Final Quality Report SILC2007 - BELGIUM

0. Introduction

This report contains a description of the accuracy, precision and comparability of the Belgian SILC2004 to SILC2006-surveydata. It is structured following the guidelines in the commission regulation (EC) no. 28/2004. This results in three chapters:

1. Indicators
2. Accuracy
3. Comparability
4. Coherence

1. Indicators

For the common longitudinal EU indicators based on the longitudinal sample of EU-SILC we refer the readers to the EUROSTAT website where these indicators are available in a dynamic way.

2. Accuracy

For second and following waves of the longitudinal component the following information has to be provided

2.1 Sampling Design

2.1.1 Type of sampling (stratified, multi-stage, clustered)

The Belgian EU-SILC 2007 survey follows a stratified 2-stage sampling.

2.1.2 Sampling units (one stage, two stages)

Primary units:
The *Primary Sampling Units* are the municipalities (or part thereof in the larger ones); in each of the 11 strata, they were drawn PPS, i.e. with repetitions allowed (for instance, Schaerbeek was drawn 6 times). In total, 275 draws were made in 2004, once forever (for the whole duration of EU-SILC).
Secondary units:
The *Final Sampling Units* are the (private) households.

Recall that, in 2004, 40 households had been selected in each PSU, numbered 1 to 40. The first 10 (whether or not they responded irrelevant) vanished from the panel in 2005, the other 30 (including possible split-offs) were followed according to the tracing rules.

Hence, the (cross-sectional) sample of SILC 2007 consists of
- "old" households (drawn between 2004 and 2006)

and
- "new" households (drawn in 2007, staying until 2010).

In fact, it is only the selection of the new households that gave us some degree of freedom (see in particular 2.1.4)

In the D-file, three variables have been added:
- ✓ DB061 is the identification of the primary units (concatenation of 5 digits for the municipalities and one letter).
- ✓ DB063 is the 'multiplicity order', the number of times each PSU was drawn in the sample.
- ✓ DB071 is the order of selection of the new households within each letter.

2.1.3 Stratification and sub-stratification criteria
The stratification criterion is the region (NUTS2 level). The 11 strata are the 10 provinces of Belgium and the Brussels Capital Region.

2.1.4 Sample size and allocation criteria
In 2007 we managed to keep the number of responding households close to 6000, drawing 16 new hh in each PSU.

Table 1: sample size and achieved response by NUTS2-units

NUTS2	Name	Old (or strange) hh	New hh	Total hh	Accepted hh (DB135=1)
BE10	Brussels	887	800	1687	856
BE21	Antwerpen	847	688	1535	947
BE22	Limburg	368	256	624	461
BE23	Oost-Vlaanderen	612	512	1124	761
BE24	Vlaams-Brabant	494	416	910	611
BE25	West-Vlaanderen	566	334	900	680
BE31	Brabant Wallon	201	128	329	180
BE32	Hainaut	774	590	1364	851
BE33	Liège	536	399	935	580
BE34	Luxembourg	149	93	242	186
BE35	Namur	219	157	376	235
Total	**Belgium**	5653	4373	10026	6348

2.1.5 Sample selection schemes
Systematic sampling of secondary units (new households) in each primary unit selected, the households have been ordered according to the age of the reference person.

2.1.6 Sample distribution over time

2.1.7 Renewal of sample: Rotational groups
See above.

2.1.8 Weightings

Recall that, for the first year of the panel (=SILC 2004 in Belgium), the computation of weights involved three stages (described in 134-04)
- (a)   initial weights
- (b)   weights corrected for nonresponse
- (c)   final (calibrated) weights

For 2007, a distinction has to be made between

"old" households   i.e. households that contain at least one sample person who took part in 2006, and had to be surveyed again in 2007 according to the rotation and tracing rules (excluding the outgoing fourth) (household composition may have changed, whence quotations marks)

"new" households   i.e. households that were drawn for the first time in 2007, among those households not containing any sample person already drawn before (quotations marks superfluous)

This distinction pertains to initial weights and nonresponse correction
>   Since the "old" households are selected indirectly from the 2004, 2005 or 2006 samples, and household composition may have changed, some kind of "weight sharing" must be applied to determine the (2007) initial weights, or rather base weights. On the other hand, "new" households have their own inclusion probability, whose inverse gives the initial weights;
>   For the "old" households, (2007) nonresponse=attrition can be linked with (2006) SILC information. For the "new" households, all we can rely upon to explain initial nonresponse is auxiliary information (household size, urban/rural character...) from the Population Register.

On the other hand,
>   Calibration can be done together for "old" and "new" households. With respect to our 2004 model, we decided in 2005 to relax the constraints (basically, calibrating at NUTS1-level instead of NUTS2), in order to decrease the standard deviation of weights.

This introduces the following sections
> **2.1.8.1 Initial weights for the new households**
> **2.1.8.2 Nonresponse correction for the new households**
> **2.1.8.3 Base weights for the old households**
> **2.1.8.4 Attrition correction for the old households**
> **2.1.8.5 Calibration (all households)**

2.1.8.1.   Initial weights for the new households

Belgium chose to draw the Primary Sampling Units (= municipalities or parts thereof) "forever", and to rotate the Secondary Sampling Units (=households) within the selected PSU's.

The 2004 PPS two-stage sampling design was self-weighting within each stratum h: x denoting any households in municipality X), we had (in 2004)

$$P(x \text{ drawn}) = P(x \text{ drawn}|X \text{ drawn}) \cdot P(X \text{ drawn}) = n_h/N_X \cdot N_X/N_h \cdot g_h = n_h/N_H \cdot g_h,$$

where

n_h denotes the number of households to be drawn in the (selected) PSU (viz. 40)

N_X the number of households in the PSU (in 2004)

N_h the number of households in the stratum (in 2004)

g_h the number of PSU's drawn in the stratum.

(This is an oversimplification, since PSU are drawn with repetition; the selection probability for a PSU should be replaced by the expectation of selection multiplicity, and the term 40 by a multiple depending on the selection multiplicity…but the idea is the same).

In 2007, the picture has become

$$P(x \text{ drawn}) = P(x \text{ drawn}|X \text{ drawn}) \cdot P(X \text{ drawn}) = m_h/M_X \cdot N_X/N_h \cdot g_h, \text{ where}$$

m_h is the number of households to be drawn in the (selected) PSU (depending on h)

M_X is the number of households in the PSU (in 2007)

The factor N_X/M_X indicates the increase-decrease in inclusion probabilities in PSU X (still assuming X has been drawn) between 2007 and 2004.

Now it would seem logical to replace N_X by a smaller number, to account for the households[1] already drawn in 2004, 2005 or 2006, whence immunized from being drawn again in 2007.

However, the following argument shows that (assuming momentarily that X has been drawn and that the population figures N_X and M_X remain stable) matters are not so easy:

$$P(x \text{ drawn in 2007}) =$$
$$(P(x \text{ drawn in 2007}|x \text{ drawn before}) \cdot P(x \text{ drawn before})) +$$
$$(P(\text{drawn in 2007}|x \text{ not drawn before}) \cdot P(x \text{ not drawn before}),$$

the first term vanishes and the second equals $n_h/(M_X-b) \cdot (N_X-b)/N_h$, where b denotes the number of hh already drawn; since both fraction terms are much larger than b (at least 900 in all selected PSU's), the ratio $(N_X-b)/(M_X-b)$ is (close to 1, and) very close to N_X/M_X. Since the term b is an approximation anyway, we chose to stick to $m_h/M_X \cdot N_X/N_h \cdot g_h$ as inclusion probabilities, and its inverse for initial weights **INIwei=DB080**. Note that, with this concept of DB080, the "new" hh correspond to the total Belgian population (some 4,5 millions private hh); before calibrating, theses weights will be scaled down "to make room" for the old hh; recovering the strange hh means that the sum of the pre-calibration weights will be slightly larger than 4,5 millions (average of g-weights slightly less than 1)

2.1.8.2. Nonresponse correction for the new households

Following Eurostat's suggestion (see Document 065, WEIGHTING II. WEIGHTING FOR THE FIRST YEAR OF EACH SUB-SAMPLE), we replaced the homogeneous response groups (based on household size crossed with urbanity) ratio by a multiple regression model (based on the same dummy variables). By "responding", we mean only those households whose results were accepted (DB135=1). For technical

---

[1] Perhaps a bit less (households that vanished already subtracted) or a bit more (split households, both components of which stayed in PSU, should be subtracted twice)

reasons, we used linear regression instead of logistic; since the (predicted) response turned out to be close to 50% for all categories, this is harmless.

The file was split by NUTS1 and the following variables were used
- DB100 = urbanity (constant in BE1 = Brussels; 3 values, so 2 dummies needed in model, elsewhere)
- HOUSEHOLD size, recoded into the four values "one", "two", "three" and "four or more" (so three dummies)

The regression produced a new variable "expresp", allowing us to define
NRwei = INIwei/expresp

2.1.8.3 Attrition for the old households

Before "sharing" the 2006 weights, a correction for attrition should be introduced. This year, we elected to perform this correction at the level of individuals, since a 2006 sample person either stays in the panel or leaves it (rotated out, left population, noncontact, refusal or inability to respond, while the structure of a household can change. Note that all household characteristics (e.g. HH020) can be distributed to the members.

This year, we chose to separate the "Children" (for which only basic personal information from the R-file and the distributed H-file is available) from the "Adults" (present in the 2006 P-file as well), i.e. those persons born in 1989 or before.

In the children's model, the following predictors (all, except the last, from the 2006 file – although this does not matter much for group A) were used, grouped by type
A. individual demographic information: age[2] from RB080, sex = RB090, country of birth (= pb210 for adults, but available for children too in our Belgian files);

---

[2] Let us start with a picture (Z in function of age class, "1" denoting the range 0-4, …, "17" the range "80-84", "18' corresponds to '85 or older", age computed here as 2006-rb080)

The highest 2 scores are depicted in white, the lowest 2 in dark blue. We distinguish two local maxima (one among children 5-9, the other one in the area of "old but not too old") and two local minima (one among "young adults" and one for "very old".

6

B. <u>housing information</u>: dwelling type = HH010 and tenure = HH020
C. <u>household type</u>: a limited number of dummies, as there is at least one dependent child;
D. <u>monetary indicators</u>: we refrained from taking the equivalised income (outliers), but took a transform of it, as well as the dummy "poor or not" and the subjective ability to make ends meet = HS120
E. <u>sampling and rotation</u>: number of years in panel (from DB075) and urbanisation (=DB100)
F. one variable (paradata) related to <u>fieldwork</u> in 2006 (computed from HB040 and HB050)
G. one variable indicating a possible <u>change of interviewer</u> (suitably imputed for hh that did not participate in 2007)

For the adults, the same predictors were used, and moreover
H. variables from the <u>P-file</u> (related to education level and health);
I. a "<u>Belgian</u>" variable, corresponding to satisfaction with the society in general)
were integrated.

We used linear regression; (with some truncation, when the estimated response propensity turned out to be larger than one).

2.1.8.4 Weight sharing
We followed Eurostat's recommendation "EU-SILC weighting procedures: an outline" and shared the calibrated 2006 weights, after correcting for attrition (instead of the initial weights, see Lavallée).

perc resp
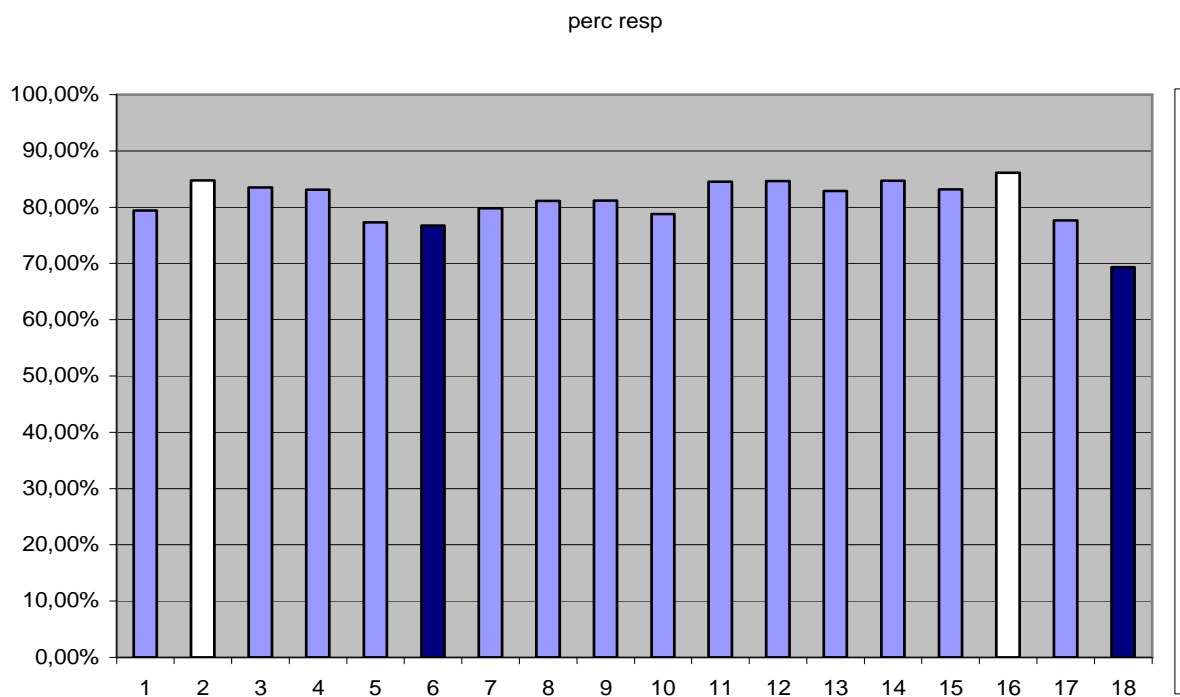
This can be illustrated by an imaginary example, dealing simultaneously with fusions (persons *A&B* in same 2006 hh, *C* in another 2006 hh, so "fusion" in the sense of DB110 occurs), new members (a baby like *E* or already in population like *D*); we focus on the 2007 hh, what happened to those who co-resided with *A* and *B* or with *C* in 2006 (left or split) is irrelevant!

Note that

- RB050 = weight 2006: same for *A* & *B*, vacuous for *D* and *E*
- Newi: in general a bit larger than RB050; *A*'s differs from *B*'s (attrition correction at individual level)
- Somwe = 950+1000+850 involves only *A*, *B* and *C*
- *Weiind:* = ¼ * **somwe** (*A B C D* : four contribute to the denominator)[3]

Person in 2007 hh	A	B	C	D	E
RB110 (2007)	1	1	2	3	4
RB050 (weight 2006)	800	800	600	---	---
Newi = Weight 2006 (after attrition correction)	950	1000	850	---	---
Somwe (sum Newi over 2007 hh)	2800	2800	2800	2800	2800
Weiind	700	700	700	700	700

Weiind will be injected as "initial" weight in the final calibration job.

2.1.8.5 Calibration

We first put the pieces together: weiind is defied as

- new = started in 2007 (initial weight, corrected for initial nonresponse, scaled, see 2.1.8.1)
- old = took part in 2006 (2006 weight, corrected for attrition and weight sharing if necessary, see 2.1.8.4)
- strange = did not take part in 2006 (initial weight, non correction)

In terms of persons, the weiind statistics were

Type	# ind	Mean of weiind
NEW	4884	612,08
OLD	10009	869,74
BACK	600	408,37
Total	15493	770,65

-Children born to sample women. They receive the weight of the mother (this assumes that the baby belongs tohis/her mother's hh)

-Persons moving into sample households from outside the survey population. They receive the average of base weights of existing household members (vacuous here, as RB110 enables us to identify the newborns, but not the immigrants or the –few- persons moving from a collective to a private hh)

-Persons moving into sample households from other non-sample households in the population – these are "co-residents" and are given zero base weight.

Recall that 11 *sampling* strata were used (provinces= NUTS2); we use 3 *extrapolation* strata (the 3 NUTS1 regions BRUssels=BE1, VLAanderen=BE2 and WALlonia=BE3)

Calibration model
VLA, WAL:
SIZE4+(AGE8XSEX2)+PROV5 →20 individual[4] + 4 household constraints
BRU:
SIZE4+(AGE8XSEX2) →16 individual + 4 household constraints

Prov = province where interviewed (differs from DB040 in two cases)

Individual constraints 27=16+11 (age*sex + prov; note that each province belongs to one single region (extrapolation stratum), for the other two regions, the total is set to 0 and the condition is vacuous)

Household constraints 4 (size: "1", "2", "3 or "4 & more",)

Calibration type (after some trials and errors…): linear

2.1.8.7 Final longitudinal weights

Combination of steps above…

2.1.8.8. Final cross-sectional weights
Statistics

	N	Minimum	Maximum	Mean	Std. Dev.
Final weights	6348	156,53	4576,81	715,74	300,13

Historical remark: .

Year	n	Min	Max	Mean	Std	Calibration
2004	5275	135,49	5816,95	841,64	292,64	Exponential
2005	5137	58,18	7878,79	871,64	325,86	Truncated
2006	5860	192,51	3292,18	771,67	246,75	Linear, modified intermediate weights
2007	6348	156,53	4576,81	715,74	300,13	Linear

2.1.9 Substitutions

No substitution was applied in our survey.

---

[4] Five provinces and 16 age*sex categories, but sum over provinces = sum over age*sex

2.2 Sampling errors

Income components	Mean	Number of observations before imputation	Number of observation after imputation	Standard error
HY010	40353,7	2045	6335	598,8
HY020	30039,01	1960	6340	460,1
HY022	27011,23	1802	6143	478,6
HY023	22335,01	1696	5991	474,4
Net income components at household level				
HY030N				
HY040N				
HY090N	1489,906	1352	4314	372,8
HY050N	3590,98	2209	2276	86,1
HY060N				
HY070N				
HY080N				
HY100N				
HY110N				
HY120N				
HY130N				
HY140N				
HY145N				
Gross income components at household level				
HY030G				
HY040G	4991,754	425	463	1187,6
HY090G	9256,867	1352	4314	372,8
HY050G	1489,906	2193	2276	88,0
HY060G	3623,437	115	116	536,6
HY070G	6378,422	37	47	294,2
HY080G	1617,723	464	493	206,8
HY100G	3418,395	1610	1975	88,4
HY110G	2963,312	11	11	675,1
HY120G	1528,434			
HY130G		445	460	211,0
HY140G	3107,686	2975	5687	228,3
net income components at personal level				
PY010N	19119,1	4699	5877	264,9
PY020N	1415,211	1926	2353	50,1
PY035N				
PY050N	17297,63	533	768	1082,5
PY070N				
PY080N	7504,04	13	22	2080,8
PY090N	7895,852	1446	1556	317,4
PY100N	13493,81	2101	2240	207,9
PY110N	11400,49	91	96	552,4
PY120N	5753,421	192	212	512,4
PY130N	8902,287	397	411	361,9
PY140N	1020,369	212	219	237,5

gross income components at personal level				
PY010G	29159,06	4053	5877	492,5
PY020G	1556,991	1926	2353	60,2
PY030G				
PY035G				
PY050G	21755,2	218	768	1280,3
PY070G				
PY080G	7504,04	13	22	2080,8
PY090G	8470,936	1036	1556	341,8
PY100G	15314,45	1434	2240	287,0
PY110G	11864,44	67	96	607,6
PY120G	5939,121	153	212	540,4
PY130G	9332,932	309	411	412,0
PY140G	1020,369	212	219	237,5
PY200G				

Equivalised disposal income	Mean	Number of observations before imputation	Number of observation after imputation	Standard error
Subclasses by household size				
1 household member	16454	710*	1764*	480,0
2 household members	19548	630*	2131*	399,3
3 household members	20728	256*	991*	771,5
4 and more	19083	364*	1454*	656,8
Population by age group				
<25	18325	1366	4925	484,4
25 to 34	20921	523	1945	602,8
35 to 44	20846	652	2289	949,0
45 to 54	20560	591	2249	399,5
55 to 64	20310	554	1861	645,3
65+	15465	672	2225	271,3
Population by sex				
Male	19508	2065	7564	342,2
Female	18789	2293	7917	656,8

2.3 Non-sampling errors
2.3.2 Measurement and processing errors

Mismatch in time between household composition and household income (see also §3.1)

A number of inconsistencies result from a mismatch between the composition of the household at the moment of the interview (between September and December of year x) and the income of the previous year (year x-1).

This mismatch can bias the measurement of poverty status in several ways. For example:

- ✓ Persons who were full-time students in year x-1 (and depending on their parents), but were employed at the time of the interview (and living independently in a one person household for example) will report an income equal to 0 in year x-1 and will be wrongly classified as a poor household.

Other examples can also occur for persons where the household composition changed:

- ✓ For a housewife who was married in year x-1, but divorced and is working at the time of the survey there will also be a mismatch
- ✓ For a household which received family allowances for a student in year x-1, but where the student is no longer part of the household in year x there will also be a mismatch
- ✓ For a household with a person working in year x-1, but retired at the moment of the survey (in year x) a mismatch will also occur. Take notice of the fact that, as the examples show the bias can go in both directions: under and over reporting of income. In each one of the examples, the choice to situate the income reference period in the past is the cause, however.

- *Error in the routing wave 2004*

An error in the routing occurred for Questions H100 and H101 on the 'Revenus du patrimoine' (Interests, dividends, profit from capital investments in unincorporated business)(To be included in Variable HY090G). Only individuals responding precisely on Question H99 about 'Revenus des placements financiers' were asked to precise whether the amounts were profit or loss. For individuals responding the question H100 (not an amount but a scale value) H101 was never asked. For these cases, the incomes were considered as profit.

H 36 (HY040): if the person answered that he didn't let out a part of his house, we still asked how much the profit was.

- *Error in the routing wave 2005*

There was one error in the routing in the household questionnaire for tenants. They skipped the question "Can you tell me what is the amount you pay monthly for your consumption of electricity and gas together? Give a rough estimation. If a part of your dwelling is professionally used, give the total only for the non-professional part."

- *Error in the routing wave 2006*

There was one error in the routing. In the household questionnaire, in the part concerning childcare, the selection was made on the base of actual age instead of age

in the income reference period. So we missed information for some children born in 1993 or 1994.

- *Error in the routing wave 2007*

There was one error in the routing. In the household questionnaire, in the part concerning childcare, the selection was made on the base of actual age instead of age in the income reference period. So we missed information for some children born in 1994.

- *Correspondence French/Dutch versions of Questionnaires wave 2004*

There was no mistake in the formulation of the French/Dutch versions.in 2004.

- *Correspondence French/Dutch/German versions of Questionnaires wave 2005*

For the question about the mode of contact, the French version was wrongly asking whether the **household** was contacted where the Dutch version asked whether the **address** was contacted.

In the German version, question I8. 'Retirement' is coded 8 as it is coded 7 in the other languages because 'Student' and 'Unpaid work experience' were unfortunately split in 2 codes (6 & 7). Other consequence: 'Permanently disabled' and 'Fulfilling domestic tasks' were collected on the same code (9). We estimate that 0,18% of the response on this question could have been influenced by this.

- *Correspondence French/Dutch/German versions of Questionnaires wave 2006*

For the question about the mode of contact, the French version was wrongly asking whether the **household** was contacted where the Dutch version asked whether the **address** was contacted.

In the German version, question I8. 'Retirement' is coded 8 as it is coded 7 in the other languages because 'Student' and 'Unpaid work experience' were unfortunately split in 2 codes (6 & 7). Other consequence: 'Permanently disabled' and 'Fulfilling domestic tasks' were collected on the same code (9). We estimate that about 0,2% of the response on this question could have been influenced by this.

- *Differently asked questions*

HH050: The question in 2004 did not point out that the inability to keep home adequately warm was the **inability to pay** to keep home adequately warm. We then changed the question in 2005 and the interviewee was then asked 'do you have financial difficulties to keep home warm?'.

Problem: in the French version, the question did not mention 'to keep home **adequately** warm', whereas the Dutch version did.

The answers in 2005 are thus barely comparable to those of 2004.

2004 :

N°	Question

	Pouvez-vous chauffer votre logement convenablement ?		
H 1			
	Oui Non		

2005 :

N°	Question	Codes	Routing	*EV*
H 11	**Avez-vous financièrement des difficultés pour chauffer votre logement ?**		H 12	
	Oui Non	1 2		*HH050*

2.3.2.2. Processing errors

Belgium used the CAPI–method to interview the persons. The questionnaire was programmed in Blaise. So processing errors due to data entry (from a written to an electronic format) were reduced to a minimum.
Statistics Belgium programmes several data entry and coding controls in the Blaise program. Those were identical for both waves.

Next to these controls, some warnings were implemented **in 2005** in order to ask the interviewer to verify the introduced data in the case of abnormally high or low amounts. A warning is a simple text box with a message such as 'This amount is very low, are you sure the amount is right?' or 'This amount is very high, are you sure the amount is right?'. The interviewer has then to confirm the value or to change it in case of error.

Household questionnaire	
H16	If lower than 500 or higher than 1000000
H22 (monthly)	If lower than 20 or higher than 2000
H22 (half-yearly)	If lower than 100 or higher than 10000
H22 (yearly)	If lower than 200 or higher than 20000
H23 (monthly)	If lower than 20 or higher than 2000
H23 (half-yearly)	If lower than 100 or higher than 10000
H23 (yearly)	If lower than 200 or higher than 20000
H26	If lower than 25 or higher than 5000
H33	If lower than 50 or higher than 10000
H34, H37, H41	If lower than 100 or higher than 5000
H43, H77, H84	If lower than 25 or higher than 1000
H66	If lower than 100 or higher than 25000
H71B	If lower than 25 or higher than 750
H79, H86	If lower than 300 or higher than 12000
H93	If lower than 100 or higher than 1500
Individual questionnaire	
I25, I27, I47, I50, I90, I91	If lower than 500 or higher than 5500

I53, I86, I93, I94	If lower than 6000 or higher than 66000
I58	If higher than 1200
I98B, I98C, I115B, I115C	If higher than 1350
I99, I102B, I102C	If higher than 5400

Some warnings concern other values than amounts. It's the case for H17 when the value is higher than 30 years ('A period of 30 years is really exceptional, are you sure it is right?') and for H18 when the interest equals 0 or is higher than 15.

2.3.3. Non-response errors

2.3.3.1. Achieved sample size

- number of households for which an interview is accepted in the **longitudinal database 2004-2007**:

2004	2005	2006	2007
1320	3150	4705	4066

- number of persons 16 years or older, number of sample persons and number of co-residents, members of households for which an interview is accepted in the **longitudinal database 2004-2007** and who completed a personal interview:

	2004	2005	2006	2007
Persons 16 y and more	2571	6128	9006	7894
Sample persons	2571	6079	8821	7586
Co-residents with interview		49	185	308

2.3.3.2. Unit non-response

Response rate for households

- Wave response rate

Wave response rate =

$$= \frac{5952}{9144 - 119} = 66\%$$

Refusal rate =

$$= \frac{1804}{9144 - 119} = 20\%$$

Non contacted and others rate =

$$= \frac{1044}{9144 - 119} = 11.6\%$$

- Longitudinal follow-up rate

Longitudinal follow - up rate $=$

$$= \frac{4242}{4242 + 429} = 91\%$$

- Follow-up ratio:

follow - up ratio $=$

$$= \frac{4242 + 2102}{4242 + 429} = 1.36$$

- Achieved sample size ratio

Achieved sample size ratio $=$

$$= \frac{5952}{4672} = 1.27$$

SAMPLE OUTCOME IN WAVE4

DB130=11										
DB135=1 (A)	DB135=2 (B)	DB120=22 (C)	DB130=22 (D)	DB130=23 (E)	DB130=24 (F)	DB130=21 (G)	DB120=21 (H)	NC (I)	DB110=10 (J)	DB120=23 (K)
5952	0	6	565	225	436	1804	35	2	0	119

SAMPLE OUTCOME IN WAVE3

DB130=11 DB135=1										
3872	0	0	131	102	137	427	0	2	0	1

DB135=2

0	0

DB120=21 to 23

DB130=21 to 24 TOTAL

NEW HOUSEHOLDS IN WAVE 4

DB110=8

54	0	0	7	5	10	22	1	NA	NA	0

DB110=9

2026	0	6	427	118	289	1355	34	NA	NA	118

Personal interview response rates

Response rate for persons
- Wave response rate

Wave response rate of sample persons =

$$= \frac{11437}{13094} = 87\%$$

- Wave response rate of non sample persons:

$$= \frac{166}{170} = 97\%$$

- Longitudinal follow-up rate:

$$= \frac{11437}{13094} = 87\%$$

Rate (RB250=21) $= \dfrac{35}{13094} = 0.3\%$

Rate (RB250=23) $= \dfrac{4}{11340} = 0.03\%$

Rate (RB250=31) $= \dfrac{19}{11340} = 0.15\%$

Rate (RB250=32) $= \dfrac{5}{11340} = 0.04\%$

Rate (RB250=33) $= \dfrac{1}{11340} = 0.01\%$

- Achieved sample size ratio for sample persons

$$= \frac{11437}{10983} = 104\,\%$$

- Achieved sample size ratio for sample and co-residents

$$= \frac{11751}{11227} = 104\,\%$$

- Response rate for non-sample persons

$$= \frac{314}{427 - 94} = 94.3\%$$

Personal interview response rate in wave 2

	RB250=11,12,13	Not completed because of									TOTAL
		RB250=21	RB250=22	RB250=23	RB250=31	RB250=32	RB250=33	HHnc	Pn	PI	

Sample persons (RB100=1 and rb245=1-3) from the sample forwarded from last wave

		RB250=11,12,13	RB250=21	RB250=22	RB250=23	RB250=31	RB250=32	RB250=33	HHnc	Pn	PI	TOTAL
(1) RB110=1-2		7583	31									7614
(2) RB110=6												10
(3) RB110=-1												0
(4) RB120=2												1
(5) RB120=3												6
(6) RB120=4												39
(7) DB135=2 or -1 or DB110=7 or DB120=21-23 or DB130=21-24 or -1												1555
(8) DB110=3-6												0
												9215
New sample persons												
(9) Reached age 16												0
(10) Sample additions		3854	4	0	4	19	5	0				3886
Non-sample persons 16+												
(11) this wave	From w1	166	0	0	1	3	0	0	79	1		250
	Not in w1	148	1	0	5	6	1	2	14	0		177
(12) Earlier wave	From w1											
	Not in w1											

Sample persons from sample not forwarded from last wave (excluded died or non eligible)	0									0
Sum of rows	11437	35	0	4	19	5	0	0	0	0	13094
	11437	35	0	4	19	5	0	0	0	0	13094
	11751	36	0	10	28	6	2	93	1	0	13521

2.3.3.3 Distribution of households by household status, by record of contact at address, by household questionnaire result, by household acceptance

Household status
DB110=

	Total	1	2	3	4	5	6	7	8	9	10
Total	5312	5010	201	0	0	1	1	0	99		0
%	100	94.3%	3.8%	0.0%	0.0%	0.0%	0.0%	0.0%	1.9%	0.0%	0.0%

Record of contact at address
DB120=

	Total	11	21	22	23
Total (DB110=2, 8,10)	300	297	1	0	2
%	100	99.0%	0.3%	0.0%	0.7%

Household questionnaire result
DB130=

	Total	11	21	22	23	24	missing
Total (DB120=11 or DB110=1)	5312	4066	596	230	143	272	5
%	100	76.5%	11.2%	4.3%	2.7%	5.1%	0.1%

Household interview acceptance
DB135=

	Total	1	2	missing
Total(DB130=11)	4066	4066	0	0
%	100	100	0	

2.3.3.4 Distribution of persons for membership status (RB110)

	Total	Current HH member				No current HH member		
		RB110=1	RB110=2	RB110=3	RB110=4	RB120=2 to 4	RB110=6	RB110=7
Total	10058	9685	61	220	35	47	10	0
%	100	96.3%	0.6%	2.2%	0.3%	0.5%	0.1%	0.0%

Distribution of persons moving out by variable RB120

	Total	RB110=5

		RB120=1		RB120=2	RB120=3	RB120=4
		This person is a current HH member	This person is not a current HH member			
Total	152	46	59	1	7	39
%	100	30.3%	38.8%	0.7%	4.6%	25.7%

2.3.3.5 Item non-response

In the following table an overview of the item non-response for all income variables is presented. The percentage households having received an amount, the percentage of households with missing values and the percentage of households with partial information is calculated.

These percentages are calculated as follows:
- % of households having received an amount : number of households (or persons) who have received something (yes to a filter) / total
- % of households with missing values : number of households (or persons) who said that they have received something but did not give any amount (no partial information) / number of households (or persons) who have received something (yes to a filter)
- % of households with partial information: number of households (or persons) who said that they have received something but gave partial information (amounts were not given for all components) / number of households (or persons) who have received something (yes to a filter)

Overview of the non-response for the income variables - % households having received an amount, % of households with missing values and % of households with partial information.

Item non-response	% of households having received an amount	% of households with missing values	% of households with partial information
Total gross household income (HY010)	100	10,3	57,2
Total disposable household income (HY020)	100	4,5	64,5
Total disposable household income before social transfers except old-age and survivor's benefits (HY022)	96,8	3,4	64,9
Total disposable household income before social transfers including old-age and survivor's benefit (HY023)	94,4	1,1	66,5
Net income components at household level			
Family related allowances (HY050N)	35,9	0,6	0,5
Interests, dividends, etc. (HY090N)	68,0	46,7	0
Gross income components at household level			
Income from rental of a property or land (HY040G)	7,3	0,6	0,0

Family related allowances (HY050G)	35,9	0,6	0,7
Social exclusion not elsewhere classified (HY060G)	1,8	0,0	
Housing allowance (HY070G)	0,7	0,2	
Regular inter-household cash transfer received (HY080G)	7,8	0,4	0,1
Interest repayments on mortgage (HY100G)	31,1	5,7	
Income received by people aged < 16 (HY110G)	0,2		
Regular inter-household cash transfer paid (HY130G)	7,2	0,2	0,0
Tax on income and social contributions (HY140G)	89,6	8,0	34,7
Net income components at personal level			
Employee cash or near cash income (PY010N)	47,7	3,8	5,8
Cash benefits or losses from self-employment (PY050N)	6,2	1,9	0,0
Pension from individual private plans (PY080N)	0,2	0,1	
Unemployment benefits (PY090N)	12,6	0,9	0,0
Old age benefits (PY100N)	18,2	1,1	0,1
Survivor' benefits (PY110N)	0,8	0,0	
Sickness benefits (PY120N)	1,7	0,2	
Disability benefits (PY130N)	3,3	0,1	
Gross income components at personal level			
Employee cash or near cash income (PY010G)	47,7	6,3	8,5
Non cash employee income (PY020G)	19,1	1,6	1,9
Company Car (PY021G)	3,6	0,9	0,0

Cash benefits or losses from self-employment (PY050G)	**6,2**	**4,4**	**0,0**
Pension from individual private plans (PY080G)	**0,2**	**0,1**	**0,0**
Unemployment benefits (PY090G)	**12,6**	**4,1**	**0,1**
Old age benefits (PY100G)	**18,2**	**6,3**	**0,2**
Survivor' benefits (PY110G)	**0,8**	**0,2**	**0,0**
Sickness benefits (PY120G)	**1,7**	**0,5**	**0,0**
Disability benefits (PY130G)	**3,3**	**0,8**	**0,0**
Education-related allowances (PY140G)	**1,8**	**0,0**	**0,0**

2.4 Mode of data collection

Distribution of household members aged 16 and over by RB250

(Household members RB245=1)

	Total	RB250=11	RB250=14	RB250=21	RB250=23	RB250=31	RB250=32	RB250=33
Total	7940	7894	45	0	0	0	1	0
%	100	99.4	0.6	0	0	0	0	0

(Sample persons 16+ RB245=1 and RB100=1)

	Total	RB250=11	RB250=14	RB250=21	RB250=23	RB250=31	RB250=32	RB250=33
Total	7617	7586	30	0	0	0	1	0
%	100	99.6	0.4	0	0	0	0	0

(Co-residents 16+ RB245=1 and RB100=2)

	Total	RB250=11	RB250=14	RB250=21	RB250=23	RB250=31	RB250=32	RB250=33
Total	323	308	15	0	0	5	1	0
%	100	95.4	4.6	0	0	0	0	0

Distribution of household members aged 16 and over by RB260

(Household members 16 + RB250=11)

	Total	RB260=2	RB260=5
Total	7940	5929	1033
%	100	85.2	14.8

(Sample persons 16 + RB100=1 and RB250=11)

	Total	RB260=2	RB260=5
Total	7617	5747	959
%	100	85.7	14.3

(Co-residents 16 + RB100=2 and RB250=11)

	Total	RB260=2	RB260=5
Total	323	182	74
%	100	71.1	28.9

2.5 Imputation procedure

2.5.0 Preceding important remark

In contrast to 2004 and as 2005 – in 2006 and 2007 the calendar question (i40 in the questionnaire) was presented to every respondent rather the only those who indicated that had been a change in their social-economic position. It enabled us to assess and check much thoroughly the link between the social-economic position and the income variables. Notably for the self-employed this resulted in a substantive number of cases (being identified as being self-employed) who would be otherwise (and who were to some extent in 2004) not identified as being self-employed. These cases mainly concern people in jobs 'somewhere on the bridge' between being self-employed and employee but who nevertheless indicated in the calendar that they were self-employed.

2.5.1 Overall strategy: Emphasis on internal information and integration of outlier detection- , imputation- and control-phases.

Overall strategy has not changed between 2006 and 2007. We refer the readers to the 2006 Quality rapport for details.

2.5.2 Description on imputation per target variable

In the following table is shown which imputation method we used for each target variable (and also for each component within the Belgian questionnaire). The percentage of imputed cases and the total number of observations is added.

Percentage of imputation over the total number of observations per (target) variable

Income Component		Question in the Belgian questionnaire		Percentage imputed cases	Method
Code	Description	Code	Description	(total number of observations)	
HY040	Income from rental of a property or land	H37	Rental of a part of the house	14.3 (28)	1) median
HY040	Income from rental of a property or land	H74	Rental of property or land other than own house	1.0 (481) 1.8 (481) [5.0 (481)]	1) Hot deck (imputation of a randomly drawn given amount) 2) imputation based on SILC 2006 [3) Median of predefined intervals (classes)]
HY040	Income from rental of a property or land			8.2 (481)	
HY050	Family/children related allowances	H91	Child allowance	0.6 (2265) 0.9 (2265)	1) Regression with number of children and age of the oldest child as auxiliary variables 2) SILC 2006 is source
HY050	Family/children related allowances	H93	Birth grant	4.1 (169)	1) Median of the given amounts (in classes based on number of children)

HY050	Family/children related allowances	(I116)	Income maintenance benefit in the event of childbirth	0 (121)	No imputations
HY050	Family/children related allowances	(I117)	Parental leave benefit	3.6 (84) 8.3 (84)	1) correction 2) imputation fixed amount
HY050	Family/children related allowances			3.6 (2276)	
HY060	Social assistance	H71A, H71B		0.9 (116)	
HY070	Housing allowance	H43	Allowance for housing (tenants)	9.1 (22)	1) Median
HY070	Housing allowance	H26	Intervention of authorities for repayments on mortgage	32 (25)	1) Median
HY070	Housing allowance			21.3 (47)	
HY080	Regular inter-household cash transfer received	H86	Alimony and child support received	1.7 (354)	1) Median
HY080	Regular inter-household cash transfer received	H88	Regular cash support	11.6 (189) 0.5 (189)	1) Hot deck 2) SILC 2006 is source
HY080	Regular inter-household cash transfer received			5.9 (493)	
HY090	Interests, dividends, etc.			68.7 (4314)	

HY110	Income received by people aged < 16	H69		0 (18)	No imputation
HY130	Regular inter-household cash transfer paid	H79	Alimony and child support paid	1.5 (273)	05: Median
HY130	Regular inter-household cash transfer paid	H81	Regular cash support	4.6 (219) 0.5 (219)	01: Hot deck 09: deductive imputation based on answer in 2008
HY130	Regular inter-household cash transfer paid			3.3 (460)	
HY140	Tax on income and social contributions	I130	Repayments for tax adjustment	4.6 (2032)	1) other source was used for control: fiscal data
HY140	Tax on income and social contributions	I132	Receipts for tax adjustment	4.6 (3649)	1) other source was used for control: fiscal data
HY140	Tax on income and social contributions			47.7 (5687)	Tax was computed as the sum of all differences between gross and net in income variables, corrected by tax adjustment. In case a gross-net model or a net-gross regression was used, the difference (tax) was considered as imputed.
PY010	Employee cash income – Gross income	I47-I48 (scale)	Monthly Wages and salaries	1.15 (5693) 12.17 (5693) 1.33 (5693) 2.93 (5693) 0.17 (5693)	1) Corrections 2) Net income is given, imputation based on regression 3) current income is given, imputation based on regression 4) Imputation on basis of EU-SILC 2006

					5) annual income is source
PY010	Employee cash income – Net income	I50-I51 (scale)	Monthly Wages and salaries	1.01 (5693)	1) Corrections
				2.17 (5693)	2) current income is given, imputation based on regression
				3.99 (5693)	3) Imputation on basis of EU-SILC 2006
				0.17 (5693)	
					4) annual income is source
PY010	Employee cash income	I52	Number of months I47-I48	0.1 (5693)	1) correction
PY010	Employee cash income	(i60_a_ne)	Pay for overtime	5.8 (190)	1) imputation based on SILC2006
PY010	Employee cash income	(i60_b_ne)	Commissions	8.2 (61)	1) imputation based on SILC2006
PY010	Employee cash income	(i60_c_ne)	Tips	4.3 (23)	1) imputation based on SILC2006
PY010	Employee cash income	(i60_d_ne)	Additional payments based on productivity	3.0 (100)	1) imputation based on SILC2006
PY010	Employee cash income	(i60_e_ne)	End of the year payments	4.2 (3887)	1) imputation based on SILC2006
				4,4 (3887)	2) regression – a.o. income as independent variable
PY010	Employee cash income	(i60_f_ne)	Thirteenth month payment	2.3 (653)	1) imputation based on SILC2006
				6.6 (653)	2) regression – a.o. income as independent variable
				0.2 (653)	3) correction
PY010	Employee cash income	(i60_g_ne)	Fourteenth month payment	2.1 (47)	1) imputation based on SILC2006
PY010	Employee cash income	(i60_h_ne)	Holiday payments	4.2 (4649)	1) imputation based on SILC2006
				3.6 (4649)	2) regression – a.o. income as independent variable
				0.2 (4649)	3) correction
PY010	Employee cash income	(i60_i_ne)	Profit sharing	0.8 (118)	1) imputation based on SILC2006
PY010	Employee cash income	(i60_j_ne)	Shares	5.7 (35)	1) imputation based on SILC2006

PY010	Employee cash income	(i60_k_ne)	Allowances for mobile-phone costs	2.2 (446)	1) imputation based on SILC2006
				0.22 (446)	2) correction
PY010	Employee cash income	(i60_l_ne)	Allowances for gas/electricity and dwelling related cost	11.11 (27)	1) imputation based on SILC2006
				48,15 (27)	2) median imputed
PY010	Employee cash income	(i60_m_ne)	Allowances car insurance	1.72 (116)	1) imputation based on SILC2006
				65.5 (116)	2) median imputed
PY010	Employee cash income	(i60_n_ne)	Allowances gasoline/petrol	0.7 (430)	1) imputation based on SILC2006
				26,6 (430)	2) median imputed
PY010	Employee cash income	(i60_o_ne)	Allowances paid for working in remote locations	2.5 (21.4)	1) median imputed
PY010	Employee cash income	(i60_p_ne)	Other additional payments	0.0 (179)	No imputation
PY010	Employee cash income	I53	Income from irregular jobs : wages and salaries	1.8 (221)	1) only gross value was recorded
				0.9 (221)	2) Imputation based on SILC 2006
PY010	Employee cash income	I93	Income from jobs other than main job : wages and salaries	13.2 (53)	1) imputation fixed amount
PY010G	Employee cash income			31.0 (5877)	Please consider high number of net-gross imputations (see variable I47 above)
PY010N	Employee cash income			20.0 (5877)	Please consider high number of imputations for which an alternative income was the source (see variable I50 above)
PY050	cash benefits or losses from self-employment	I93	Income for jobs other than main job : self-employed	31.00 (71)	1) imputation fixed amount
PY050G	cash benefits or losses from self-			71.2 (768)	Please take notice of the important remarks in 2.6.0

	employment				and 2.6.1 to assess the nature of the imputations for the self-employed.
PY050N	cash benefits or losses from self-employment			30.2 (768)	Please take notice of the important remarks in 2.6.0 and 2.6.1 to assess the nature of the imputations for the self-employed.
PY080	Pension from Individual private plans	I109	Savings for ones old day (Epargne-pension)	18 (11)	No imputations
PY080	Pension from Individual private plans	I112	Life insurance (Assurance-vie)	(11)	No imputations
PY090	Unemployment benefits	I98_a	Subsistence income for persons entering the labour market	3.4 (29)	1) legal amount was imputed
PY090	Unemployment benefits	i98_b	Full unemployment benefits	2.2 (1031)	1) SILC 2006 is source
				0.5 (1031)	2) indirect imputation via HH-income
				0.2 (1031)	3) correction
				2.1 (1031)	4) legal amount
PY090	Unemployment benefits	I98_c	Partial unemployment benefits	(129)	No imputations
PY090	Unemployment benefits	I98_d	Other financial assistance (Allocation de garantie de revenus)	(20)	No imputations
PY090	Unemployment benefits	(I98_e)	Other financial assistance (Allocation du fonds de sécurité d'existence)	10 (20)	1) Net income is given, imputation based on regression
PY090	Unemployment benefits	(I98_f)	Vocational training allowance	6.7 (15)	1) Net income is given, imputation based on regression

PY090	**Unemploym ent benefits**	**(I98_h)**	**Other cash benefits**	**4.2 (24)** **4.2 (24)**	**1) Net income is given, imputation based on regression** **2) imputation legal amount**
PY090	**Unemploym ent benefits**	**I99_b**	**Early retirement benefits**	**2.2 (275)** **1,8 (275)**	**1) current income is source** **2) SILC 2006 is source**
PY090	**Unemploym ent benefits**			**7,1 (1556)** **32.3 (1556)**	**1) imputations** **2) net income was given**
PY100	**Old age benefits**	**I104**	**Pension Fund (Fonds de pension)**	**(56)**	**No imputations**
PY100	**Old age benefits**	**I106**	**Group insurance (Assurance-groupe)**	**(15)**	**No imputations**
PY100	**Old age benefits**	**(I_102_B)**	**Old age pensions**	**4.2 (1973)** **0.8 (1973)** **0.3 (1973)** **0.1 (1973)**	**1) SILC-2006 is source** **2) current pension is source** **3) indirect imputation via HH-income** **4) correction**
PY100	**Old age benefits**	**(I_102_C)**	**Other financial assistance to old aged people[5]**	**(16)**	**No imputations**
PY100	**Old age benefits**	**(I_102_D)**	**Other financial assistance to old aged people[6]**	**(22)**	**No imputations**
PY100	**Old age benefits**	**(I_102_E)**	**Type of old age benefits not given**	**5.4 (37)**	**04: Net pension is source**
PY100	**Old age benefits**			**5.2 (2225)** **29.1 (2225)**	**1) imputed** **2) only net pension was given**

---

[5] Revenus garantis aux personnes âgées
[6] Complément au revenu garanti aux personnes âgées

PY110	Survivor's benefits[7]	(I102_A)		2.5 (365)	1) SILC 2006 is source
				[32,3 (365)]	[2) Net pension is given]
PY120	Sickness benefits	(I115_c)	Paid sick leave (temporary inability to work due to sickness)	0.8 (130) 1.5 (130) 0.8 (130)	1) current income is source 2) SILC 2006 is source 3) legal amount is imputed
PY120	Sickness benefits	(I115_d)	Paid sick leave (temporary inability to work due to professional sickness or injury)	5.9 (17)	1) correction
PY120	Sickness benefits	(I115_e)	Other sickness benefits	5.3 (38)	1) imputation of fixed amount
PY120	Sickness benefits			9.5 (212)	
PY130	Disability benefits	I115_a	Disability pension	0.3 (312) 0.6 (312)	1) current income is source 2) correction
PY130	Disability benefits	(I115_b)	Integration income for the handicapped	9.8 (61) 1.6 (61)	1) correction 2) Silc 2006 is source
PY130	Disability benefits			3.4 (411)	
PY140	Education-related allowances			3.2 (219)	Note that in the P-file all grants received by someone in the household are given to the reference person of the household as they can concern persons aged under 16 who are not present in the P-file.

2.5.3 Imputation of partial unit non-response

The method chosen for Belgium was imputation of an income for each member of the household who did not answer the questionnaire. Imputation is based on the variable

[7] Individuals could answer 'yes' to the filter of question I102_a and be more than 65 years. After imputation, the values of the benefits were classified as old-age benefits.

RB210 (basic activity status) of the individual given in the R-file. When the answer is missing or 4 (other inactive person), it is chosen not to impute any income. When available, we preferably used the longitudinal information's from 2006 for imputation. For the other cases the chosen method for imputation was imputation of a sub-category median based on age and sex. Net incomes were computed with a gross to net model, based on the imputed gross incomes.

2.6 Imputed rent

From 2007 onwards a measure for 'imputed rent' needs to add to the data.

Below we briefly explain the implementation of imputed rent (IR – hereafter) in the Belgian EU-SILC 2007 data. The text gives insight in the variables and methods used and in the results but is, overall, non-technical. For more in-depth technical background on the subject please turn to the appropriate documentation available via Eurostat (Doc. EU-SILC/162/06/EN).

In order to asses IR it was agreed on with Eurostat to use a (two-step) Heckman regression. The Heckman method involves in essence (A) the resolution of a probit regression model with tenure status of the household dwelling (dichotomy tenant/non-tenant) as dependent variable and conventional explanatory variables (Doc. EU-SILC/162/06/EN). (B) The coefficients found for the inverse of Mills ratio are then introduced in a regression model to counter selection bias in the estimated IR outcomes.

One difficulty in the first step is choosing the right variables. The Eurostat guidelines were closely followed for that purpose and also previous work on the subject of IR for the household budget survey was helpful. The following variables - or rather sets of variables - were selected:

- Characteristics and 'state' of the dwelling: type, number of rooms, presence of problems with the dwelling
- A number of neighborhood characteristics (with some emphasis on the presence of problems).
- Characteristics of the household: ages of the members of the household, their activity status, educational attainment, household type, number of children, number of persons in the household

One difficulty was that individual characteristics (age, activity status, educational attainment) needed to be aggregated on the household level. That was done by the creation of dummy variables for each category of the individual characteristics measuring the presence or the absence of that category on the level of the household. The table below gives an overview.

Not all variables originated from the SILC-database. Calculated for each municipality from the Belgian census 2001 — the distribution renters/owners was added to the equation.

Table: Overview of the variables in the analysis.

Label in output-files	Variable	Operationalisation/ measurement level
HH_INC_Q	Household income – HY020	quintiles
HT	householdtype	Categorical – see EUR.doc….
N_HH	Number of persons in the household	Metric
HH010	Dwelling type	Categorical – see EUR.doc.065
HH030	Number of rooms	Metric
HH050	Ability to keep dwelling warm	Categorical
HH080	Bath or shower	Categorical
HH090	Indoor flushing toilet	Categorical
HS160	Problems with dwelling	Categorical
HS170	Noise from neighbours	Categorical
HS180	Pollution	Categorical
HS190	Crime, violence or vandalism	Categorical
PERC_RENT	% HH renting in community of residence	Source census 2001
AGE_1	<18 yrs.	Dummy
AGE_2	>= 18 yrs. - < 25 yrs.	Dummy
AGE_3	>= 25 yrs. - < 45 yrs.	Dummy
AGE_4	>= 45 yrs. - < 65 yrs.	Dummy
AGE_5	>= 65 yrs.	Dummy
ACTSTA_1	Activity status – working	Dummy
ACTSTA_2	Activity status – unemployed	Dummy
ACTSTA_3	Activity status – retired	Dummy
ACTSTA_4	Activity status – non active	Dummy
EDUC_1	ISCED – 0 – 1	Dummy
EDUC_2	ISCED – 2	Dummy
EDUC_3	ISCED – 3 – 4	Dummy
EDUC_4	ISCED – 5 – 6	Dummy

EXPLORATORY ANALYSIS.

To get a first insight in the impact of each of the variables on the dependent variable tenure status (tenant/owner) a number of (mainly) bivariate logistic regressions were done.

Overall, the results show that the majority of the variables are associated with tenure status. All variables were therefore further kept in the analysis.

The explanatory analysis also resulted in the identification of a small number of missing values on some of the variables. Imputations were necessary to avoid distortion of further analysis.

The following imputations were done:

HH010 → 212 missing cases were coded as a separate category.
HH030 → 212 missing cases were given the median value (5)
HH031 → 17 missing cases were given the median value (1996)
HH040 → 1 missing case was given the value 1
HH050 → 5 missing cases were given the value 5
HS160 → 2 missing cases were given the value 2
HS180 → 3 missing cases were given the value 2
HS190 → 3 missing cases were given the value 2

PROBIT-REGRESSION.

The probit-regression part of the analysis was done in SAS. The output of this analysis is available on demand.

LINEAIR-REGRESSION.

The final estimation of IR is based on a linear regression model in which the observed rent for the renters is the dependent quantity and a number of dwelling-related characteristics are the independent variables.

An important note here is that, that dummy variables for the arrondissement of residence – variables ARR in the output – were introduced in the model. Arrondissements are (in fact) a (juridical – not political) administrative level between municipalities and provinces. We believe they are excellent indicators of regional differences and tendencies on scale smaller than provinces but bigger than municipalities.

The inverse-mills coefficient was significant at <0.001 level.

The output of the final regression is available on demand.

RESULTS.

The table below resumes our final results for the imputed rent variable:

Table: final results

	TENSTA			
	OWNERS (and others)		RENTERS	
	OBS.RENT	IR	OBS.RENT	IR
Mean	.	420.22	414.93	415.01
StdDev	.	119.02	200.76	95.78
Median	.	424.29	400.00	421.91
Min	.	0.00	50.00	1.19
Max	.	777.30	3700.00	761.31
P5	.	212.37	173.00	256.89
Q1	.	341.08	286.00	353.82
Q3	.	504.55	500.00	477.00
P95	.	605.29	715.00	562.14

The results are given on a monthly basis. Mean estimated imputed rent on basis of the fitted regression is 420 Euro for owners and slightly less, 415 Euro, for renters. Differences with the observed mean rent are quasi nil for the latter. Differences in the median estimated imputed rent are equally small.

2.7 Collection variable company Car

Since 2005, we decided to work with **the national rules of the tax authorities**. The benefit for individuals of using a company car for private goals was not directly assessed at the interview but afterwards calculated by applying the applicable taxation rules.

The fiscal benefit of all nature that a person has - due to disposition of a company car for private goals - is calculated by multiplying a fixed amount of kilometres driven for private use by a coefficient. To calculate the latest we need the fiscal cylinder capacity of the car. This fixed amount of kilometres driven for private use is for the tax authorities 5000 km if the distance home-work is less than 25 km, and 7500 if it's more than 25 km.

Since 2005, we asked directly the fiscal cylinder capacity and the distance between work and home. In case of non response of the cylinder capacity, we asked the mark, type and registration year of the car. Than we had to use an imputation method.

Imputation: To calculate the cylinder capacity, we did the following. We assumed that a company car is mostly diesel driven. We looked up for each mark, type and diesel engine what the corresponding cylinder capacity is. If we had several cylinder capacities for the type of the mark, we calculated the weighted mean of the cylinder capacity. If there is not diesel version for a type of car, we did the same logic but than for petrol.

Once we had that we could easily find the corresponding fiscal coefficient. Than we only had to multiply it by the fixed amount of kilometres driven for private use to obtain the fiscal benefit of all nature

Example:

Type of car	Fiscal cylinder capacity	Forfait	Distance home work	Fixed amount	Fiscal benefit of all nature
Smart fortwo	5	0,1864	< 25 km	5000	931 €
Smart fortwo	5	0,1864	> 25 km	7500	1396 €

After we calculated the fiscal benefit of all nature for a whole year, we weighted it for respondents who didn't dispose for a whole year of the company car. **The fiscal benefit of all nature is a gross non-cash employee income**.

3.Comparability

All household members of 16 year and older **at the time of the interview**, are selected for a personal interview. From 2006 on the age of 16 will be calculated at the end of the income reference period.

3.1 Basic concepts and definition

Only changes from first wave are reported.

Basic information on activity status during the income reference period
Basic information on activity status during the income reference period was mainly obtained via the calendar question (I40) in contrast to 2004 where it was obtained by combining the answer for question I8 (PL030) with the answer(s) for question(s) I38 (PL200) and for those with a change I40 (calendar question)). ALSO SEE REMARK 2.5.0.

3.2 Components of income

3.2.1 Differences between the national definitions and standard EU-SILC definitions, and an assessment, if available, of the consequences of the differences mentioned will be reported for the following target variables.

Total household gross income
HY010 = PY010 + PY020G + PY050G + PY090G + PY100G + PY110G + PY120G + PY130G + PY140G + HY040G + HY050G + HY060G + HY070G + HY080G + HY090G + HY110 G.

PY020G was not part of HY010 for 2004.
For 2005 and 2006 PY020G only contains the value of company cars.

Family/children related allowances

For the SILC 2004 Belgium asked allowances received from the federal government. From 2005 on it also includes birth grants given by some local authorities and medical organizations.

Income received by people aged under 16: in 2004 we asked the amount for last month (current) but the reference period for the variable is income reference period (year 2003). This was corrected for 2005 and the question aimed at the total income received last year by people aged fewer than 16.

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

No change from the previous wave.

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

No change from the previouswave.

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

See above for information on control, correction, imputation and creation of the gross target variables.

Tracing rules

Although the 'tracing rules' from Eurostat say that sample households non enumerated the first year of the panel 'may be dropped', some households who did not participate in 2004 were contacted in 2005. These cases concern households who were not interviewed in 2004 because they were temporarily away, unable to respond due to illness or due to other reason (DB130=22 to 24).

4. Coherence

The results of the Belgian EU-SILC2007 are in an acceptable way coherent with the results of previous waves. In depth studies to demonstrate this are currently in process at Statistics Belgian. Nevertheless however, these analysis (not the general analysis that is done during the validation phase at Eurostat) always has to be weighted against other priorities such as timelines f. ex.

Index