at least one time employee cash or near cash income (PY010) who has been working as employee at least one month during (PL070). The acquired results show that EU-SILC data exceeded enterprise statistical data on average labour income amount by 12 % in 2004. The higher estimates from EU-SILC are in all likelihood due to the fact that in EU-SILC the average wages and salaries are calculated for persons receiving income, whereas in wage statistics, the unit of enumeration is the job. Thus, income received from second, third etc. jobs is added to the income received from the main job, whereas in wage statistics, the wages from second, third etc. job are counted separately. It should be also taken into account that wage statistics is based on the information provided by the employers and for a certain number of cases it corresponds to that part of wages for which taxes have been paid.

Table 4.1. presents the number of persons receiving an income component as estimated from EU-SILC, the HBS and from additional external sources. It should be taken into account that the HBS part of the income components are obtained only at the household level and for this reason comparisons are made among those income components which are obtained in the same way as EU-SILC. Besides, definitions of income components can vary between sources and for that reason only the components for which comparable enough external data is available, are presented below.

Table 4.1. Number of persons receiving income components by different data sources in 2004

EU-SILC target variable	EU-SILC	HBS	Other sources
Employee cash or near cash income (PY010)	963 113	938 983	877 776 ¹
Old-age benefits (PY100)	498 887	492 708	481 683 ²
Survivor' benefits (PY110)	21 439	23 367	28 786 ³
Disability benefits (PY130)	65 201	-	74 605

¹ Wage statistics² At the end of year. ESSPROS data³ At the end of year. ESSPROS data, recipients all age groups, including aged below 16 years.

The number of people receiving employee income is almost the same in EU-SILC and HBS. The number of people receiving employee income is by 85 000 higher in EU-SILC than it is in the wage statistics. This is not unexpected, given that unofficial work relationships are not included in wage statistics.

4.2. Comparison of other target variables with external sources

Important background indicator is a mean size of household. The official statistics in this area is based on the Population Census data but for the periods between the censuses it is based on calculations according to which the mean household size in 2005 was 2.51 persons. Data on the mean size of households are given in Table 4.2.

Table 4.2. Mean size of household in 2005

	Population statistics	HBS	EU-SILC
Mean size of household, persons	2.51	2.55	2.56

A comparison of data shows that surveys such as HBS and EU-SILC probably under-represent single-person households and other households with a small number of persons for which the risk of failing to make contacts with these households is invariably higher.

A comparison of the breakdown of households by the number of persons therein does not show any substantial differences (Table 4.3). However, single-person households in EU-SILC are somewhat better represented than in the HBS.

Table 4.3. Distribution of households by size in 2005

	EU-SILC		HBS		LFS	
	%	number of households, in thousands	%	number of households, in thousands	%	number of households, in thousands
All households	100	869.2	100	885.0	100	838.4
of which by number of members:						
1 person	25.2	219.2	23.8	210.3	20.3	170.4
2 persons	29.8	259.0	32.9	290.7	30.7	257.3
3 persons	21.7	188.9	20.4	180.2	21.4	179.1
4 persons	14.4	125.1	14.5	128.3	16.6	138.9
5 persons and more	8.9	77.0	8.5	75.5	11.1	92.7

A comparison between the EU-SILC and LFS data shows slightly more remarkable differences. Table 4.4. presents the distribution of population by ISCED levels as estimated from EU-SILC and the LFS. As can be seen, there are differences in overall distribution, but they are not substantial. It should be noted that there is no information about 40.9 thousand persons in corresponding age in EU-SILC (which is 2.4 % of total population in age between 16 and 74 years in EU-SILC), because the Individual Questionnaire was not filled for those persons.

Table 4.4. Distribution of population in age between 16 and 74 years by ISCED level in 2005

	LFS		EU-SILC	
	thousand of persons	%	thousand of persons	%
ISCED 0	10.9	0.6	35.6	2.2
Basic education (ISCED 1 + ISCED 2)	441.4	25.1	392.4	23.7
ISCED 3	877.5	49.3	773.1	46.7
ISCED 4	137.1	7.7	175.4	10.6
ISCED 5	302.4	17.0	273.3	16.5
ISCED 6	4.5	0.3	5.4	0.3
Total	1 779.8	100	1 655.2	100

Table 4.5. presents the distribution of population by self-defined economic status in comparison with LFS estimates. The differences that emerge are probably related to the fact that the main activity status is entirely self-defined in EU-SILC at the time of interview, whereas in the LFS it is also self-defined, but the activity status refers to the last three months.

Table 4.5. Distribution of population in age between 16 and 74 years by self-defined economic status in 2005

	LFS		EU-SILC	
	thousand of persons	%	thousand of persons	%
Missing value	-	-	40.9	2.4
Working	1030.7	57.9	970.8	57.1
Unemployed	128.8	7.2	131.4	7.7
Pupil, student	183.0	10.3	138.3	8.1
In retirement	287.9	16.2	317.5	18.7
Permanently disabled	52.9	3.0	23.3	1.4
Domestic task	90.0	5.1	61.6	3.6
Other inactive	6.5	0.4	16.3	1.0
Total	1779.8	100	1700.1	100

Table 4.6. presents the share of households by the type of dwelling. Again the differences between the two sources are small.

Table 4.6. Distribution of households by the type of dwelling in 2005

	EU-SILC	HBS
Detached house	23.6	25.4
Semi-detached house or terraced house	4.0	4.3
Apartment or flat	72.1	70.1
Other kind of accommodation	0.3	0.2
Total	100	100