

# **Final Quality Report**

**relating to the**

## **EU-SILC 2005 Operation in Denmark**

**Cross-sectional component:  
2005**

**Longitudinal Component:  
Trajectories ending in 2005**

**Version 2**

**By Hans Christian Jørgensen  
hac@dst.dk**

Copenhagen 2008

## 1. COMMON CROSS-SECTIONAL EUROPEAN UNION INDICATORS

### 1.1. Common cross-sectional EU indicators based on the cross-sectional component of EU-SILC

The indicators are calculated using Eurostat's SAS-program available at the CIRCA-website

	Total	Female	Male
At-risk-of-poverty rate after social transfers - total	11,8	12,1	11,6
At-risk-of-poverty rate after social transfers - 0-15 years	10,1	10,1	10,2
At-risk-of-poverty rate after social transfers - 16-24 years	28,9	31,7	26,2
At-risk-of-poverty rate after social transfers - 25-49 years	9,7	9,2	10,2
At-risk-of-poverty rate after social transfers - 50-64 years	4,6	4,3	5,0
At-risk-of-poverty rate after social transfers - 65+ years	17,6	18,4	16,5
At-risk-of-poverty rate after social transfers - 16+ years	12,3	12,6	11,9
At-risk-of-poverty rate after social transfers - 16-64 years	11,0	11,0	11,0
At-risk-of-poverty rate after social transfers - 0-64 years	10,8	10,8	10,8
At-risk-of-poverty rate after social transfers - employed	4,9	4,8	5,0
At-risk-of-poverty rate after social transfers - unemployed	25,8	14,1	38,8
At-risk-of-poverty rate after social transfers - retired	15,8	16,1	15,3
At-risk-of-poverty rate after social transfers - other inactive	28,3	27,7	29,2
At-risk-of-poverty rate after social transfers - single, < 65 years	27,8		
At-risk-of-poverty rate after social transfers - single, 65+ years	21,5		
At-risk-of-poverty rate after social transfers - single Total	25,8	25,4	26,3
At-risk-of-poverty rate after social transfers - 2 adults, no children, both < 65	5,1		
At-risk-of-poverty rate after social transfers - 2 adults, no children, at least one 65+	13,0		
At-risk-of-poverty rate after social transfers - other households without children	0,8		

At-risk-of-poverty rate after social transfers - single parent, at least one child	21,2		
	4,4		
At-risk-of-poverty rate after social transfers - 2 adults, 2 children	4,4		
At-risk-of-poverty rate after social transfers - 2 adults, 3+ children	14,2		
At-risk-of-poverty rate after social transfers - other households with children	4,4		
At-risk-of-poverty rate after social transfers - households without children	14,9		
At-risk-of-poverty rate after social transfers - households with children	8,6		
At-risk-of-poverty rate after social transfers - owner or rent-free	7,5		
At-risk-of-poverty rate after social transfers - tenant	20,6		
At-risk-of-poverty rate after social transfers - households without children, $w = 0$	27,4		
At-risk-of-poverty rate after social transfers - households without children, $0 < w < 1$	5,4		
At-risk-of-poverty rate after social transfers - households without children, $w = 1$	4,9		
At-risk-of-poverty rate after social transfers - households with children, $w = 0$	51,7		
At-risk-of-poverty rate after social transfers - households with children, $0 < w < 0.5$	12,0		
At-risk-of-poverty rate after social transfers - households with children, $0.5 < w < 1$	6,4		
At-risk-of-poverty rate after social transfers - households with children, $w = 1$	5,3		
Relative median at-risk-of-poverty gap - total	15,6	15,9	15,5
Relative median at-risk-of-poverty gap - 0-15 years	18,2		
Relative median at-risk-of-poverty gap - 16-64 years	21,6	21,4	22,1
Relative median at-risk-of-poverty gap - 65+ years	8,1	8,7	7,0
Relative median at-risk-of-poverty gap - 16+ years	15,6	16,0	14,3
Dispersion around the risk-of-poverty threshold - 40%	3,4	3,2	3,6

Dispersion around the risk-of-poverty threshold - 50%	5,7	5,9	5,5
Dispersion around the risk-of-poverty threshold - 70%	19,5	20,6	18,4
At-risk-of-poverty rate before social transfers transfers except old-age and survivors' benefits - total	29,9	31,3	28,4
At-risk-of-poverty rate before social transfers transfers except old-age and survivors' benefits - 0-15 years	25,0	26,7	23,5
At-risk-of-poverty rate before social transfers transfers except old-age and survivors' benefits - 16-64 years	28,5	30,3	26,8
At-risk-of-poverty rate before social transfers - 65+ years	42,2	40,7	44,1
At-risk-of-poverty rate before social transfers transfers except old-age and survivors' benefits - 16+ years	31,1	32,5	29,7
At-risk-of-poverty rate before all social transfers - total	37,9	41,0	34,8
At-risk-of-poverty rate before all social transfers - 0-15 years	25,2	26,8	23,7
At-risk-of-poverty rate before all social transfers - 16-64 years	29,1	30,9	27,3
At-risk-of-poverty rate before all social transfers - 65+ years	93,7	95,8	90,9
At-risk-of-poverty rate before all social transfers - 16+ years	41,1	44,4	37,7
Inequality of income distribution S80/S20 income quintile share	3,5		
Gini coefficient	23,9		

## 1.2. Other indicators

### 1.2.1. *Equivalised disposable income*

Mean equivalised disposable income: 23.419 EURO

### 1.2.2. *The unadjusted gender pay gap*

The gender pay gap is not computed on the basis of EU-SILC.

## 2. ACCURACY

### 2.1. Sample design

Denmark has adopted the 4-year rotational integrated design recommended by Eurostat. The sample is drawn as a sample of persons.

#### 2.1.1. *Type of sampling design (stratified, multi-stage, clustered)*

The sub-samples are sampled by simple random sampling.

2.1.2. *Sampling units (one stage, two stages)*

It is a one stage sample. The sampling unit is the individual person. The household is defined as the household of which the selected person is member at the beginning of the survey year (1 January). The sampling frame is all persons aged 13+. Only households, where selected person are 16 or more at the beginning of the survey year are included in statistics of this year.

2.1.3. *Stratification and substratification criteria*

No stratification.

2.1.4. *Sample size and allocation criteria*

Total number of persons aged 16+ living in private households.....4.272.821  
 Number of addresses in the sampling frame.....2.643.240  
 Size of the sample (selected persons/households).....9 467

About 0.5 % of the total number of households in Denmark are represented in the sample

2.1.5. *Sample selection schemes*

Not applicable, since Denmark uses simple random sampling.

2.1.6. *Sample distribution over time*

-

2.1.7. *Renewal of sample: rotational groups*

The sample of the cross-sectional component of EU-SILC 2005 in Denmark consists of 4 sub-samples. 2 selected in 2003, one selected in 2004 and one selected in 2005. 2.500 persons/households were selected for each panel.

<i>Table 1: Renewal of the 2005 cross-sectional sample</i>				
	<i>Selected 2003</i>	<i>Selected 2004</i>	<i>Selected 2005</i>	Total
Number initially selected	5.000	2.500	2.500	10.000
- selected person out of scope	213	60	0	273
- selected person not 16+	81	61	118	260
Number in the sample	4.706	2.379	2.382	9.467

Notes:

Out of scope includes

-

-

- selected persons, who have moved to a collective household or institution within the country,
- selected persons, who have moved outside the country and
- selected persons, who have died

Selected persons not 16+ includes

- persons below 16 selected for the survey but not yet included cf. section 2.1.2

#### 2.1.8. *Weightings*

The weighting procedure of the cross sectional component 2005 is described in detail in Appendix 1. The weighting procedure of the longitudinal component 2003-5 is described in detail in Appendix 2.

##### 2.1.8.1. Design factor

-

##### 2.1.8.2. Non-response adjustments

-

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

-

##### 2.1.8.4. Final cross-sectional weight

#### 2.1.9. *Substitutions*

No substitution

## 2.2. Sampling errors

### 2.2.1. Standard error and effective sample size

The effective sample size for the cross sectional component is not calculated yet.

	Value	Achieved sample size	Standard error before calibration	Standard error after calibration cal
Total	11,8	15321	0,39	0,08
Men	11,6	7707	0,43	0,07
Woman	12,1	7614	0,48	0,10
0-15 years	10,1	3420	0,74	0,21
16-24 years	28,9	1631	1,25	0,21
25-49 years	9,7	5423	0,43	0,07
50-64 years	4,6	3172	0,45	0,01
65+ years	17,6	1675	1,26	0,23

Note: The figures are calculated by Guillaume OSIER, Eurostat

You will find more detailed information about standard error in appendix 3

## 2.3. Non-sampling errors

### 2.3.1. Sampling frame and coverage errors

The sample frame is persons aged 13+ living in private household according to the Register of Population Statistics of Statistics Denmark (version 1 January 2005). The register is based on Central Population Register (CPR) run by the *Ministry of the Interior*. CPR is updated by the municipalities. The register is a continuously updated register.

Main coverage problems:

- persons living in a private household but registered in the register as living in a collective household at the time of selecting the sub-sample. This group will be under-covered in the sub-sample.
- persons, who after the sub-sample were selected during its lifetime, moved into a private Danish household from a collective household in Denmark or from abroad. This group will likewise be under-covered in the sub-sample:

In theory, these groups should be taken into consideration like persons between 13+ and 15+ at the time of sampling, cf. above, but technically it is difficult, and the number of persons involved is relatively small. The number of immigrant is on a yearly basis less than 1 pct. of the population and the

number of persons living in collective is about 1 pct., primarily persons living in old-age homes and homes for other people, who cannot take care of themselves.

If 2 persons from the same household are selected to a panel, one of them is dropped as a selected person. If a person, who already belongs to a household from an earlier still active panel, is selected, the person is likewise dropped as a selected person. The situation, where a household is selected more than once, is only of theoretical interest. The practical importance is negligible.

### *2.3.2. Measurement and processing errors*

#### *2.3.2.1. Measurement errors*

The data comes from interviews or from registers. Income and demographic data primarily comes from registers, while social data primarily comes from interviews. The questionnaire does not include other questions than the SILC-questions. The questionnaire includes between 40 and 50 questions dependent on the type of household.

Interview-method was telephone interviewing when feasible and postal questionnaire for other households. The questionnaire was programmed in BLAISE. To obtain contact by telephone at least 5 calls was conducted. Households contacted by mail received one reminder, if they did not respond to the first letter.

The interviews were conducted by the interviewers of Statistics Denmark. In addition to their usual training and education, they got a special introduction to the SILC-questionnaire of 2 hours.

#### *2.3.2.2. Processing errors*

The questionnaire is programmed in BLAISE. Several entry controls are built into the questionnaire. The system for processing, checking and editing data is programmed in SAS. Finally, the files are transformed into Eurostat's standard format and tested using the checking program developed by Eurostat.

During the checking procedure errors are corrected.

### *2.3.3. Non-response errors*

#### *2.3.3.1. Achieved sample size was*

Number of households contacted is 9.467

Number of households for which an interview is accepted for the database: 5.957.

Number of persons of 16 years or older, who are members of the households and for whom the interview is accepted for the database: 11.901



If the household part of the interview and the personal interview of household representative is acceptable, all members of the household are accepted for the database also in case unit non-response for the person. The necessary information about his/hers income, activity status etc. is extracted from registers.

### 2.3.3.2.-2.3.3.3 Unit non-response

Table 2 gives an overview of the result. The tables below give the results in more details.

	Number	Percent
Total sample	9467	100,00
- No contact attempt	936	9,89
- refusal to cooperate	994	10,50
- Other reasons for not completed	1580	16,69
Completed	5957	62,92

Two groups of households are not contacted initially.

- *Tough refusers*: Households, who have told, that they do not want to be contacted by Statistics Denmark and households, who have refused to participate in the SILC-project two consecutive years
- *Researcher-protected persons*:: If people do not want to be contacted by researchers, they can get their address in the Central Population Register (CPR), researcher-protected. If the selected person has declared, that we are not allowed to contact the household.

	Number	Percent
Total sample	9467	100,00
- Tough refusers	43	0,45
- Researcher-protected	884	9,34
Households contacted	8540	90,11

Households contacted falls into two groups

- Households, where the telephone number is found
- Households, where the telephone number is not found

	Number	Percent
Total number of households	8.540	100,00
Telephone number found	7.643	89,50
Telephone number not found	897	10,50

Contact attempts can be closed by telephone or by post.

Contact attempts are closed by post if

- a telephone number cannot be found
- telephone contact could not be obtained
- telephone contact was obtained, but the interview person refused to be interviewed by telephone but promised to look at a postal questionnaire, if we send a questionnaire to him.

	Number	Percent
Total number of households	8.540	100,00
Telephone number found , closed by telephone	7.019	82,19
Telephone refusal, closed by post	233	2,73
Telephone contact not obtained, closed by post	391	4,58
Telephone number not found closed by post	897	10,50

	Completed		Not completed	
	Number	Percent	Number	Percent
All households	5957	69,75	2583	30,25
Telephone number found , closed by Telephone	5673	80,82	1346	19,18
Telephone refusal, closed by post	22	9,44	211	90,56
Telephone contact not obtained, closed by post	70	17,90	321	82,10
Telephone number not found closed by post	192	21,40	705	78,60

	All Households		Closed by telephone		Closed by post	
	Number	Percent	Number	Percent	Number	Percent
Total	8531	100,00	7019	100,00	1512	100,00
Completed	5957	69,83	5673	80,82	284	18,78
Refusal	994	11,65	950	13,53	44	2,91
Inness etc.	238	2,79	227	3,23	11	0,73
Other reasons	1342	15,73	169	2,41	1173	77,58

Longitudinal Component 2005 consisted of 2 trajectories ending in 2005. One trajectory covering 2003-2005 and one trajectory covering 2004-2005.

Trajectory 2003-2005 includes two rotational groups selected 2003, namely group 3 and 4.

Trajectory 2004-2005 includes two rotational groups selected 2003, namely group 3 and 4 selected 2003 and group 1 selected 2004.

Households are only included in a longitudinal trajectory if DB130 = '11' for all the years of the trajectory.

	<i>Selected 2003 (Rotational group 3)</i>	<i>Selected 2003 (Rotational group 4)</i>	<i>Selected 2004 (Rotational group 1)</i>	Total
<i>Total number selected 1)</i>	2393	2389	0	4782
Not included because				
- result-2003 ^= '11'	780	797	0	1577
- result-2004 ^= '11'	282	294	0	576
- result-2005 ^= '11'	196	208	0	404
Total number of households included in trajectory 2003- 2005	1135	1090	0	2225

Note 1. persons below 16 selected for the survey, but not yet included cf. section 2.1.2 is not included.

Trajectorie 2004-2005				
	<i>Selected 2003 (Rotational group 3, 4)</i>	<i>Selected 2003 (Rotational group 3, 4)</i>	<i>Selected 2004 (Rotational group 1)</i>	Total
<i>Total number selected 1)</i>	2393	2389	2360	7142
Not included beause				
- result-2003 ^= '11'	780	797	-	1577
- result-2004 ^= '11'	282	294	902	1478
- result-2005 ^= '11'	196	208	255	659
Total number of households included in trajectorie 2003- 2005	1135	1090	1203	3428

Note 1. persons below 16 selected for the survey, but not yet included, cf. section 2.1.2, is not included.

For details about attrition se Appendix 0.1 and 0.2.

#### 2.3.3.4. Distribution of substituted units (if applicable) by 'record of contact at address'

No substitution

#### 2.3.3.5.-2.3.3.6 Item non-response

Item non-response is generally very low between 0 and 2 pct. The most striking exception is HS130: LOWEST MONTHLY INCOME TO MAKE ENDS MEET, where it is 10.70.

Information about income is taken from a register. Against this background, Denmark has no item non-response for income variables.

For details see appendix 4.

## 2.4. Mode of data collection

Denmark is one of the countries, which uses a sample of persons rather than a sample of addresses or households in the survey.

The establishment of the sample and the delimitation of the household are undertaken in the way described below.

A sample of persons is selected from the Central Population Register (CPR).

All other persons living at the same address is identified using information in the register. In the same way, married couples, couples not married, but expected to be partners, the ID's of fathers and mothers living at the address etc. is identified. In the following, the results will be called the "register-household". The register household can be considered as a hypothesis to be checked in the survey.

As a general rule, the selected person becomes the respondent of the household questionnaire, and therefore the person to be interviewed about the composition of the household, etc. The only exception is the case, where the selected person is under 25 years and has parents living at the address. In this case, we randomly select one of the parents to represent the household (the household respondent).

In the 2005 survey, 366 out of 7.019 cases, which were closed by telephone, one of the parents of the selected person was selected as the household respondent.

After the interview, a "statistical household" following Eurostat's definition is defined. Persons in the register-household, who do not belong to the statistical household, will be excluded from the sample and persons belonging to the statistical household, who are not found in the register-household are included.

As mentioned income and demographic data, including citizenship etc. primarily comes from registers, while social data primarily comes from interviews.

The questionnaire was split up into 4 different parts.

- a) Questions relating to defining households
- b) Questions about the household
- c) General questions about the household members
- d) Detailed questions about the selected person; including detailed labour information and health information

According to the instructions given to the interviewers, questions under a), b) and c) and if the selected person is the same as the selected household respondent also d), shall be asked the person in the household selected as household respondent if possible. If this person is unable to respond, e.g. is not at home or is busy with other things, it should be attempted to arrange an appointment to conduct an interview at another time. If such an appointment appears be difficult to obtain, it shall be attempted to achieve an interview with the spouse, if any. The interviewers are told to accept partners not married as proxies for the interview, if necessary, but that they should be very careful in doing so. Other members

of the household should only be accepted as proxies in the worst case, e.g. if no other possibility is feasible. Table 7 shows the households by type of interview

Who was interviewed?	Number
The household respondent	6925
The spouse	73
A partner	8
Another person	13
Total number of households	7019

Questions under d shall preferably be asked the selected person. If it is not feasible, because the person is not home or is busy with other things the instruction is that a proxy interview with one of the parents is OK.

Who was interviewed?	Number	Percent
The selected person	311	85
A parent	54	15
Total	365	100

It is our experience that the procedure is the most feasible. It makes the interview more fluent and comfortable. Interviewing each household member individually instead of one household member on behalf of the others would be a troublesome process to the interviewers as well as to the interviewees.

It must be taken into account, that information on income and many other subjects is information extracted from registers, and therefore was not included in the questionnaire.

## 2.5. Interview duration

		Total	0-5 minutes	6-10 minutes	11-15 minutes	16-20 minutes	16-20 minutes
All households	Number	5957	1381	3174	981	245	176
	Per cent	100	23,2	53,3	16,5	4,1	3,0
1 person-households	Number	1288	422	637	158	42	29
	Per cent	100	32,8	49,5	12,3	3,3	0,5
2 person-households	Number	3658	825	2017	588	126	102
	Per cent	100	22,6	55,1	16,1	3,4	2,8
3 person-households	Number	774	113	410	175	46	30
	Per cent	100	14,6	53,0	22,6	5,9	3,9
4 or more	Number	237	21	110	60	31	15

person-households	Per cent	100	8,9	46,4	25,3	13,1	6,3
-------------------	----------	-----	-----	------	------	------	-----

Tabel 12 The average length of the household in minutes by number of householdmembers aged 16+				
All Households	1 person households	2persons households	3 person households	4+ persons households
8,9	8,2	8,8	9,8	11,5

### 3. COMPARABILITY

#### 3.1. Basic concepts and definitions

*Reference population:*

Private households residing in Denmark 1 January 2005 and members of these households.  
No difference from EU-SILC concept

*Private household definition:*

No difference from EU-SILC concept.

*Household membership:*

No difference from EU-SILC concept.

*Income reference period(s) used:*

Calendar year 2004

*Period for taxes on income and social insurance contributions:*

Calendar year 2004

*Reference period for taxes on wealth:*

Calendar year 2004

*Lag between the income reference period and current variables:*

4-6 months

*Total duration of the data collection of the sample:*

3 months

*Information on activity status during the income reference period:*

Calendar year 2004

#### 3.2. Components of income

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

Imputed rent for dwellings owned by the occupant is, in principle, calculated as 4% of the taxable value of the property in our national income statistics and in the micro-files that we transmit to Eurostat. The taxable value is a relatively good estimate of the market value. The properties are valued by the municipalities.

HY090G can be negative. Cf. Appendix 4: **HY090G**

Apart from these facts only insignificant departures from EUSILC 065/rev03 occur.

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

The variables concerning income, wealth and taxes are monitored by registers.

The most important source is the registers of the tax authorities. These registers contain information on all kinds of taxable income and on all kinds of taxes. In addition to information used for taxation purposes, the register contains specified information reported by municipalities on social assistance, housing allowances, disability benefits, sickness benefits etc. and on the originator's number in the Central Business Register.

Almost all income in Denmark is taxable. The only exceptions of any importance are child allowances, housing allowances and supplementary payments to the disabled and the like. The municipalities report, as mentioned above, information about these forms of income to the Tax Authorities.

Information about the number of days for which the taxpayer received benefits according to different social, unemployment and training schemes are submitted to Statistics Denmark by municipalities and other authorities. The information is located in the so-called Labour Market Policy Measures Register and is used, when the different kinds of benefits from unemployment funds, trade unions etc. are split up into the different income components.

Income in the form of regular pension is from private schemes and allowances from the State Education Fund's can be distinguished and broken down by components, using information about the kind of income in the tax authorities' registers and about the originator of the income from the Central Business Register and the age of the person.

Information about the amount of unemployment benefit payments can be extracted from a special register.

Information from these different sources makes it possible to estimate the breakdown of gross income by the components with a high degree of accuracy.

### *3.2.3. The form in which income variables at component level has been obtained.*

**Income components were collected gross.**

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



They were collected gross.  
Cf. 3.2.1

#### 4. COHERENCE

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

All income target variables are monitored using external sources.

## **Appendic 1: Information on weighting procedure for the cross sectional component of SILC-2005**

This paper describes the method which is used to determine cross-sectional weights for the SILC-sample 2005.

### **Sampling design**

SILC 2005 is a continuation of the SILC 2004 and SILC 2003 survey. SILC has a rotating panel design. Panels stay in principle 4 years in the sample. In total, SILC 2005 contains 9467 sampled persons. Out of these, 4706 have been in the sample since 2003. In 2004 a panel of 2379 entered the sample, and in 2005 another 2382 persons entered the sample. All persons were drawn randomly from the register of persons over 15 living in private households. The complete household to which a sampled person belongs, is observed (network sampling).

Compared to the previous year, the household around persons that remain in the sample may have changed. For instance, due to a divorce, the household around sampled person X may have been split up in two separate households. In such a case, only the part around the sampled person is observed in SILC 2005.

With the rotating panel design described above, persons that are entering the sampling frame from one year to the next have smaller inclusion probabilities than persons that always have been in the frame. This is especially true for persons that were 15 year old in the previous wave and for immigrating persons (both Danish and foreigners). The former group is largest (there are about 65.000 persons in each 1-year age class).

Since the persons in the panels that remain in the sample from one year to the next are turning 1 year older, the youngest persons in those panels are 17 years in the next wave. Persons of age 16 can therefore only enter the sample via the new panel. If they are drawn at random, they will be underrepresented by about a factor 4. The same is in principle true for immigrants.

To avoid underrepresentation of the 16 year olds, the frame is in practice extended to also include 13, 14, and 15 year olds. In the actual sample, there are therefore also some persons below 16 years, but they are kept out of the actual surveying until the wave in which they have reached the age of 16 years. Only from then, they participate in the sample. Notice that the the age group 13-15 is not drawn as a separate stratum. This means that the sample size of this group is stochastic, and the same with the remainder of the sample, the persons of 16 year and older. In the weighting procedure, this stochasticity is neglected, though. The group 16 year and older is in the weighting procedure treated as being of fixed size  $n$  sampled out of a population of size  $N$  of persons 16 and older.

The underrepresentation of other new persons in the frame is not dealt with in the design phase, mainly because this group is not expected to deviate much from the remainder of the population and thus is not expected to have a large influence on the final estimates.

## Description of weighting procedure

### *Step 1: Design weights.*

Let  $N$  be the number of persons over 15 in the population, and  $n$  be the sample size. Let furthermore  $M_h$  be the *present* number of persons over 15 in household  $h$ . The design weight for a sampled person  $i$  is given by

$$pb070_i = \frac{N}{n} .$$

Every sampled person has a weight  $pb070_i$  .

Since the complete household  $h$  to which a sampled person  $i$  belongs is observed, the inclusion probability of household  $h$  is proportional to the number of persons over 15 in that household. The design weight of household  $h$  is therefore given by

$$db080_h = \frac{N}{nM_h} .$$

The inclusion probability is the inverse of this quantity. Every household that is sampled via a sampled person has a weight  $db080_h$  .

Note that the household design weight applies both to households around sampled persons that are new in the survey, as well as to the households around persons that continue in the survey from the previous wave, even if its household composition has changed (i.e., the number of persons over 15 has changed since the previous wave). For instance, in case of a household splitting up in two parts, the inclusion probability of the original household has to be divided over the two new parts such that the sum of their inclusion probabilities equals the original inclusion probability. That is, the respective design weights are proportional to the present number of persons over 15. The same applies evidently when households merge, or combinations between merging and splitting.

### *Step 2: Initial correction for non-response.*

Let  $m$  be the number of responding households. The household design weights are initially corrected for this non-response by multiplying the design weights by a factor  $n/m$ , that is, after a first non-response correction we have the household weights

$$db080_h^{corrected} = \frac{N}{mM_h} .$$

### *Step 3: Further correction for non-response and calibration on registers.*

Because of selective non-response, the household weights  $db080_h^{corrected}$  give a rather skewed picture of the population of households. Therefore, these weights are corrected further for non-response. Simultaneously, these

weights are calibrated such that certain known population totals from registers are reproduced.

The sample data refers to households and in performing the non-response correction and register calibration, both household and person information of all persons in the households is included.

The non-response turns out to be correlated most with the total net household income, the size of the household, and the education level of the person with the highest professional status in that household. Non-response correction will be performed at a **household** level (that is, households are counted).

In addition to correcting for skewness due to non-response, it is important that the SILC cross-sectional weights reproduce certain demographic and poverty distributions from the register of persons. In particular, the weights should reproduce correct population totals for the number of **persons** by:

- Age (5 year age groups 0-15, 16-19, 20-24, ..., 70-74, 75+) and sex.
- Economic status and poverty.
- Professional status and poverty.
- Age (5 classes 0-15, 16-24, 25-49, 50-64, 65+), sex and poverty.
- Family type and poverty.
- Education.
- Equivalised income group.

A last requirement we want to include, is that the household weights should reproduce a correct distribution of age (in 5 classes) and sex, if the age and sex of the sampled person is assumed to be representative for the whole household (that is, if only the responding sampled persons are used, they should also reproduce the correct age-sex distribution for persons over 15).

All in all, the following weighting model is used:

$$\begin{aligned}
 & [\text{hhsz}] \times 1 + \\
 & [\text{famincgrp}] \times 1 + \\
 & [\text{educationgrp}] \times 1 + \\
 & [\text{sexhh} \times \text{agehh}] \times M + \\
 & [\text{eq\_incgrp}] \times A + \\
 & [1] \times \text{sexagecat1} + \dots + [1] \times \text{sexagecat28} + \\
 & [\text{poverty}] \times \text{ec\_status1} + \dots + [\text{poverty}] \times \text{ec\_status5} + \\
 & [\text{poverty}] \times \text{profession1} + \dots + [\text{poverty}] \times \text{profession8} + \\
 & [\text{poverty}] \times \text{sexagegrp1} + \dots + [\text{poverty}] \times \text{sexagegrp10} + \\
 & [\text{poverty} \times \text{famtype}] \times A + \\
 & [1] \times \text{education1} + \dots + [1] \times \text{education4}.
 \end{aligned}$$

The notation [ var1 x var2 ] x var3 stands for a count over numerical variable ‘var3’ in the population of households with respect to combination of categorical variables ‘var1 x var2’. For instance, [ var1 ] x 1 stands for a ‘frequency count’ of households in the population broken down to

categorical variable 'var1'. If the numerical variable is unequal to '1', persons are counted instead of households.

The meaning of the variables is as follows:

**Hhsize:** Number of persons (all ages) *in household* (4 categories: 1, 2, 3, 4 or more).

**Famincgrp:** Net family income *in household* (5 categories: 112 499 and less, 112 500 – 159 999, 160 000 – 234 999, 235 000 – 329 999, 330 000 and more; boundaries are chosen such that each group contains roughly 20% of the population of households).

**Educationgrp:** Highest education obtained (according to the register) of the person with the highest education *in the household* (3 categories: Isced 0-1-2, Isced 3-4, Isced 5-7).

**Sexhh:** Gender of sampled person *in household* (male/female).

**Agehh:** Age group of sampled person *in household* (5 classes: 0-15, 16-24, 25-49, 50-64, 65+; by definition 0-15 is empty).

**Eq\_incgrp:** Equivalised income group of *household* (3 classes: below 'at risk of poverty threshold', between 'at risk of poverty threshold' and median, above median), where the 'at risk of poverty threshold' is defined as 60% of the median of the equivalised income of all persons in the population.<sup>1</sup>

**Poverty:** Indication whether *the household* is below or above the 'at risk of poverty threshold'.

**Famtype:** Type of family living *in household* (10 classes: one person under 65 no children, one person 65 or older no children, two persons both under 65 no children, two persons at least one 65 or older no children, other household without children, one adult one or more children, two adults one child, two adults two children, two adults three or more children, other household with children).

**M:** Number of *persons* over 15 in household (equal to  $M_h$ ).

**A:** Number of *persons* (all ages) in household.

**Sexagecat1 to Sexagecat28:** Number of *persons* in household that belong to the following sex and age groups: **1:** male 0-15, **2:** male 16-19, **3:** male 20-24, ..., **13:** male 70-74, **14:** male 75+, **15:** female 0-15, **16:** female 16-19, **17:** female 20-24, ..., **27:** female 70-74, **28:** female 75+.

---

<sup>1</sup> The household equivalized income is calculated as the household total net income divided by equivalized household size according to the modified OECD scale (which gives a weight of 1.0 to the first adult, 0.5 to other persons aged 14 or over who are living in the household and 0.3 to each child aged less than 14). All persons in a household have the same equivalized income.

**Ec\_status1 to Ec\_status5:** Number of *persons* in household that are **1:** employed (excl self employed), **2:** self-employed, **3:** unemployed, **4:** retired, and **5:** other economically inactive.

**Profession1 to Profession8:** Number of *persons* in household that are in one out of 8 profession groups based on Disco.

**Sexagegrp1 to Sexagegrp10:** Number of *persons* in household that belong to the following sex and age groups: **1:** male 0-15, **2:** male 16-24, **3:** male 25-49, **4:** male 50-64, **5:** male 65+, **6:** female 0-15, **7:** female 16-24, **8:** female 25-49, **9:** female 50-64, **10:** female 65+.

**Education1 to Education4:** Number of *persons* in household with highest level of education equal to **1:** Unknown or Isced 0 and 1, **2:** Isced 2, **3:** Isced 3 and 4, **4:** Isced 5, 6, and 7.

The first three terms in the weighting model are inserted for non-response correction on the household level. These terms also ensure that the calibrated household weights will sum up to the correct number of households in Denmark.

The fourth term in the weighting model ensures that the household weights are chosen such that they return the correct number of persons over 15 in the population that are in a certain sex by age class, where sex and age class refer to the sampled person in the household.

The remaining terms in the weighting model ensure that the weight of each person in the household (which is identical to the household weight) is such that the population totals with respect to age and sex, education level, and ‘at risk of poverty’ rate broken down to several relevant variables **on a person level** are reproduced.

The weighting procedure with the linear model as specified above is performed in CLAN, using a regression estimator. The weights  $db080_i^{corrected}$  are used as initial weights. The final weights are bounded from below, such that they always are larger than  $0.02 \times db080_h^{corrected}$ . Since the minimum value of  $db080_h^{corrected}$  is 102, the final weights are always larger than one.

The bounding procedure used in CLAN works as follows. The final weights are initially calculated without lower bounds. For all records with final weight below the lower bound, the final weight is fixed to the lower bound. The final weights are then recalculated for the remaining records. If they all are above the lower bound, one is ready. Otherwise the weights below the lower bound are set to the lower bound and the weights of the remaining records are again recalculated. This procedure is iterated until all weights are above or equal to the lower bound.

In the present case, 20 households get a weight equal to the lower bound. In fact, it turned out that there is no solution at all for the bounding problem, since there are two responding household that - no matter what - get zero

final weight (this could be due to an internal inconsistency in the record or a seldom combination of the auxiliary information which does not occur in the register<sup>2</sup>). Therefore one is forced to take these particular respondents out of the response and subsequently use bounded weights on the remainder. The zero-weight respondents are afterwards added to the data with initial and final weight equal to zero.

The resulting, bounded, final weights are denoted with  $db090_h$ . These weights are household weights. Every household in the response has a weight  $db090_h$ .

**Step 4: Personal cross-section weights (all ages).**

The household weights  $db090_h$  also give the weight of each person  $i$  within the household  $h$ , irrespective of age. Thus,

$$rb050_i = db090_h.$$

Every person belonging to a responding household has a weight  $rb050_i$ . In total, there are 15320 persons living in the 5957 responding households.

**Step 5: Personal cross-section weights (over 15).**

The household weights  $db090_h$  also give the weight of each person within the household, irrespective of age. Thus,

$$pb040_i = db090_h,$$

where it is assumed that there is no non-response among persons over 15 in households that have responded. Every person over 15 belonging to a responding household has a weight  $pb040_i$ . In total, there are 11901 persons over 15 living in the 5957 responding households.

**Step 6: Personal cross-section weights (for sampled persons).**

By multiplying the household weights  $db090_h$  by  $M_h$ , the number of persons over 15 in the household, one obtains the calibrated weight for the sampled person in that household. Thus,

$$pb060_i = M_h \times db090_h.$$

Only the sampled persons belonging to responding households have a weight  $pb060_i$ .

This weight should be used to estimate population totals based on the sampled persons only. Since the sampled persons all are over 15, these weights sum up to the number of persons over 15 in the population (actually by sex and age class, see the fourth term in the weighting model).

---

<sup>2</sup> The household around a sampled person according to the register can deviate from the actually observed household since changes in the register are usually not immediately registered.

## External information

The following external information on population totals is used to calculate the weights.

Sample size:  $n = 9\,467$ .

Number of respondents (households):  $m = 5\,957$ .

Number of persons over 15:  $M = N = 4\,272\,821$

Number of persons all ages:  $A = 5\,342\,955$

Population of private households (with at least 1 person over 15):  $N_{hh} = 2\,643\,240$ .

Population totals of the auxiliary variables in the weighting model:

Auxiliary variable	class 1	class 2	class 3	class 4	class 5	class 6	class 7	class 8	class 9	class 10	class 11	class 12	class 13	class 14	class 15	class 16	class 17	class 18	class 19	class 20
		221517																		
[poverty] x ec_status1	102522	1																		
[poverty] x ec_status2	37834	158199																		
[poverty] x ec_status3	18407	113096																		
[poverty] x ec_status4	134697	525228																		
		170041																		
[poverty] x ec_status5	337385	6																		
[poverty] x sexagegrp1	54734	493391																		
[poverty] x sexagegrp2	69116	195316																		
[poverty] x sexagegrp3	96142	858346																		
[poverty] x sexagegrp4	26318	503788																		
[poverty] x sexagegrp5	57208	285282																		
[poverty] x sexagegrp6	52520	469489																		
[poverty] x sexagegrp7	79837	176866																		
[poverty] x sexagegrp8	85664	854258																		
[poverty] x sexagegrp9	22862	508198																		
[poverty] x sexagegrp10	86444	367176																		



[poverty] x	258127																			
profession1	493817	3																		
[poverty] x																				
profession2	24450	233361																		
[poverty] x																				
profession3	5669	151954																		
[poverty] x																				
profession4	63458	602912																		
[poverty] x																				
profession5	14413	285354																		
[poverty] x																				
profession6	15171	459314																		
[poverty] x																				
profession7	13111	331562																		
[poverty] x																				
profession8	756	66380																		
[famtype x poverty] x																				
A	216593	570939	80673	284664	48944	922138	65910	431104	3341	162943	70435	284861	24603	538542	40284	942064	37319	74669	42743	500186
	115286																			
[hhsizex] x 1	9	814567	281291	394513																
[famincgrp] x 1	396924	485786	568557	448496	743477															
		110948																		
[educationgrp] x 1	805234	1	728525																	
[sexhh x agehh] x M	0	264432	954488	530106	342490		0	256703	939922	531060	453620									

Population totals of the auxiliary variables in the weighting model:

Auxiliary variable	class 1	class 2	class 3	class 4	class 5	class 6	class 7	class 8	class 9	class 10	class 11	class 12	class 13	class 14	class 15	class 16	class 17	class 18	class 19	class 20
[1] x sexagecat1	548125																			
[1] x sexagecat2	120811																			
[1] x sexagecat3	143621																			
[1] x sexagecat4	169964																			
[1] x sexagecat5	190521																			
[1] x sexagecat6	207578																			
[1] x sexagecat7	201557																			
[1] x sexagecat8	184868																			
[1] x sexagecat9	178871																			
[1] x sexagecat10	193839																			
[1] x sexagecat11	157396																			
[1] x sexagecat12	115756																			
[1] x sexagecat13	87566																			
[1] x sexagecat14	139168																			
[1] x sexagecat15	522009																			
[1] x sexagecat16	114871																			
[1] x sexagecat17	141832																			
[1] x sexagecat18	170273																			
[1] x sexagecat19	189334																			
[1] x sexagecat20	202185																			
[1] x sexagecat21	196436																			
[1] x sexagecat22	181694																			
[1] x sexagecat23	178111																			
[1] x sexagecat24	192713																			
[1] x sexagecat25	160236																			
[1] x sexagecat26	124273																			
[1] x sexagecat27	102814																			
[1] x sexagecat28	226533																			
	124883																			
[1] x education1	0																			
	146542																			
[1] x education2	0																			
	171482																			
[1] x education3	7																			
[1] x education4	913878																			
		204064	267146																	
[eq_incgrp] x A	630845	6	4																	

## Appendix 2

*Marianne Houbiers, Statistics Denmark*  
October 6, 2006  
*Revised by Stine Fangel September 2007*

### **SILC weights for the longitudinal period 2003-2005 and the period 2004-2005**

This paper describes the method which is used to determine longitudinal weights for the period 2004-2005 and the period 2003-2005. Essentially the same method is used for longitudinal weights 2004-2005 as for longitudinal weight 2003-2004. This method is expanded to the longitudinal weight 2003-2005.

#### **Sampling design**

The SILC survey is conducted since 2003 and has a rotating panel design. Panels stay in principle four years in the sample. Three out of four panels is present in the year 2004 are also present in 2005. These panels constitute the longitudinal sample of interest for the period 2004-2005.

Sampling for SILC is done using network sampling. For each panel, a SRS sample of persons is drawn randomly from the register of persons over 15 living in private households. The complete household, to which a sampled person belongs, is then observed.

#### **Cross sectional weights**

Every year, a set of cross-sectional weights is determined for the households and persons in the 4 participating panels in that year. These weights are db080 and db090 (household weights), and pb070, pb060, rb050, and pb040 (personal weights). The weights for 2004 and 2005 are already transferred to EUROSTAT, including a description of their calculation.

#### **Base weights**

In a preliminary paper “*EU-SILC weighting procedures: an outline*” by Vijay Verma (2006), an integrated procedure is described to obtain both longitudinal and cross-sectional weights for SILC. The idea is to create some “base weights”, which are assigned to all persons in a panel from when the panel starts in the sample. These base weights are then adapted in subsequent waves, depending on the likelihood of certain people disappearing from the panel. Persons appearing in an existing household in the panel also get a weight.

Disappearance from the panel is modelled by logistic regression, resulting in a probability that a certain person is continuing to the next wave. The base weights are then multiplied by the inverse of this probability.

New entrances into the panel, such as newborn babies get the base weight of their mother, and new cohabitating persons in a household get a base weight depending on whether they come from completely outside the population or from a household that already existed in the population.

The above mentioned paper describe how exactly the base weights in the different waves should be calculated and how the cross-sectional and longitudinal weights are derived from them.

We have **directly** adopted the method described in the paper mentioned above in our derivation of the longitudinal weights. The reasons for this are the following:

1. We assume that the cross-sectional weights in any wave of a panel give a better approximation for the inclusion probabilities of households and persons in the sample at that particular time, than the above mentioned adapted base weights. The difference between the adapted base weights and the cross-sectional weights is assumed to become larger the longer a panel has been participating in SILC.

An example of a situation where there arises a difference is when a person enters a sample household from another non-sample household in the population. This person will initially get zero base weight in the wave where he enters. The weights of all household members are subsequently averaged and assigned to all persons in the household. In contrast, this person gets a cross-sectional household (and person) weight, which depends on the number of persons over 15 in that household at that particular wave. This means that his cross-sectional weight differs, depending on whether the person is below or above 15, whereas there is no difference for the adapted base weight.

2. In determining the adapted base weights for successive waves, a logistic regression is used for determining the probability that persons in the panel continue to the next wave. However, we can use register information to calibrate for nonresponse. This is assumed to give a better estimation of the equivalent of the base weight changes due to exit of persons from the sample.
3. With the base weight method, it is not clear how to include households that have not responded in the first wave, but have responded in the next waves. In the Danish SILC design, this is of importance, because we sample in reality among people over 13. However, only households around sampled persons over 15 are observed in the first wave. In the second wave, the households around sampled persons who are at that time over 15 are surveyed, etc. This is done to avoid that persons of 16, 17, or 18 years old are underrepresented among the sampled persons in a given panel. In addition, a household refusing to respond in one wave is still followed up in the next wave. Only after 2 successive refusals, the household is not followed up anymore.

So, instead of deriving adapted base weights for successive waves from the initial base weights in the first wave, we always start with the cross-sectional weights in the beginning of the longitudinal period of interest. In the following, the derivation of the longitudinal weights is explained in detail.

## **Description of the longitudinal weighting procedure two-year duration (RB062) 2004 – 2005**

The ultimate goal is to derive longitudinal weights RB062 (and RB063 for persons of all ages). Only these two personal weights are calculated as longitudinal weights, this is in contrast to last year where also to other type of longitudinal weights were calculated. These weights are all personal weights. There are no household longitudinal weights. Weights will only be derived for persons that are continuously present in the sample during the longitudinal period of interest.

### ***Step 1: Panel cross-sectional weights.***

We begin with the cross-sectional weights RB050 (persons all ages) year 2004, the first year of the longitudinal period. These weights are determined using calibration of the complete 2004-sample  $S_{2004}$  on the 2004-population  $U_{2004}$ . The weights RB050 are equivalent to the household weight DB090, and PB060 is a factor  $M_h$  larger ( $M_h$  is the number persons over 15 in household  $h$ ). In this calibration, the panel structure is not taken into account. This means, that the 4 panels are together, and not separately, calibrated on the population.

The first step in determining the longitudinal weights for the period 2004-2005 consists of separating  $S_{2004}$  into 4 panels  $S_{2004}^{(p)}$  ( $p=1,2,3,4$ ), such that each panel represents the whole population  $U_{2004}$ . This is

done by multiplying the household weights DB090 by the inverse of the relative size of the panel to which a certain household belongs:

$$DB090 \rightarrow \frac{n}{n_p} \times DB090,$$

where  $n = n_1 + n_2 + n_3 + n_4$  is the total number of (responding) households in the cross-sectional sample, and  $n_p$  is the number of (responding) households in panel  $p$  ( $p = 1, \dots, 4$ ).

Although each panel  $S_{2004}^{(p)}$  now counts for the whole population  $U_{2004}$ , they are not calibrated on the population totals on which the 4 panels together were calibrated. To compensate partially for this, each panel is subsequently calibrated on the correct number of households, the correct number of persons over 15, and the correct number of persons of all ages in  $U_{2004}$ . This gives a panelequivalent  $DB090^{(p)}$  of the cross sectional weights DB090. From  $DB090^{(p)}$ , the corresponding personal panelequivalents  $RB050^{(p)}$  are derived with  $DB090^{(p)} = RB050^{(p)}$ .

**Step 2: Longitudinal population.**

The second step consists of defining a so-called longitudinal population  $U_L$ . This population consists of all persons that are present in the population in the complete longitudinal period of interest. That means in the present case, all persons that were present in 2004 and 2005. The longitudinal population is thus given by

$$U_L \equiv U_{2004} \cap U_{2005}.$$

Persons that entered the population between 2004 and 2005 are not part of the longitudinal population, and neither are persons that exited the population in this period.

**Step 3: Longitudinal sample.**

The third step consists of determining the so-called longitudinal sample  $S_L$ . This sample consists of all persons that are present in the sample in both years of the longitudinal period. In the present case, that means all persons from panel 2, 3, and 4 that were observed both in 2004 and 2005:

$$S_L \equiv S_{2004} \cap S_{2005}.$$

Persons that entered a sampled household between 2004 and 2005 are thus not a part of the longitudinal sample. Neither are people who left a sampled household during this period, for instance due to a divorce or death. Also households that responded in 2004 but chose not to in 2005 are not part of the longitudinal sample.

Notice that the longitudinal sample consists of persons, i.e., the sampling units are not households.

**Step 4: Calibration of longitudinal sample on longitudinal population.**

Each of the 3 panels  $S_L^{(p)}$  in the longitudinal sample  $S_L$  should be representative for the longitudinal population  $U_L$ . But due to nonresponse, and people leaving a sampled household but not leaving the population, a longitudinal panel  $S_L^{(p)}$  is initially underrepresenting the longitudinal population. To compensate for this, the panel weights from Step 1 are calibrated for nonresponse. This is done independently for persons of all ages, persons over 15, and for the sampled persons.

The calibration is done for persons, that is, not for households. As a consequence, persons in the same household can get a different weight. This can be understood easily from an example: the probability that a 20-year old daughter is continuing in the next wave as a part of her parents household is smaller

than the probability that the 5 year old son is continuing. The daughter might leave her parents house in connection with an education elsewhere, or she might start to live together with someone else. Similarly, persons living in a household with an old sampled person have a smaller probability to continue in the next wave.

The calibration is done in 2 parts. First, the 3 longitudinal panels of each type of persons (all ages, over 15, sampled persons) are collected again into the longitudinal sample  $S_L$ , so that they together represent the (longitudinal) population  $U_L$ . This is done by multiplying the corresponding panel weights by the relative panel sizes, as in Step 1 but now the other way around. For persons of all ages we thus get

$$RB050^* = \frac{n_p}{n} \times DB050^{(p)},$$

where  $n_p$  is the number of responding households in the longitudinal panel  $S_L^{(p)}$  ( $p = 3, 4$  and  $1$ ) and  $n = n_2 + n_3 + n_4$  is the total number of (responding) households in the longitudinal sample  $S_L$  (these numbers are not the same as in Step 1!).

Second, the longitudinal sample  $S_L$  is calibrated with CLAN on the longitudinal population  $U_L$ . For persons of all ages, we use the weights  $RB050^*$  as starting weights, and the following weighting model:

$$\begin{aligned} & [sex \times age14] \times 1 + \\ & [sex \times age5 \times poverty] \times 1 + \\ & [eq\_incgrp] \times 1 + \\ & [famincgrp] \times 1 + \\ & [famtype] \times 1 + \\ & [education \times poverty] \times 1 + \\ & [hh\_educatingrp] \times 1 + \\ & [ec\_status \times poverty] \times 1 + \\ & [hh\_ec\_stat] \times 1. \end{aligned}$$

The notation  $[var1 \times var2] \times 1$  stands for a count over persons of all ages in  $U_L$  with respect to the combination of categorical variables ‘var1 x var 2’.

The auxiliary variables all refer to their value in 2004 and their meaning is as follows:

**Sex:** Gender of person (male, female).

**Age14:** Age group of person (14 classes: 0-15, 16-19, 20-24, ..., 60-64, 65-74, 75-84, 85+).

**Age5:** Age group of person (5 classes: 0-15, 16-24, 25-49, 50-64, 65+).

**Eq\_incgrp:** Equivalised income group of *household* to which person belongs (3 classes: below ‘at risk of poverty threshold’, between ‘at risk of poverty threshold’ and median, above median), where the ‘at risk of poverty threshold’ is defined as 60% of the median of the equivalised income of all persons in

the population.<sup>3</sup> The variable is determined for the whole household in 2004, whether or not all persons belong to the longitudinal household or not.

**Poverty:** Indication whether *the household* is below or above the ‘at risk of poverty threshold’.

**Famincgrp:** Net family income *in household* to which person belongs (5 categories: 112 499 and less, 112 500 – 159 999, 160 000 – 234 999, 235 000 – 329 999, 330 000 and more; boundaries are chosen such that each group contains roughly 20% of the population of households). The family income is determined for the whole household in 2004, whether or not all persons belong to the longitudinal household or not.

**Famtype:** Type of family living *in household* to which person belongs (10 classes: one person under 65 no children, one person 65 or older no children, two persons both under 65 no children, two persons at least one 65 or older no children, other household without children, one adult one or more children, two adults one child, two adults two children, two adults three or more children, other household with children). The family type is determined for the whole household in 2004, whether or not all persons belong to the longitudinal household or not.

**Educationgrp:** Highest level of education of person (4 classes, namely: **1:** Unknown or Isced 0 and 1, **2:** Isced 2, **3:** Isced 3 and 4, **4:** Isced 5, 6, and 7).

**Hh\_educationgrp:** Highest education obtained (according to the register) of the person with the highest education *in the household* (3 categories: Isced 0-1-2, Isced 3-4, Isced 5-7). The hh\_educationgrp is determined for the whole household in 2004, whether or not all persons belong to the longitudinal household or not.

**Ec\_status:** Economic status of person (5 classes, namely **1:** employed (excl self employed), **2:** self-employed, **3:** unemployed, **4:** retired, and **5:** other economically inactive).

**Hh\_ec\_status:** Economic status of person with highest education in the household (5 classes, namely **1:** employed (excl self employed), **2:** self-employed, **3:** unemployed, **4:** retired, and **5:** other economically inactive).

Since the original cross-sectional weights already were calibrated on similar population totals (though for the whole population and not only the longitudinal part), the above calibration does not require that much. There occur for instance no negative weights. The most important effect of the calibration is to ensure that persons who have a lower probability to continue from one wave to the next get a higher weight. The resulting weights are the longitudinal weights RB062.

### Remarks for longitudinal weights of two-year duration

Some remarks on the longitudinal weights are in order.

As mentioned before, persons in the same household generally do not have the same longitudinal weight RB062. This is in contrast to the cross-sectional weights, where all persons within a household are assigned the same weight RB050.

One of the consequences is, that the longitudinal weights are less ‘integrated’. That means that it is in principle possible to get a domain estimate for a certain variable based on the weights PB050 which is

---

<sup>3</sup> The household equivalized income is calculated as the household total net income divided by equivalized household size according to the modified OECD scale (which gives a weight of 1.0 to the first adult, 0.5 to other persons aged 14 or over who are living in the household and 0.3 to each child aged less than 14). All persons in a household have the same equivalized income.

larger than the same estimate based on RB062, i.e., the estimated number of persons over 15 with a certain property is larger than the estimated number of persons of all ages with that property.

Such inconsistencies can in principle be avoided for the weights RB060 and PB050 by performing a completely separate weighting for persons below and above 15 years, but here we have chosen not to do that.

### **Description of the longitudinal weighting procedure for three-year duration (RB063) 2003 – 2005**

In order to calculate the longitudinal weights for a period of three years the same method as above is used. All four steps are followed the same way as for RB062 but only two panels are used now, panel 3 and 4.

#### ***Step 1: Panel cross-sectional weights.***

This is done the exact same way as for the RB062 weights. We begin with the cross-sectional weight RB050 from the year 2003. There is adjusted for panels and we end up with panel DB090 weights DB090<sup>(p)</sup>.

#### ***Step 2: First longitudinal population.***

The second step is defining the first longitudinal population  $U_L$ . This population consist of all persons that are in the population from 2003 and the population 2004.

$$U_L \equiv U_{2003} \cap U_{2004}$$

Persons that entered in the population between 2003 and 2004 are part of the longitudinal populations as well as persons that exited the population in this period of time.

#### ***Step 3: First longitudinal sample.***

The third step is determining the so-called first longitudinal sample  $S_L$ . This sample consists of all persons that are represented in both years of the longitudinal period and are in panel 3 or 4. These two panels are observed in 2003, 2004 and 2005.

$$S_L \equiv S_{2003, \text{panel3 or panel4}} \cap S_{2004, \text{panel3 or panel4}}$$

#### ***Step 4: Calibration of the first longitudinal sample on the first longitudinal population.***

The two panels  $S_L^{(p)}$  in the longitudinal sample  $S_L$  should be representative for the longitudinal population  $U_L$ .

This is the exact same procedure as in step 4 for the two-year duration longitudinal weights, and the same model and every thing is used. The auxiliary information refers to year 2003.

These weights are not the final RB063 as they are the final RB062 in the two-year duration procedure. These weights are used as stating weights in step 7, but first we need to make the three year longitudinal population and the three year sample. This sample is not the same as in step 3 due to non response.

#### ***Step 5: longitudinal population.***

The fifth step is defining the longitudinal population  $U_L$ . This population consist of all persons that are in the population 2003, in the population 2004 and in the population 2005.



$$U_L \equiv U_{2003} \cap U_{2004} \cap U_{2005}$$

Persons that entered in the population between 2003 and 2005 are part of the longitudinal populations as well as persons that exited the population in this period of time. Furthermore, persons that are in the population 2003 and in the population 2005 but not in the population 2004 are not in the longitudinal population.

***Step 6: longitudinal sample***

The sixth step is to find the longitudinal sample for three years. This sample consists of all persons that are represented in all three years of the longitudinal period and are in panel 3 or 4. These two panels are observed in 2003, 2004 and 2005.

$$S_L \equiv S_{2003} \cap S_{2004} \cap S_{2005}$$

Persons who have responded in 2003, 2004 and 2005 are in the longitudinal sample no one else is.

***Step 7: Calibration of longitudinal sample on the longitudinal population.***

The two panels  $S_L^{(p)}$  in the longitudinal sample  $S_L$  should be representative for the longitudinal population  $U_L$ .

The starting weights are the ones found in step 4. A calibration is done with auxiliary information from 2004 but the same model as from step 4 (with newer auxiliary information as this is from 2004 compared to step 4 with 2003 auxiliary information).

These weights are the final RB063 longitudinal weights.

**The method for longitudinal period of 3 years in short**

For a 3 year longitudinal period, one starts with the cross-sectional weights  $w_X^t$  at year t, and calculate longitudinal weights  $w_L^{t,t+1}$  for the period t to t+1 for two panels

Subsequently, longitudinal weights  $w_L^{t+1,t+2}$  are calculated for the period t+1 to t+2, where the longitudinal weights  $w_L^{t,t+1}$  are used as starting weights instead of the cross-sectional panel weights  $w_X^t$  that were used in the period t to t+1. The remaining longitudinal sample  $S_L \equiv S_t \cap S_{t+1} \cap S_{t+2}$  is calibrated on the remaining longitudinal population  $U_L \equiv U_t \cap U_{t+1} \cap U_{t+2}$ , using household and personal information from  $U_{t+1}$  at t+1 as auxiliary information. The weights  $w_L^{t+1,t+2}$  are the weights of interest. They reflect changes in the longitudinal population and sample over the whole longitudinal period from t to t+2.

**Small remark on Base Weights RB060, PB050 and PB080**

From the above mentioned procedure one can see that we do not use the base weights to calculate the RB062 longitudinal weight or RB063 longitudinal weight. Base weights are assumed to be the same as the cross sectional weights in the set-up we have made for the Danish SILC survey.

In the DOC LC/08/07EN\_Supporting document EU-SILC 065/05.1 p. 35

RB060 = RB050 for wave t=1,

and

PB080 = PB060 for wave t=1,

from this we also assume that

PB040 = PB050 for wave t=1.

These weights are cross-sectional weights, which is calculated each year with the most recent data. Furthermore, these weights adjust for nonresponse in a way so the weights recreate different population totals that are critical for nonresponse bias.

No longitudinal weights are calculated for the P-file like PB052 and PB053 or PB082 and PB083 as they are not requested. Therefore, only longitudinal weights for the R-file are calculated weights RB062 and RB063.

### Appendix 3

Indicator	Value	Achieved sample size	Standard error	Standard error (after calibrating)	Confidence Interval at 95%		CV(%)	DEFF
					Lower	Upper		
At-risk-of-poverty rate after social transfers - total	11,8	15321	0,39	0,08	11,6	12,0	0,7	1,00
At-risk-of-poverty rate after social transfers - men total	11,6	7707	0,43	0,07	11,5	11,7	0,6	1,00
At-risk-of-poverty rate after social transfers - women total	12,1	7614	0,48	0,10	11,9	12,3	0,8	1,00
At-risk-of-poverty rate after social transfers - 0-15 years	10,1	3420	0,74	0,21	9,7	10,5	2,1	1,00
At-risk-of-poverty rate after social transfers - 16-24 years	28,9	1631	1,25	0,21	28,5	29,3	0,7	1,00
At-risk-of-poverty rate after social transfers - 25-49 years	9,7	5423	0,43	0,07	9,6	9,8	0,7	1,00
At-risk-of-poverty rate after social transfers - 50-64 years	4,6	3172	0,45	0,01	4,6	4,6	0,2	1,00
At-risk-of-poverty rate after social transfers - 65+ years	17,6	1675	1,26	0,23	17,1	18,1	1,3	1,00
At-risk-of-poverty rate after social transfers - 16+ years	12,3	11901	0,40	0,06	12,2	12,4	0,5	1,00
At-risk-of-poverty rate after social transfers - 16-64 years	11,0	10226	0,35	0,05	10,9	11,1	0,5	1,00
At-risk-of-poverty rate after social transfers - 0-64 years	10,8	13646	0,38	0,08	10,6	11,0	0,7	1,00
At-risk-of-poverty rate after social transfers - men 16-24 years	26,2	853	1,48	0,15	25,9	26,5	0,6	1,00
At-risk-of-poverty rate after social transfers - men 25-49 years	10,2	2611	0,51	0,09	10,0	10,4	0,8	1,00
At-risk-of-poverty rate after social transfers - men 50-64 years	5,0	1627	0,60	0,01	5,0	5,0	0,3	1,00
At-risk-of-poverty rate after social transfers - men 65+ years	16,5	833	1,46	0,17	16,2	16,8	1,0	1,00
At-risk-of-poverty rate after social transfers - men 16+ years	11,9	5924	0,44	0,06	11,8	12,0	0,5	1,00
At-risk-of-poverty rate after social transfers - men 16-64 years	11,0	5091	0,41	0,06	10,9	11,1	0,5	1,00
At-risk-of-poverty rate after social transfers - men 0-64 years	10,8	6874	0,42	0,08	10,7	10,9	0,7	1,00
At-risk-of-poverty rate after social transfers - women 16-24 years	31,7	778	1,79	0,32	31,1	32,3	1,0	1,00
At-risk-of-poverty rate after social transfers - women 25-49 years	9,2	2812	0,55	0,09	9,0	9,4	0,9	1,00
At-risk-of-poverty rate after social transfers - women 50-64 years	4,3	1545	0,53	0,01	4,3	4,3	0,2	1,00
At-risk-of-poverty rate after social transfers - women 65+ years	18,4	842	1,45	0,33	17,7	19,1	1,8	1,00
At-risk-of-poverty rate after social transfers - women 16+ years	12,6	5977	0,49	0,09	12,4	12,8	0,7	1,00
At-risk-of-poverty rate after social transfers - women 16-64 years	11,0	5135	0,46	0,07	10,9	11,1	0,6	1,00
At-risk-of-poverty rate after	10,8	6772	0,47	0,09	10,6	11,0	0,8	1,00

social transfers - women 0-64 years								
At-risk-of-poverty rate after social transfers - employed	<b>4,9</b>	7437	0,29	0,17	4,6	5,2	3,4	1,00
At-risk-of-poverty rate after social transfers - unemployed	<b>25,9</b>	365	1,95	1,63	22,7	29,1	6,3	1,00
At-risk-of-poverty rate after social transfers - retired	<b>15,8</b>	1920	1,13	0,32	15,2	16,4	2,1	1,00
At-risk-of-poverty rate after social transfers - other inactive	<b>28,3</b>	1850	1,11	0,65	27,0	29,6	2,3	1,00
At-risk-of-poverty rate after social transfers - men, employed	<b>5,0</b>	3910	0,33	0,21	4,6	5,4	4,2	1,00
At-risk-of-poverty rate after social transfers - men, unemployed	<b>38,8</b>	152	3,03	2,65	33,6	44,0	6,8	1,00
At-risk-of-poverty rate after social transfers - men, retired	<b>15,3</b>	938	1,34	0,39	14,5	16,1	2,5	1,00
At-risk-of-poverty rate after social transfers - men, other inactive	<b>29,3</b>	764	1,52	0,99	27,4	31,2	3,4	1,00
At-risk-of-poverty rate after social transfers - women, employed	<b>4,8</b>	3527	0,38	0,28	4,3	5,3	5,8	1,00
At-risk-of-poverty rate after social transfers - women, unemployed	<b>14,1</b>	213	2,06	1,79	10,6	17,6	12,7	1,00
At-risk-of-poverty rate after social transfers - women, retired	<b>16,1</b>	982	1,28	0,45	15,2	17,0	2,8	1,00
At-risk-of-poverty rate after social transfers - women, other inactive	<b>27,7</b>	1086	1,38	0,85	26,0	29,4	3,1	1,00
At-risk-of-poverty rate after social transfers - single, < 65 years	<b>27,6</b>	765	1,06	0,16	27,3	27,9	0,6	1,00
At-risk-of-poverty rate after social transfers - single, 65+ years	<b>21,5</b>	387	1,75	0,39	20,7	22,3	1,8	1,00
At-risk-of-poverty rate after social transfers - single, male	<b>26,0</b>	531	1,25	0,34	25,3	26,7	1,3	1,00
At-risk-of-poverty rate after social transfers - single, female	<b>25,4</b>	621	1,32	0,38	24,7	26,1	1,5	1,00
At-risk-of-poverty rate after social transfers - single, total	<b>25,7</b>	1152	0,96	0,17	25,4	26,0	0,7	1,00
At-risk-of-poverty rate after social transfers - 2 adults, no children, both < 65	<b>5,0</b>	2942	0,56	0,12	4,8	5,2	2,4	1,00
At-risk-of-poverty rate after social transfers - 2 adults, no children, at least one 65+	<b>13,0</b>	1460	1,36	0,21	12,6	13,4	1,6	1,00
At-risk-of-poverty rate after social transfers - other households without children	<b>0,9</b>	590	0,53	0,56	-0,2	2,0	61,7	1,00
At-risk-of-poverty rate after social transfers - single parent, at least one child	<b>20,9</b>	597	2,44	0,89	19,1	22,7	4,3	1,00
At-risk-of-poverty rate after social transfers - 2 adults, 1 child	<b>4,3</b>	1992	0,72	0,05	4,2	4,4	1,2	1,00
At-risk-of-poverty rate after social transfers - 2 adults, 2 children	<b>4,7</b>	3860	0,68	0,22	4,3	5,1	4,7	1,00
At-risk-of-poverty rate after social transfers - 2 adults, 3+ children	<b>13,8</b>	1833	1,57	0,49	12,8	14,8	3,5	1,00
At-risk-of-poverty rate after social transfers - other households with children	<b>4,5</b>	769	1,79	1,70	1,2	7,8	37,9	1,00
At-risk-of-poverty rate after social transfers - households without children	<b>14,9</b>	6144	0,55	0,10	14,7	15,1	0,7	1,00
At-risk-of-poverty rate after social transfers -	<b>8,6</b>	9177	0,55	0,13	8,3	8,9	1,6	1,00

households with children								
At-risk-of-poverty rate after social transfers - owner or rent-free	7,5	11794	0,38	0,22	7,1	7,9	3,0	1,00
At-risk-of-poverty rate after social transfers - tenant	20,6	3527	0,89	0,51	19,6	21,6	2,5	1,00
At-risk-of-poverty rate after social transfers - households without children, $w = 0^1$	33,7	977	1,67	1,04	31,7	35,7	3,1	1,00
At-risk-of-poverty rate after social transfers - households without children, $0 < w < 1$	11,0	365	2,29	2,16	6,8	15,2	19,7	1,00
At-risk-of-poverty rate after social transfers - households without children, $w = 1$	5,8	3259	0,65	0,57	4,7	6,9	9,8	1,00
At-risk-of-poverty rate after social transfers - households with children, $w = 0$	5,1	9221	0,39	0,26	4,6	5,6	5,2	1,00
At-risk-of-poverty rate after social transfers - households with children, $0 < w < 0.5$		0						1,00
At-risk-of-poverty rate after social transfers - households with children, $0.5 < w < 1$		0						1,00
At-risk-of-poverty rate after social transfers - households with children, $w = 1$		0						1,00
Median of the equivalised disposable household income	22125,0	15321	148,65	16,67	22092,3	22157,7	0,1	1,00
At-risk-of-poverty threshold - single	13275,0	15321	89,19	10,00	13255,4	13294,6	0,1	1,00
At-risk-of-poverty threshold - 2 adults, 2 children	27877,5	15321	187,29	21,00	27836,3	27918,7	0,1	1,00
Inequality of income distribution S80/S20 income quintile share ratio	3,5	15321	0,10	0,09	3,3	3,7	2,5	1,00
Relative median at-risk-of-poverty gap - total	15,6	910	0,76	0,65	14,3	16,9	4,2	1,00
Relative median at-risk-of-poverty gap - men total	15,5	410	0,96	0,82	13,9	17,1	5,3	1,00
Relative median at-risk-of-poverty gap - women total	15,9	500	0,86	0,76	14,4	17,4	4,8	1,00
Relative median at-risk-of-poverty gap - 0-15 years	18,2	176	1,70	1,58	15,1	21,3	8,7	1,00
Relative median at-risk-of-poverty gap - 16-64 years	21,6	504	1,13	0,99	19,7	23,5	4,6	1,00
Relative median at-risk-of-poverty gap - 65+ years	8,1	230	0,63	0,57	7,0	9,2	7,0	1,00
Relative median at-risk-of-poverty gap - 16+ years	15,6	734	0,73	0,60	14,4	16,8	3,9	1,00
Relative median at-risk-of-poverty gap - men, 16-64 years	22,1	229	1,49	1,35	19,4	24,8	6,1	1,00
Relative median at-risk-of-poverty gap - men, 65+ years	7,0	98	0,72	0,67	5,7	8,3	9,6	1,00
Relative median at-risk-of-poverty gap - men, 16+ years	14,3	327	0,89	0,76	12,8	15,8	5,3	1,00
Relative median at-risk-of-poverty gap - women, 16-64 years	21,4	275	1,36	1,23	19,0	23,8	5,7	1,00
Relative median at-risk-of-poverty gap - women, 65+ years	8,7	132	0,78	0,73	7,3	10,1	8,4	1,00
Relative median at-risk-of-poverty gap - women, 16+ years	16,0	407	0,86	0,74	14,5	17,5	4,6	1,00
Dispersion around the risk-of-poverty threshold - 40%	3,4	15321	0,19	0,14	3,1	3,7	4,1	1,00

Dispersion around the risk-of-poverty threshold - 50%	5,7	15321	0,27	0,16	5,4	6,0	2,8	1,00
Dispersion around the risk-of-poverty threshold - 70%	19,5	15321	0,50	0,29	18,9	20,1	1,5	1,00
Before social transfers except old-age and survivors' benefits								
At-risk-of-poverty rate before social transfers - total	29,9	15321	0,50	0,35	29,2	30,6	1,2	1,00
At-risk-of-poverty rate before social transfers - men total	28,4	7707	0,56	0,40	27,6	29,2	1,4	1,00
At-risk-of-poverty rate before social transfers - women total	31,4	7614	0,62	0,45	30,5	32,3	1,4	1,00
At-risk-of-poverty rate before social transfers - 0-15 years	25,1	3420	1,11	0,89	23,4	26,8	3,5	1,00
At-risk-of-poverty rate before social transfers - 16-64 years	28,6	10226	0,51	0,34	27,9	29,3	1,2	1,00
At-risk-of-poverty rate before social transfers - 65+ years	42,2	1675	1,85	0,94	40,4	44,0	2,2	1,00
At-risk-of-poverty rate before social transfers - 16+ years	31,1	11901	0,49	0,31	30,5	31,7	1,0	1,00
At-risk-of-poverty rate before social transfers - men, 16-64 years	26,8	5091	0,59	0,41	26,0	27,6	1,5	1,00
At-risk-of-poverty rate before social transfers - men, 65+ years	44,1	833	2,00	1,27	41,6	46,6	2,9	1,00
At-risk-of-poverty rate before social transfers - men, 16+ years	29,7	5924	0,56	0,38	29,0	30,4	1,3	1,00
At-risk-of-poverty rate before social transfers - women, 16-64 years	30,3	5135	0,64	0,45	29,4	31,2	1,5	1,00
At-risk-of-poverty rate before social transfers - women, 65+ years	40,7	842	2,14	1,13	38,5	42,9	2,8	1,00
At-risk-of-poverty rate before social transfers - women, 16+ years	32,5	5977	0,62	0,41	31,7	33,3	1,3	1,00
Before social including old-age and survivors' benefits								
At-risk-of-poverty rate before social transfers - total	38,0	15321	0,56	0,34	37,3	38,7	0,9	1,00
At-risk-of-poverty rate before social transfers - men total	34,8	7707	0,62	0,39	34,0	35,6	1,1	1,00
At-risk-of-poverty rate before social transfers - women total	41,0	7614	0,66	0,43	40,2	41,8	1,0	1,00
At-risk-of-poverty rate before social transfers - 0-15 years	25,2	3420	1,11	0,89	23,4	27,0	3,5	1,00
At-risk-of-poverty rate before social transfers - 16-64 years	29,1	10226	0,51	0,34	28,4	29,8	1,2	1,00
At-risk-of-poverty rate before social transfers - 65+ years	93,7	1675	0,69	0,63	92,5	94,9	0,7	1,00
At-risk-of-poverty rate before social transfers - 16+ years	41,1	11901	0,53	0,29	40,5	41,7	0,7	1,00
At-risk-of-poverty rate before social transfers - men, 16-64 years	27,3	5091	0,59	0,42	26,5	28,1	1,5	1,00
At-risk-of-poverty rate before social transfers - men, 65+ years	90,9	833	1,08	0,98	89,0	92,8	1,1	1,00
At-risk-of-poverty rate before social transfers - men, 16+ years	37,7	5924	0,61	0,36	37,0	38,4	1,0	1,00
At-risk-of-poverty rate before social transfers - women, 16-64 years	31,0	5135	0,65	0,45	30,1	31,9	1,5	1,00

At-risk-of-poverty rate before social transfers - women, 65+ years	<b>95,8</b>	842	0,70	0,66	94,5	97,1	0,7	1,00
At-risk-of-poverty rate before social transfers - women, 16+ years	<b>44,4</b>	5977	0,64	0,37	43,7	45,1	0,8	1,00
Gini coefficient	<b>23,9</b>	15321	0,46	0,40	23,1	24,7	1,7	1,00
Mean equivalised disposable income	<b>23290,4</b>	15321	168,45	118,88	23057,3	23523,4	0,5	1,00





## Appendix 4: Over view of target variables.

### Abbreviations

na: Not applicable

r: Register

i: Interview

c: Calculated

b: Generated by the blaisesystem

h: Householdregister

p: Personal register

r/i: Primary source is Register. Secondary source is interview

i/r: Primary source is interview. Secondary source is register.

ol; Only longitudinal

	Source	Item non response	Comments
<b>HOUSEHOLD REGISTER (D-FILE)</b>			
DB010: YEAR OF THE SURVEY	2005	na	
DB020: COUNTRY	DK	na	
DB030: HOUSEHOLD ID	-	na	
DB040: REGION	r	0	
DB050: PRIMARY STRATA	na	na	
DB060: PSU-1 (FIRST STAGE)	na	na	
DB062: PSU-2 (SECOND STAGE)	na	na	
DB070: ORDER OF SELECTION OF PSU	na	na	
DB075: ROTATIONAL GROUP	s	na	
DB080: HOUSEHOLD DESIGN WEIGHT	c	na	
DB090: HOUSEHOLD CROSS-SECTIONAL WEIGHT	c	na	
DB100: DEGREE OF URBANISATION	r	0	
DB110: HOUSEHOLD STATUS	ol	na	
DB120: CONTACT AT ADDRESS	h/i	na	
DB130: HOUSEHOLD QUESTIONNAIRE RESULT	h/i	na	
DB135: HOUSEHOLD INTERVIEW ACCEPTANCE	h/i	na	

<b>PERSONAL REGISTER (R-FILE)</b>			
RB010: YEAR OF THE SURVEY	2005	na	
RB020: COUNTRY	DK	na	
RB030: PERSONAL ID	p	na	
RB040: CURRENT HOUSEHOLD ID	p	na	
RB041: PERSONAL ID	p	na	
RB050: PERSONAL CROSS-SECTIONAL WEIGHT	c	na	
RB060: PERSONAL BASE WEIGHT	c	na	
RB070: MONTH OF BIRTH	r	0	
RB080: YEAR OF BIRTH	r	0	
RB090: SEX	r	0	
RB100: SAMPLE PERSON OR CO-RESIDENT	i/p		
RB110: MEMBERSHIP STATUS	s/i	ol	

<b>RB120: MOVED TO</b>	<b>i/r</b>	<b>ol</b>	
<b>RB140: MONTH MOVED OUT OR DIED</b>	<b>i/r</b>	<b>ol</b>	
<b>RB150: YEAR MOVED OUT OR DIED</b>	<b>i/r</b>	<b>ol</b>	
<b>RB160: NUMBER OF MONTHS IN HOUSEHOLD DURING THE INCOME REFERENCE PERIOD</b>	<b>i/r</b>	<b>ol</b>	
<b>RB170: MAIN ACTIVITY STATUS DURING THE INCOME REFERENCE PERIOD</b>	<b>i/r</b>	<b>ol</b>	
<b>RB180: MONTH MOVED IN</b>	<b>i/r</b>	<b>ol</b>	
<b>RB190: YEAR MOVED IN</b>	<b>i/r</b>	<b>ol</b>	
<b>RB200: RESIDENTIAL STATUS</b>	<b>p/i</b>	<b>0</b>	
<b>RB210: BASIC ACTIVITY STATUS</b>	<b>i/r</b>	<b>0,85</b>	
<b>RB220: FATHER ID</b>	<b>r/i</b>	<b>0</b>	
<b>RB230: MOTHER ID</b>	<b>r/i</b>	<b>0</b>	
<b>RB240: SPOUSE/PARTNER ID</b>	<b>r/i</b>	<b>0</b>	
<b>RB245: RESPONDENT STATUS</b>	<b>p/i</b>	<b>0</b>	
<b>RB250: DATA STATUS</b>	<b>p/i</b>	<b>0</b>	
<b>RB260: TYPE OF INTERVIEW</b>	<b>p/i</b>	<b>na</b>	<b>cf. section 2.4 of the quality report for further information</b>
<b>RB270: PERSONAL ID OF PROXY</b>	<b>p/i</b>	<b>na</b>	
<b>RL010: EDUCATION AT PRE-SCHOOL</b>	<b>i</b>	<b>2,01</b>	
<b>RL020: EDUCATION AT COMPULSORY SCHOOL</b>	<b>i</b>	<b>0,43</b>	
<b>RL030: CHILD CARE AT CENTRE-BASED SERVICES</b>	<b>i</b>	<b>0,77</b>	
<b>RL040: CHILD CARE AT DAY-CARE CENTRE</b>	<b>i</b>	<b>0</b>	
<b>RL050: CHILD CARE BY A PROFESSIONAL CHILD-MINDER AT CHILD'S HOME OR AT CHILDMINDER'S HOME</b>	<b>i</b>	<b>0</b>	
<b>RL060: CHILD CARE BY GRAND-PARENTS, OTHERS HOUSEHOLD MEMBERS (OUTSIDE PARENTS),</b>	<b>i</b>	<b>0</b>	
<b>OTHER RELATIVES, FRIENDS OR NEIGHBOURS</b>	<b>i</b>	<b>0</b>	
<b>RL070: CHILDREN CROSS-SECTIONAL WEIGHT FOR CHILD CARE</b>	<b>c</b>	<b>na</b>	

<b>HOUSEHOLD DATA (H-FILE)</b>			
<b>HB010: YEAR OF THE SURVEY</b>	<b>2005</b>	<b>na</b>	
<b>HB020: COUNTRY</b>	<b>DK</b>	<b>na</b>	
<b>HB030: HOUSEHOLD ID</b>	<b>h</b>	<b>na</b>	
<b>HB040: DAY OF HOUSEHOLD INTERVIEW</b>	<b>i/b</b>	<b>4,77</b>	
<b>HB050: MONTH OF HOUSEHOLD INTERVIEW</b>	<b>i/b</b>	<b>4,77</b>	
<b>HB060: YEAR OF HOUSEHOLD INTERVIEW</b>	<b>2005</b>	<b>na</b>	
<b>HB070: PERSON RESPONDING THE HOUSEHOLD QUESTIONNAIRE</b>	<b>i/h</b>	<b>0</b>	
<b>HB080: PERSON 1 RESPONSIBLE FOR THE ACCOMMODATION</b>	<b>i/r</b>	<b>0</b>	
<b>HB090: PERSON 2 RESPONSIBLE FOR THE ACCOMMODATION</b>	<b>i/r</b>	<b>0</b>	
<b>HB100: NUMBER OF MINUTES TO COMPLETE THE HOUSEHOLD QUESTIONNAIRE</b>	<b>s</b>	<b>i/b</b>	
<b>HH010: DWELLING TYPE</b>	<b>r</b>	<b>4,78</b>	
<b>HH020: TENURE STATUS</b>	<b>i/r</b>	<b>0</b>	
<b>HH030: NUMBER OF ROOMS AVAILABLE</b>	<b>r/i</b>	<b>0,65</b>	

<b>TO THE HOUSEHOLD</b>			
<b>HH031: YEAR OF CONTRACT OR PURCHASING OR INSTALLATION</b>	na	na	
<b>HH040: LEAKING ROOF, DAMP WALLS/FLOORS/FOUNDATION, OR ROT IN WINDOW FRAMES OR FLOOR</b>	i	0,34	
<b>HH050: ABILITY TO KEEP HOME ADEQUATELY WARM</b>	i	0,34	
<b>HH060: CURRENT RENT RELATED TO OCCUPIED DWELLING</b>	i	1,65	
<b>HH061: SUBJECTIVE RENT</b>	na	na	
<b>HH070: TOTAL HOUSING COST</b>	c	0,08	Imputed. Using data from registers and the household budget survey
<b>HH080: BATH OR SHOWER IN DWELLING</b>	r	na	
<b>HH090: INDOOR FLUSHING TOILET FOR SOLE USE OF HOUSEHOLD</b>	r	na	
<b>HS010: ARREARS ON MORTGAGE OR RENT PAYMENTS</b>	i	0	
<b>HS020: ARREARS ON UTILITY BILLS</b>	i	0	
<b>HS030: ARREARS ON HIRE PURCHASE INSTALMENTS OR OTHER LOAN PAYMENTS</b>	i	0	
<b>HS040: CAPACITY TO AFFORD PAYING FOR ONE WEEK ANNUAL HOLIDAY AWAY FROM HOME</b>	i	0	
<b>HS050: CAPACITY TO AFFORD A MEAL WITH MEAT, CHICKEN, FISH (OR VEGETARIAN EQUIVALENT) EVERY SECOND DAY</b>	i	0	
<b>HS060: CAPACITY TO FACE UNEXPECTED FINANCIAL EXPENSES</b>	i	0,84	
<b>HS070: DO YOU HAVE A TELEPHONE (INCLUDING MOBILE PHONE)?</b>	i	0	All households in Denmark has a at least one telephone. We do not ask
<b>HS080: DO YOU HAVE A COLOUR TV?</b>	i	0	
<b>HS090: DO YOU HAVE A COMPUTER?</b>	i	0	
<b>HS100: DO YOU HAVE A WASHING MACHINE?</b>	i	0	
<b>HS110: DO YOU HAVE A CAR?</b>	i	0	
<b>HS120: ABILITY TO MAKE ENDS MEET</b>	i	0	
<b>HS130: LOWEST MONTHLY INCOME TO MAKE ENDS MEET</b>	i	10,07	
<b>HS140: FINANCIAL BURDEN OF THE TOTAL HOUSING COST</b>	i	0,40	
<b>HS150: FINANCIAL BURDEN OF THE REPAYMENT OF DEBTS FROM HIRE PURCHASES OR LOANS</b>	i	0,20	
<b>HS160: PROBLEMS WITH THE DWELLING: TOO DARK, NOT ENOUGH LIGHT</b>	i	0,37	
<b>HS170: NOISE FROM NEIGHBOURS OR FROM THE STREET</b>	i	0,25	
<b>HS180: POLLUTION, GRIME OR OTHER ENVIRONMENTAL PROBLEMS</b>	i	0,35	
<b>HS190: CRIME VIOLENCE OR VANDALISM IN THE AREA</b>	i	0,37	
<b>HY010: TOTAL HOUSEHOLD GROSS INCOME</b>	r	0	
<b>HY020: TOTAL DISPOSABLE HOUSEHOLD INCOME</b>	r	0	
<b>HY022: TOTAL DISPOSABLE HOUSEHOLD INCOME BEFORE SOCIAL TRANSFERS OTHER THAN OLDAGE AND SURVIVOR'S</b>	r	0	imp

<b>BENEFITS</b>			
<b>HY023: TOTAL DISPOSABLE HOUSEHOLD INCOME BEFORE SOCIAL TRANSFERS INCLUDING OLDAGE AND SURVIVOR'S BENEFITS</b>	r	0	imp
<b>HY025: WITHIN-HOUSEHOLD NON-RESPONSE INFLATION FACTOR</b>	r	0	
<b>HY030G/HY030N: IMPUTED RENT</b>	r	0	Cf. section 3.2.1 of the quality report
<b>HY040G/HY040N: INCOME FROM RENTAL OF A PROPERTY OR LAND</b>	r	0	
<b>HY090G/HY090N: INTEREST, DIVIDENDS, PROFIT FROM CAPITAL INVESTMENTS IN UNINCORPORATED BUSINESS</b>	r	0	Negative values occurs. The concept is calculated as a net-concept. E.g. interest recieved from bank accounts etc. are deducted interest paid on consumer loans etc. If interest paid exceeds capital income <b>HY090 will be negative.</b>
<b>HY050G/HY050N: FAMILY/CHILDREN RELATED ALLOWANCES</b>	r	0	Only information about allowances known by authorities are included. Typically obligatory allowances and allowances for which tax-deductions can be obtained. Other forms of regular allowances are not common in Denmark. In the 2003 and 2004 survey we had some qustions about voluntary allowances, but the incidences was low and data quality was bad, so we decided not to use the data.
<b>HY060G/HY060N: SOCIAL EXCLUSION NOT ELSEWHERE CLASSIFIED</b>	r	0	
<b>HY070G/HY070N: HOUSING ALLOWANCES</b>	r	0	
<b>HY080G/HY080N: REGULAR INTER-HOUSEHOLD CASH TRANSFER RECEIVED</b>	r	0	
<b>HY100G/HY100N: INTEREST REPAYMENTS ON MORTGAGE</b>	r		
<b>HY110G/HY110N: INCOME RECEIVED BY PEOPLE AGED UNDER 16</b>	r	0	
<b>HY120G/HY120N: REGULAR TAXES ON WEALTH</b>	r	0	
<b>HY130G/HY130N: REGULAR INTER-HOUSEHOLD CASH TRANSFER PAID</b>	r	0	Cf. HY050
<b>HY140G/HY140N: TAX ON INCOME AND SOCIAL CONTRIBUTIONS</b>	r	0	
<b>HY145N: REPAYMENTS/RECEIPTS FOR TAX ADJUSTMENT</b>	na	na	

<b>PERSONAL DATA (P-FILE)</b>			
<b>PB010: YEAR OF THE SURVEY</b>	2005	na	
<b>PB020: COUNTRY</b>	DK	na	
<b>PB030: PERSONAL ID</b>	s	na	
<b>PB040: PERSONAL CROSS-SECTIONAL WEIGHT</b>	c	na	
<b>PB050: PERSONAL BASE WEIGHT</b>	c	nal	
<b>PB060: PERSONAL CROSS-SECTIONAL WEIGHT FOR SELECTED RESPONDENT</b>	c	na	

<b>PB070: PERSONAL DESIGN WEIGHT FOR SELECTED RESPONDENT</b>	<b>c</b>	<b>na</b>	
<b>PB080: PERSONAL BASE WEIGHT FOR SELECTED RESPONDENT</b>	<b>c</b>	<b>na</b>	
<b>PB090: DAY OF THE PERSONAL INTERVIEW</b>	<b>b</b>	<b>0</b>	
<b>PB100: MONTH OF THE PERSONAL INTERVIEW</b>	<b>b</b>	<b>0</b>	
<b>PB110: YEAR OF THE PERSONAL INTERVIEW</b>	<b>2005</b>	<b>0</b>	
<b>PB120: MINUTES TO COMPLETE THE PERSONAL QUESTIONNAIRE</b>	<b>s</b>	<b>0</b>	
<b>PB130: MONTH OF BIRTH</b>	<b>r</b>	<b>0</b>	
<b>PB140: YEAR OF BIRTH</b>	<b>r</b>	<b>0</b>	
<b>PB150: SEX</b>	<b>r</b>	<b>0</b>	
<b>PB160: FATHER ID</b>	<b>i/r</b>	<b>0</b>	
<b>PB170: MOTHER ID</b>	<b>i/r</b>	<b>0</b>	
<b>PB180: SPOUSE/PARTNER ID</b>	<b>i/r</b>	<b>0</b>	
<b>PB190: MARITAL STATUS</b>	<b>r</b>	<b>0</b>	
<b>PB200: CONSENSUAL UNION</b>	<b>i</b>	<b>0</b>	
<b>PB210: COUNTRY OF BIRTH</b>	<b>r</b>	<b>0</b>	
<b>PB220A: CITIZENSHIP 1</b>	<b>r</b>	<b>0</b>	
<b>PB220B: CITIZENSHIP 2</b>	<b>m</b>		<b>We have no knowlegde about a possible CITIZENSHIP 2</b>
<b>PE010: CURRENT EDUCATION ACTIVITY</b>	<b>r</b>	<b>0</b>	
<b>PE020: ISCED LEVEL CURRENTLY ATTENDED</b>	<b>r</b>	<b>0</b>	
<b>PE030: YEAR WHEN HIGHEST LEVEL OF EDUCATION WAS ATTAINED</b>	<b>r</b>	<b>33</b>	
<b>PE040: HIGHEST ISCED LEVEL ATTAINED</b>	<b>r</b>	<b>2</b>	
<b>PH010: GENERAL HEALTH</b>	<b>i</b>	<b>0,02</b>	
<b>PH020: SUFFER FROM ANY A CHRONIC (LONG-STANDING) ILLNESS OR CONDITION</b>	<b>i</b>	<b>0,02</b>	
<b>PH030: LIMITATION IN ACTIVITIES BECAUSE OF HEALTH PROBLEMS</b>	<b>i</b>	<b>0,02</b>	
<b>PH040: UNMET NEED FOR MEDICAL EXAMINATION OR TREATMENT</b>	<b>i</b>	<b>0,02</b>	
<b>PH050: MAIN REASON FOR UNMET NEED FOR MEDICAL EXAMINATION OR TREATMENT</b>	<b>i</b>	<b>0</b>	
<b>PH060: UNMET NEED FOR DENTAL EXAMINATION OR TREATMENT</b>	<b>i</b>	<b>0,02</b>	
<b>PH070: MAIN REASON FOR UNMET NEED FOR DENTAL EXAMINATION OR TREATMENT</b>	<b>i</b>	<b>0,02</b>	
<b>PL015: PERSON HAS EVER WORKED</b>	<b>i/r</b>	<b>0</b>	
<b>PL020: ACTIVELY LOOKING FOR A JOB</b>	<b>i</b>	<b>34,38</b>	<b>By a mistake the question is only asked, if the person is unemployed PL030 = 3. It is not asked for other people not in work. Therefore the item nonresponse is big. In the 2007 survey the question has been asked all people not in work.</b>
<b>PL025: AVAILABLE FOR WORK</b>	<b>i</b>		<b>cf. PL020</b>
<b>PL030: SELF-DEFINED CURRENT ECONOMIC STATUS</b>	<b>i/r</b>	<b>1,11</b>	
<b>PL035: WORKED AT LEAST 1 HOUR DURING THE PREVIOUS WEEK</b>	<b>na</b>	<b>-</b>	
<b>PL040: STATUS IN EMPLOYMENT</b>	<b>i/r</b>	<b>?</b>	<b>Problems with the definition of flaggs</b>
<b>PL050: OCCUPATION (ISCO-88 (COM))</b>	<b>i/r</b>	<b>1,49</b>	

<b>PL060: NUMBER OF HOURS USUALLY WORKED PER WEEK IN MAIN JOB</b>	i	0,10	
<b>PL070: NUMBER OF MONTHS SPENT AT FULL-TIME WORK</b>	i/r	1,09	
<b>PL072: NUMBER OF MONTHS SPENT AT PART-TIME WORK</b>	i/r	1,09	
<b>PL080: NUMBER OF MONTHS SPENT IN UNEMPLOYMENT</b>	i/r	1,09	
<b>PL085: NUMBER OF MONTHS SPENT IN RETIREMENT</b>	i/r	1,09	
<b>PL087: NUMBER OF MONTHS SPENT STUDYING</b>	i/r	1,09	
<b>PL090: NUMBER OF MONTHS SPENT IN INACTIVITY</b>	i/r	1,09	
<b>PL100: TOTAL NUMBER OF HOURS USUALLY WORKED IN SECOND, THIRD... JOBS</b>	i	100	Caused by a programming error PL100 is missing in the Eurostat-files in 2005. It is available in national files.
<b>PL110: NACE</b>	r	1,18	
<b>PL120: REASON</b>	i		0,15 (Number of hours in PL100 is missing)
<b>PL130: NUMBER OF PERSONS WORKING AT THE LOCAL UNIT</b>	r		8,26
<b>PL140: TYPE OF CONTRACT</b>	r	-	The value is always '1'. The values are extracted from a register. Temporary contract are not much used in Denmark, but there should be some, so obviously the data from the register is erroneous. We are working on finding a solution.
<b>PL150: MANAGERIAL POSITION</b>	i	0,02	
<b>PL160: CHANGE OF JOB SINCE LAST YEAR</b>	i/r	ol	
<b>PL170: REASON FOR CHANGE</b>	i	ol	
<b>PL180: MOST RECENT CHANGE IN THE INDIVIDUAL'S ACTIVITY STATUS</b>	i/r	ol	
<b>PL190: WHEN BEGAN FIRST REGULAR JOB</b>	i	ol	
<b>PL200: NUMBER OF YEARS SPENT IN PAID WORK</b>	i	ol	
<b>PL210A: MAIN ACTIVITY ON JANUARY</b>	i/r	ol	
<b>PL210B: MAIN ACTIVITY ON FEBRUARY</b>	i/r	ol	
<b>PL210C: MAIN ACTIVITY ON MARCH</b>	i/r	ol	
<b>PL210D: MAIN ACTIVITY ON APRIL</b>	i/r	ol	
<b>PL210E: MAIN ACTIVITY ON MAY</b>	i/r	ol	
<b>PL210F: MAIN ACTIVITY ON JUNE</b>	i/r	ol	
<b>PL210G: MAIN ACTIVITY ON JULY</b>	i/r	ol	
<b>PL210H: MAIN ACTIVITY ON AUGUST</b>	i/r	ol	
<b>PL210I: MAIN ACTIVITY ON SEPTEMBER</b>	i/r	ol	
<b>PL210J: MAIN ACTIVITY ON OCTOBER</b>	i/r	ol	
<b>PL210K: MAIN ACTIVITY ON NOVEMBER</b>	i/r	ol	
<b>PL210L: MAIN ACTIVITY ON DECEMBER</b>	i/r	ol	
<b>PY010G/PY010N: EMPLOYEE CASH OR NEAR CASH INCOME</b>	r	0a	
<b>PY020G/PY020N: NON-CASH EMPLOYEE INCOME</b>	r	0	
<b>PY030G: EMPLOYER'S SOCIAL INSURANCE CONTRIBUTION</b>	r	0	
<b>PY035G/PY035N: CONTRIBUTIONS TO INDIVIDUAL PRIVATE PENSION PLANS</b>	r	0	

<b>PY050G/PY050N: CASH BENEFITS OR LOSSES FROM SELF-EMPLOYMENT</b>	<b>r</b>	<b>0</b>	
<b>PY070G/PY070N: VALUE OF GOODS PRODUCED BY OWN-CONSUMPTION</b>	<b>r</b>	<b>0</b>	
<b>PY080G/PY080N: PENSION FROM INDIVIDUAL PRIVATE PLANS</b>	<b>r</b>	<b>0</b>	
<b>PY090G/PY090N: UNEMPLOYMENT BENEFITS</b>	<b>r</b>	<b>0</b>	
<b>PY100G/PY100N: OLD-AGE BENEFITS</b>	<b>r</b>	<b>0</b>	
<b>PY110G/PY110N: SURVIVOR' BENEFITS</b>	<b>r</b>	<b>0</b>	
<b>PY120G/PY120N: SICKNESS BENEFITS</b>	<b>r</b>	<b>0</b>	
<b>PY130G/PY130N: DISABILITY BENEFITS</b>	<b>r</b>	<b>0</b>	
<b>PY140G/PY140N: EDUCATION-RELATED ALLOWANCES</b>	<b>r</b>	<b>0</b>	
<b>PY200G: GROSS MONTHLY EARNINGS FOR EMPLOYEES</b>	<b>r</b>	<b>0</b>	

# Longitudinal files 2003-2005: An overview

## General comments

The longitudinal files 2003-2005 includes the *rotational groups* 3 or 4 selected in 2003 cf. 2.3.3.2-2.3.3.3.

As mentioned in the document of Marianne Hubiers only sample persons included in 2003, 2004 and 2005 cross-sectional R-Files have been assigned longitudinal weights. Housholds are selected indirectly cf. 2.1.2.

## Files 2003

### DFILE-2003

Households are included in DFILE-2003, if the household was selected for the 2003 cross-sectional SILC and the *rotational group* (DB075) is 3 or 4

Number of households in DFILE-2003..... 4782

Number of households in DFILE-2003, where household questionnaire was completed. *Result\_2003* (DB130) = 11 and rotational group is 3 or 4.....3205

### HFILE 2003

Households are included if *result\_2003* is 11.

Number of households.....3205

- households not included in HFILE-5.....757

Number of households in HFILE-2003 and HFILE-2005.....2448

### RFILE-2003

A person is included, if the household is included in HFILE-2003 and the person is household member, *membership status* (RB110)=1. RB110 is always 1 in 2003

Number of persons in RFILE-2003..... 8.186

Number of sample persons (RB110=1).....6.646

Number of co-residents (RB110=2).....1.540

RB245: *Respondent status* can take the value 2, 3, 4 in Denmark.

RB245=2: Number of elected respondents..... 3.205

RB245=3: Number of not selected eligible households.....3.227

RB245=4: Number of not eligible households (household members aged less than 16).....1.754

All households have one and only one member, where RB245 is 2, namely the person initially selected.



*PFILE-2003*

A person is included in PFILE-2003 if he is included in RFILE\_2003 and aged 16+

Number of persons in PFILE\_2003.....6.432

**Files 2004**

*DFILE 2004*

All households included in DFILE-2003 are included in DFILE-2004, if *result\_2003* is 11,

Total number of households in DFILE-2004.....3205

DFILE-2004 contains information about the household to which the selected person belonged in 2004. Therefore the variable *household status* (DB110) is used differently from, what is literally stipulated in DOC 065 cf. below.

Number  
of cases

DB110=1: The selected person is still in scope including cases, where the person moved to another address in 2003.....	3170
DB110=3: The selected person moved to a collective household or institution within the country in 2003.....	12
DB110=4: The selected person moved outside the country in 2003.....	7
DB110=5: Selected person died in 2003.....	15
DB110=7: Selected person could not be found in the register of population.....	1
DB110=10: Selected person moved into a another household in the sample.....	0

In case of Denmark DB110 cannot be 6, 8 or 9. Households, which do not contain a selected person, is out of scope. We do not have split-off households and we do not include new addresses, households or selected persons in the longitudinal component.

*HFILE 2004*

A household is included in HFILE-2004 ,if it is included in DFILE-2004 and *result\_2004* is 11.

Number of households in HFILE-2004.....	2629
of these not included in HFILE-2005.....	404
Number of households in HFILE-2003, HFILE-2004 and HFILE-2005.....	2225

*RFILE-2004*

A person is included in RFILE-2004, if the household is

included in HFILE-2004, and the person was household member 1 January 2004.

Number of persons in RFILE-2004.....	6.459
Number of sample persons (RB110=1).....	5.047
Number of co-residents (RB110=2).....	1.412

Membership status (RB110) can take the values 1, 2, 3 or 4. Denmark does not include former household members in the longitudinal component, so RB110 cannot be 5, 6 or 7 in Denmark.

RB110=1: Number of persons from previous waves.....	6.238
RB110=2: Number of persons moved in from another sample household 2003 since last wave.....	0
RB110=3: Number of persons moved in from outside the sample 2003.....	158
RB110=4: Number of children born into the household since last wave.....	63

RB245: *Respondent status* can take the value 2, 3, 4 in Denmark.

RB245=2: Number of elected respondents.....	2.629
RB245=3: Number of not selected eligible households.....	2.442
RB245=4: Number of not eligible households (household members aged less than 16).....	1.388

All households have one and only one member, where RB245 is 2, namely the person initially selected.

*PFILE-20043*

A person is included in PFILE-2004, if he is included in RFILE\_2003 and aged 16+

Number of persons in PFILE-2004.....	5.071
--------------------------------------	-------

**Files 2005**

*DFILE 2005*

All households included in DFILE-2003 are included in DFILE-2005, if *result\_2003* is 11 and the household was in scope 2004 or in 2005.

Total number of households in DFILE-2005.....	3170
---	------

Number  
of cases

DB110=1: The selected person is still in scope including cases, where the person moved to another address in 2004.....	3116
DB110=3: The selected person moved to a collective household or institution within the country in 2004.....	7

DB110=4: The selected person moved outside the country in 2004.....	13
DB110=5: Selected person died in 2004.....	34
DB110=7: Selected person could not be found in the register of population.....	0
DB110=10: Selected person moved into a another household in the sample.....	0

In case of Denmark DB110 cannot be 6, 8 or 9. Households, which do not contain a selected person, is out of scope. We do not have split-off households and we do not include new addresses, households or selected persons in the longitudinal component.

*HFILE 2005*

A household is included in HFILE-2005 ,if it is included in DFILE-2005 and *result\_2005* is 11.

Number of households.in HFILE-2005.....	2448
of these not included in HFILE-2004.....	223

*RFILE-2005*

A person is included in RFILE-2005, if the household is included in HFILE-2005, and the person was household member 1 January 2005.

Number of persons in RFILE-2005.....	6.301
Number of sample persons (RB110=1).....	4.713
Number of co-residents (RB110=2).....	1.588

Membership status (RB110) can take the values 1, 2, 3 or 4. Denmark does not include former household members in the longitudinal component, so RB110 cannot be 5, 6 or 7 in Denmark.

RB110=1: Number of persons from previous waves.....	5824
RB110=2: Number of persons moved in from another sample household 2003 since last wave.....	0
RB110=3: Number of persons moved in from outside the sample 2003.....	.411
RB110=4: Number of children born into the household since last wave.....	66

RB245: *Respondent status* can take the value 2, 3, 4 in Denmark.

RB245=2: Number of elected respondents.....	2.448
RB245=3: Number of not selected eligible households.....	2.450
RB245=4: Number of not eligible households (household members aged less than 16).....	1.403

All households have one and only one member, where RB245 is 2, namely the person initially selected.

*PFILE-2005*

A person is included in PFILE-2005, if he is included in RFILE\_2003 and aged 16+

Number of persons in PFILE-2005.....4.898

## Longitudinal files 2004-2005: An overview

### General comments

The longitudinal files 2004-2005 includes the *rotational groups* 3 or 4 selected in 2003 and 1 selected in 2004 cf. 2.3.3.2-2.3.3.3.

As mentioned in the document of Marianne Hubiers only sample persons included in the 2003, 2004 and 2005 cross-sectional R-Files have been assigned longitudinal weights. Households are selected indirectly cf. 2.1.2.

### Files 2004

#### *DFILE-2004*

Households are included in DFILE-2004, if the household was selected for the 2004 cross-sectional SILC and the *rotational group* (DB075) is 3, 4 or 1

Number of households in DFILE-2004..... 7142

Number of households in DFILE-2004, where household questionnaire was completed. *Result\_2004* (DB130) = 11 and rotational group is 3, 4 or 1.....4586

#### *HFILE 2004*

Households are included in HFILE\_2004 if *result\_2004* is 11.

Number of households.....4586

number of these not included in HFILE-2005.....856

Number of households in HFILE-2004 and HFILE-2005.....3730

#### *RFILE-2004*

A person is included, if the household is included in HFILE-2004 and the person is household member, *membership status* (RB110)=1. RB110 is always 1 in 2004

Number of persons in RFILE-2004.....11.570

RB245: *Respondent status* can take the value 2, 3, 4 in Denmark.

RB245=2: Number of elected respondents.....4.586

RB245=3: Number of not selected eligible households.....4.478

RB245=4: Number of not eligible households (household members aged less than 16)..... **2.506**

All households have one and only one member, where RB245 is 2, namely the person initially selected.

*PFILE-2004*

A person is included in PFILE-2004 if he is included in RFILE\_2004 and aged 16+

Number of persons in PFILE\_2004.....9.064

## Files 2005

### *DFILE 2005*

All households included in DFILE-2005 are included in DFILE-2005, if *result\_2004* is 11,

Total number of households in DFILE-2005.....4586

DFILE-2005 contains information about the household to which the selected person belonged in 2005. Therefore the variable *household status* (DB110) is used differently from, what is literally stipulated in DOC 065 cf. below.

	Number of cases
DB110=1: The selected person is still in scope including cases, where the person moved to another address in 2004.....	4.538
DB110=3: The selected person moved to a collective household or institution within the country in 2004.....	6
DB110=4: The selected person moved outside the country in 2004.....	14
DB110=5: Selected person died in 2004.....	27
DB110=7: Selected person could not be found in the register of population.....	0
DB110=10: Selected person moved into a another household in the sample.....	0
DB110=Lost household. No information about what happened to the selected person.....	1

In case of Denmark DB110 cannot be 6, 8 or 9. Households, which do not contain a selected person, is out of scope. We do not have split-off households, and we do not include new addresses, households or selected persons in the longitudinal component.

### *HFILE 2005*

A household is included in HFILE-2005 ,if it is included in DFILE-2005 and *result\_2005* is 11.

Number of households in HFILE-2005.....3.730

### *RFILE-2005*

A person is included in RFILE-2005, if the household is included in HFILE-2005, and the person was household member 1 January 2005.

Total number of persons in RFILE-2005.....9.722

Membership status (RB110) can take the values 1, 2, 3 or 4. Denmark does not include former household members in the longitudinal component,

so RB110 cannot be 5, 6 or 7 in Denmark.

RB110=1: Number of persons from previous waves.....	9.085
RB110=2: Number of persons moved in from another sample.....	
household 2003 since last wave.....	0
RB110=3: Number of persons moved in from outside the sample	
2003.....	522
RB110=4: Number of children born into the household since last	
wave.....	115

RB245: *Respondent status* can take the value 2, 3, 4 in Denmark.

RB245=2: Number of elected respondents.....	3.730
RB245=3: Number of not selected eligible households.....	3.766
RB245=4: Number of not eligible households (household members	
aged less than 16).....	<b>2.226</b>

All households have one and only one member, where RB245 is 2, namely the person initially selected.

*PFILE-2005*

A person is included in PFILE-2005, if he is included in RFILE\_2003 and aged 16+

Number of persons in PFILE-2005.....	7.496
-2005 have been assigned weights.	