Working Group | Document 1

SILC 2003 - Statistics on Income and Living Conditions

Overview and first results

18 february 2004

Federal Public Service Economy, SMEs, Self-employed and Energy
Statistics and Economic Information
Enterprise number: 0314.595.348

.be

Working Group - Document 1
SILC 2003 – Statistics on Income and Living Conditions
Overview and first results

PART 1: OVERVIEW

PART 2: FIRST RESULTS

PART 1: OVERVIEW

I. EU-SILC: General issues

"EU-SILC" is the short term for "*European Union Statistics on Income and Living Conditions*". It is a *new statistic* at Statistics Belgium as from 2003.

♦ From PSBH/ECHP to EU-SILC

EU-SILC is the successor of the *Panel Study of Belgian Households.*[1] This was a panel survey conducted by the universities of Antwerp and Liège during more than ten years. Although the survey was done in a European context, there was no European Regulation that organized the survey. Countries were not obliged to organize the survey, so not all of the Member States did.
At European level it was decided to stop the panel and to launch a new project. Various reasons grounded this decision, for example: the sample of households that had become smaller with the years, the change in priorities in the domain of social policy that leaded to changes of the content of the survey.
For Belgium it was decided that Statistics Belgium would play a key role in the organization of the new project.
EU-SILC is organized in all EU Member States and is being coordinated by EUROSTAT.

♦ The importance of EU-SILC

On the one hand the SILC will be of interest for scientific research, on the other statistical information on income and living conditions is necessary for social policy makers. One of the goals is to make indicators in the field of poverty and social exclusion. We believe that different researchers at universities, the public sector in Belgium and the European Union will welcome the results of this survey.

[1] For more information we refer to the final report of the research project "Sociaal-economisch en demografisch panel": Mortelmans, D., Casman M.T., Doutrelepont, R. (ed.), 2004, *Elf jaar uit het leven in België.Socio-economische analyses op het gezinsdemografisch panel PSBH*, Gent, Academia Press, 615p.

II. EU-SILC 2003/ 2004-2007 in a European context

♦ EU-SILC 2003

6 Member States (Belgium, Denmark, Greece, Ireland, Luxemburg, Austria) and Norway organized EU-SILC in 2003 on the basis of a gentlemen's agreement.

♦ EU-SILC 2004

From 2004 on EU-SILC is based on a European Regulation (see annex 1):
Regulation (EU) No 1177 of the European Parliament and of the Council of 16 June 2003 concerning Community Statistics on income and living conditions (EU-SILC)

- *Aim* = to establish a common framework for the systematic production of a statistic that is to become the reference source of comparative statistics on income distribution and social exclusion at the European Union level;
- *Data* = cross-sectional and longitudinal (first and clear priority is the delivery of timely and comparable cross-sectional data);
- *Variables* = the primary domains are described in the European Regulation and the target variables in a related commission regulation;
- *Sample* = for each country a minimal effective sample size is fixed in the Regulation.

III. EU-SILC 2003 in Belgium

♦ Preparation of the survey

The main part of the preparatory work took place from January 2003 until August 2003 and consisted of:

- developing questionnaires, manuals and other documentation,
- programmation in Blaise and testing of the developments,
- delivery and installation of the laptops,
- selection of interviewers,
- sampling,
- creation and sending of letters to households,
- management of the 'Before Survey' program in Blaise, etc.

The choice of variables, sampling design and certain other preparatory work is done in co-operation with the general EU-SILC working group of the Higher Council Statistics in Belgium (chairman Eric Marlier and co-chairwoman Annie Versonnen).

Questionnaire
The questionnaire of the SILC2003 survey is the result of several steps. A pilot survey was held in 2002. One of the main goals of the pilot survey was to test the questionnaire in the field.
For building up the questionnaire we took the blue print questionnaire of Eurostat as the basis (documents SILC055 and SILC065). The order of the questions and the groups (themes of) questions is taken from this blue print. A major part of the

questions are quite literally taken over (and translated), but several other questions are changed because experiences in Belgium gave better results posing the questions in another way (the questionnaire is developed in collaboration with the universities that have the experience of the ECHP/PSBH project in Belgium).
After the pilot survey an evaluation of the questionnaire was done (detection of the problematic or difficult to answer questions). It was based on the comments of the interviewers and on a study of the item non-response. When building up the SILC2003 questionnaire we took account of this evaluation.

Sampling design
Type of sampling:

- The Belgian EU-SILC 2003 survey is a 2-stage sampling.
- There is stratification of sampling units.
- There is no clustering of sampling units.

Sampling units (two stages):

- Primary units:
  480 letters[2], drawn with probability proportional to size and possible repetitions, after regrouping the smallest, and ordering according to equivalised fiscal income.
- Secondary units:
  13675 private households, drawn within each primary units: 32 (64 if the unit had been selected twice, etc) for the September phase, between 17 and 37 for the November phase according to the response rate in September in each NUTS2 region (province).

Stratification and sub-stratification criteria:

- The stratification criterion is the region (NUTS2level). The strata are the 10 provinces of Belgium and the Brussels Capital Region.

Sample size and allocation criteria:

- The sample consisted of 13675 households. The allocation was designed – assuming different response rates in the strata- to yield about 6000 responding households; a first allocation was made for the three NUTS1-regions; followed, within the two larger (BE2 and BE3) regions, by a second allocation to their five NUTS2-regions. The basic rule was square root allocation (of the expected number of respondents), with a slight (Neyman-like) correction based on the number of households claiming income support and the dispersion of fiscal income, taken as an approximation of disposable income.

[2] In Belgium, statistical letters are parts of municipalities; there are some 2500 of them (average population 4000 inhabitants/1800 households, average area 12 sq. km)

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

NUTS2	Name	Drawn	Accepted (DB135=1)
BE10	Brussels	2704	1003
BE21	Antwerpen	1914	720
BE22	Limburg	804	367
BE23	Oost -Vlaanderen	1536	697
BE24	Vlaams-Brabant	1311	463
BE25	West-Vlaanderen	1083	669
BE31	Brabant Wallon	500	278
BE32	Hainaut	1664	752
BE33	Liège	1217	718
BE34	Luxembourg	371	199
BE35	Namur	571	333
Total	**Belgium**	**13675**	**6199**

Sample selection schemes:

- Systematic sampling of secondary units in each primary unit selected, the households have been ordered according to the age of the reference person.

Sample distribution over time:

- The sample was distributed over time. The fieldwork for the first phase took place in September and the fieldwork for the second phase took place in November. The reason of this is the calculation of the expected answer rate. As in 2003, it was the first time that the EU-SILC survey was organized at a large scale; the answer rate was difficult to predict. We decided to organize a first phase in September to have a better estimate of the answer rate. Based on this new estimate, we could do a better sampling for the second phase.
- In September 32 households were drawn in each of 240 primary units (half of the total number of primary units selected). With an expected answer rate of 40%, 3072 should have been interviewed. After the first wave, this number turned out to be slightly higher (especially in the Walloon/BE3 areas): 3350 households had been successfully interviewed in September. The average answer rate in September was 43,6%. For the sampling of the November (second phase) households, this answer rate was taken into account.

Interviewer selection

Interviewers were selected from the interviewer database that Statistics Belgium has centralised for all the survey's that are carried out by the institute. For each interviewer a basic curriculum vitae is present in the database (mentioning for example for which surveys they have experience, their language knowledge, their knowledge of pc, ...). A specific unit at Statistics Belgium is occupied with the selection of the interviewers for each survey; they have good contact with and knowledge of the interviewers. They try to find the best interviewer for each of the geographical areas to cover for SILC. This is not always an easy task because for certain geographical areas several interviewers are candidate, but for other geographical unit there are few or no candidates. Note that interviewers in Belgium most often carry out this work as a second or casual occupation.

Since the SILC-survey succeeded the PSBH, which was carried out by other institutes (Universities of Antwerp and Liège) than Statistics Belgium, we also asked both Universities to inform the former PSBH-interviewers on the possibility to become an

interviewer for the new SILC-survey. Several interviewers were also recruited through that way.

Interviewer training
All the interviewers followed a two days training course before starting the fieldwork. A training group consisted of minimum 5 to maximum 20 interviewers, and according to the size of the training group there were 1 or 2 trainers.
Even though the accent was given to the practical side of the training (getting to know the questions and mastering the CAPI-program by imitating interview situations), three manuals were distributed and explained during the training:
A general manual ('Manuel general aux enquêteurs') containing information about the objectives of the survey, the organisation of the survey, legal and administrative aspects around the survey, fieldwork aspect (how to contact the household, how to introduce oneself, who answers which questions, time delays, ...) and the content of the questionnaires.
A second manual ('Manuel contenu') containing all kinds of additional explanations and examples for certain questions/answers.
A third manual ('Manuel CAPI') about the use of the portable PC for the SILC Computer Assisted Personal Interviews and about the data entry program itself.
The first day of the training there was half a day for learning about and discussing the first two manuals. In the afternoon the trainees received their laptop and got to know the survey and the tool to carry out the interview in practice. One test-interview was simulated collectively. The second day of the training a small part of the time was dedicated at testing to send the data electronically after carrying out the interview. All the rest of the day interviewers practiced several interviews and interview situations with each other on the basis of household profiles that were given. There was also a lot of time for questions and discussions in between these test-interviews.
At the end of the training sessions the instructors had a good image on the degree in which each interviewer ameliorated during the training and on the degree in which they mastered the work. For certain interviewers two days of training was more than enough to master the work, for others it was necessary that they practiced some more at home on specific aspects of carrying out this survey (for example using of the CAPI-program itself, working on the content of the survey, ...). They were recommended to do so before carrying out their first real interview. They were often also recommended to start interviewing one-person households.

♦ Characteristics of the fieldwork

During the fieldwork not only the interviews took place, but several coordinating tasks were done: responding to the questions of the interviewers, the households (free number), sending of data electronically to the SILC-server, the management of the 'After Survey' program in Blaise, etc.

Fieldwork period
The fieldwork for the first phase took place in September and the fieldwork for the second phase took place in November (see description 'Sampling Design, Sample distribution over time').

Number of households by interviewer
The secondary units consisted of 32 households for the September phase; for the November phase this number varied between 17 and 37 according to the response rate in September in each NUTS2 region. Most of the interviewers had one group of households in the September phase and one group of households in the November phase. Nevertheless several interviewers also had more groups (for example two in September and 1 in November) or only one group:

interviewers with 1 groups: 57
interviewers with 2 groups: 140
interviewers with 3 groups: 25
interviewers with 4 groups: 5

CAPI-interview
CAPI (*Computer Assisted Personal Interview*) are being carried out. This method consists in gathering the information from the households by introducing it directly into an electronic format. It has several advantages such as a smoother fieldwork (automatic routing) with a shorter interview duration, the use of automatic controls when entering the data which gives better data quality, results are sooner available, etc. (see annex 3 for more information on this method).

Compensation for the households
The households participate at the survey on a voluntary basis and receive an allocation of 30 euro in compensation of the time and effort that they have given.

♦ Data processing

The main part of the work on data processing took place from December 2003 until December 2004 (Processing survey data, Evaluation of interviewers, Payments of interviewers and households).

The steps in processing of the collected information are the following (a more detailed description can be found annex 3):
A *first phase* consisted of transforming the Blaise dataset into a usable SPSS dataset. Further, for a few variables correction and/or coding was necessary and outliers had to be checked.
This phase was followed by a *second phase* of extensive work on *imputation* for missing, inconsistent or incorrect *income data*.
Also, certain *other survey data* (such as company car) needed extra processing work.
Further, weighting factors were calculated and the *Eurostat target variables* were created.
After running a *final checking program* of Eurostat and carrying out the necessary corrections, the *Laeken indicators* were produced and a quality report was created.

♦ Data distribution
As from 2005 work on data distribution starts. For this topic we refer to the document on "Variable list and questionnaires" (Document 2 of the meeting of the Working Group SILC of 18/02/05).

Annex 1. European Regulations concerning EU-SILC

- **Regulation** (EU) No 1177 of the European Parliament and of the Council of 16 June 2003 concerning **Community Statistics on income and living conditions** (EU-SILC);
- **Commission Regulation** (EC) No 1980/2003 of 21 October 2003 implementing Regulation (EC) no 1177/2003 of the European Parliament and of the Council concerning Community Statistics on income and living conditions (EU-SILC) as regards **definitions and updated definitions**;
- **Commission Regulation** (EC) No 1981/2003 of 21 October 2003 implementing Regulation (EC) no 1177/2003 of the European Parliament and of the Council concerning Community Statistics on income and living conditions (EU-SILC) as regards **the fieldwork aspects and the imputation procedures**;
- **Commission Regulation** (EC) No 1982/2003 of 21 October 2003 implementing Regulation (EC) no 1177/2003 of the European Parliament and of the Council concerning Community Statistics on income and living conditions (EU-SILC) as regards **the sampling and tracing rules**;
- **Commission Regulation** (EC) No 1983/2003 of 21 October 2003 implementing Regulation (EC) no 1177/2003 of the European Parliament and of the Council concerning Community Statistics on income and living conditions (EU-SILC) as regards **the list of target primary variables**;
- **Commission Regulation** (EC) No 28/2004 of 5 January 2004 implementing Regulation (EC) no 1177/2003 of the European Parliament and of the Council concerning Community Statistics on income and living conditions (EU-SILC) as regards the detailed content of intermediate and final **quality reports**;
- **Commission Regulation** (EC) No 16/2004 of 6 January 2004 implementing Regulation (EC) no 1177/2003 of the European Parliament and of the Council concerning Community Statistics on income and living conditions (EU-SILC) as regards **the list of target secondary variables relating to the 'intergenerational transmission of poverty'**(Module 2005);
- **Commission Regulation** (EC) No 13/2005 of 6 January 2005 implementing Regulation (EC) no 1177/2003 of the European Parliament and of the Council concerning Community Statistics on income and living conditions (EU-SILC) as regards **the list of target secondary variables relating to 'social participation'** (Module 2006);

Annex 2. The CAPI-method [3]

De enquêteurs gaan op bezoek bij de huishoudens met een draagbare computer. Meer en meer landen nemen toevlucht tot CAPI (*computer assisted personal interview*) voor het verzamelen van gegevens. Deze methode bestaat er uit om de informatie bij de huishoudens te verzamelen door ze rechtstreeks in een elektronisch formaat in te voeren.

De methode heeft namelijk verschillende voordelen:

1) De enquêteur moet de respondenten rechtstreeks interviewen (en bijgevolg begeleiden doorheen de vragenlijst). Het is niet meer mogelijk om een papieren formulier bij de respondent achter te laten zodat die alles dan zelf maar moet invullen. Deze aanpak is veel *gebruiksvriendelijker* tegenover de huishoudens, zeker en vast bij een enquête zoals SILC waar de vragenlijst toch niet eenvoudig is (bv. In termen van het begrijpen van de verschillende inkomensbegrippen, het doorverwijzen naar volgende vragen).

2) Een enquêteur met enige CAPI ervaring en enige vertrouwdheid met de enquête zal minder tijd nodig hebben dan wanneer dezelfde enquête via PAPI « *paper assisted personal interview* » zou gebeuren. Het leek ons dan ook nuttig om een opleiding met heel wat praktijktoepassingen te organiseren. Deze praktijk vormt een basiservaring zodat de enquêteur het CAPI systeem snel onder de knie heeft.

3) CAPI verzekert een betere kwaliteit van de resultaten dankzij het rechtstreeks invoeren van de antwoorden (we beperken de mogelijke fouten bij invoeren en coderen van een PAPI-enquête); ook het werk van de enquêteur is op bepaalde punten gemakkelijker:

- zo gebeurt de *routing* (= doorverwijzing naar de volgende vraag) automatisch en verschijnen er als geheugensteuntje voor de enquêteur hier en daar automatische syntheses van wat enkele minuten daarvoor werd ingevoerd,
- zo kunnen gegevens ook op voorhand in de vragenlijst worden ingevoerd zodat de enquêteur die enkel hoeft na te kijken en bevestigen (dit wordt des te interessanter in het tweede jaar van de panelenquête omdat de enquêteur dan het tweede jaar bepaalde informatie niet opnieuw hoeft in te voeren maar enkel te bevestigen of aan te passen),
- zo kunnen er ook automatische controles gebeuren bij de gegevensinvoer wat de kans op inconsistente antwoorden vermindert. Enkele voorbeelden: controle op onmogelijke/onlogische data (aankoop woning, leeftijd eerste werk, leeftijd toekenning studiebeurzen, verschillen in leeftijden tussen familieleden), bruto inkomens moeten groter of gelijk zijn aan netto inkomens, enz.

4) De resultaten zijn in principe sneller beschikbaar (geen fase van gegevensinvoer en quasi geen coderingen nodig) en bijgevolg zijn de statistieken meer accuraat voor de

[3] Text in Dutch

wetenschappers die ze interpreteren en voor de politieke verantwoordelijken die er zich op baseren om het beleid te evalueren en bij te sturen.

De Belgische huishoudens zijn waarschijnlijk (nog) niet vertrouwd met een enquêteur van de Algemene Directie Statistiek en Economische Informatie die binnenstapt met een draagbare pc om een interview af te nemen. Toch hopen we dat een vlot gebruik ervan door onze enquêteurs bij hen ook een positieve indruk zal achterlaten. We zien het als een investering voor de Algemene Directie Statistiek en Economische Informatie en voor de enquêteur op korte termijn en een waardevolle bijdrage op lange termijn.

Annex 3. Description of the steps in processing of the collected information [4]

I. Première phase de traitement (édition)

Envoi des résultats des différents groupes par ligne téléphonique
Une fois rentré chez lui, l'enquêteur doit envoyer ses données au service SILC par l'intermédiaire d'une ligne téléphonique. L'enquêteur connecte son ordinateur portable au serveur de la DG Statistique et Information Economique et envoie ses données. Les instructions sont d'envoyer les données le plus souvent possible et en tous cas, après la toute première interview pour faire un contrôle de la procédure et en fin de période d'enquête, pour être sûr que toutes ses données ont bien été enregistrées (un enregistrement sur disquette est également recommandé).

Programme 'after survey': création dans Blaise d'une grande base de données qui comprend les interviews complétées et non complétées
Lors de la période d'envoi par les enquêteurs, les statisticiens de l'équipe SILC utilisaient journellement le programme 'after survey', décompressant les fichiers reçus par le serveur et permettant d'avoir accès aux données transmises. Cette procédure permet un premier contrôle des enquêteurs (questions mal comprises, problèmes lors de l'envoi, professions peu explicites quand aucun code ISCO n'a été encodé, ...). Ces contrôles permettent certaines corrections à chaud.

Encodage de certaines variables en Blaise
Certaines informations ont pu être ajoutées immédiatement dans Blaise. Il s'agissait d'informations complémentaires apportées par les enquêteurs. Ce fut le cas principalement pour quelques interviews réalisées par proxy quand par exemple, une personne ne disposait pas de toute l'information lors de l'interview mais l'a fournie à l'enquêteur par la suite.
Pour les questions concernant la profession de l'individu ainsi que l'activité principale de l'entreprise ou de l'organisme qui l'emploie, l'enquêteur doit rentrer un code ISCO ou NACE respectivement. S'il ne trouve pas le code, il peut entrer directement la profession ou le domaine d'activité. C'est l'équipe SILC qui est alors chargée de retrouver les codes, sur base de l'information entrée par l'enquêteur.

Transfert en SPSS
Les données sont éditées dans le logiciel SPSS. Le transfert s'est fait après préparation d'un fichier de syntaxe SPSS. Celui-ci avait pour but, outre l'importation des données, l'attribution de nouveaux noms aux variables car Blaise attribue des noms par défaut, qui n'ont que peu de sens au point de vue de l'enquête, la préparation de certaines variables, notamment les dates qui devaient être formatées, l'attribution d'un label à chaque variable, pour expliciter la question à laquelle correspondait la variable, ainsi que des labels liés aux différents codes utilisés pour chaque variable (exemple : 1 veut dire 'oui, 2 veut dire 'non') et enfin, le codage des 'je ne sais pas' et 'je refuse' en tant que valeurs manquantes.

[4] Text partly in French (§ I, III, IV, V, VI) and partly in English (§II).

Les trois étapes suivantes font partie de ce qu'on nomme l'édition

L'édition des données est l'application de contrôles qui identifient les données manquantes, inconsistantes, non-valides ou qui indiquent les enregistrements qui sont potentiellement erronés. Certains de ces contrôles se fondent sur les relations logiques attendues en fonction des données recueillies (définitions). D'autres seront par nature plus empiriques ou sont obtenus après utilisation de procédures statistiques de détection des 'outliers'. Des données antérieures sur le même sujet ou des données provenant d'autres sources sont également très utiles dans le cadre du contrôle. Trois buts sont à retenir pour l'édition : servir de base pour améliorer l'enquête elle-même en vue de l'année prochaine, donner de l'information sur la qualité des données, et enfin nettoyer les données et préparer les données pour les opérations futures comme l'imputation.

✓ Première vue des résultats: **nettoyage et correction**

Dans le programme Blaise, il était prévu que les enquêteurs puissent faire des remarques générales en rapport avec le ménage interrogé, ou des remarques particulières associées à une question spécifique. Toutes ces remarques sont très importantes dans une optique d'édition des données. Elles ont notamment permis aux enquêteurs d'apporter ou de nuancer certaines informations concernant des montants.

✓ **Outliers** : détection et élimination

Un des grands problèmes dans la détection des outliers est qu'un grand nombre d'entre eux sont des zéros. Certains zéros sont des non-réponses (la personne n'a pas reçu l'allocation, l'enquêteur a encodé zéro comme non-réponse (refus ou ne sait pas) pour pouvoir continuer dans le programme) et certains sont de vrais zéros (la personne devait recevoir l'allocation, mais pour des raisons particulières, elle n'a pas été payée lors de la période de référence). Dans les remarques, nous avons régulièrement des indications sur le fait qu'un montant encodé peut paraître aberrant (très grand ou égal à zéro) mais que cela est justifié. D'autres montants aberrants, eux, sont peut-être dus à des erreurs d'encodage ou encore exprimés en francs belges. La méthode choisie pour repérer les outliers est celle du boxplot. Les valeurs qui dépassent le maximum calculé comme $x_{Q3}+1.5(x_{Q3}-x_{Q1})$ ou sont inférieures à un minimum calculé comme $x_{Q1}-1.5(x_{Q3}-x_{Q1})$ sont examinées au vu de la logique et de quelques variables bien choisies, qui peuvent rendre compte de la situation du ménage. C'est le cas de la variable 'total des revenus disponibles du ménage'. Les outliers flagrants seront éliminés de la base de données. D'autres laissant place aux doutes feront l'objet d'une discussion et d'un choix au cas par cas. Si le doute persiste, le choix sera généralement de conserver les données, en laissant en quelque sorte à l'enquêteur le bénéfice du doute.

✓ **Incohérences**

Afin d'obtenir les fichiers prêts pour imputation, il est nécessaire de s'interroger sur toute une série d'incohérences mises en lumière par quelques contrôles logiques. Ces contrôles, principalement réalisés sur les données individuelles, se basent sur le calendrier des activités de l'individu. Ce calendrier est rempli dès que la personne a répondu 'oui' à la question 'Nous sommes intéressés par les éventuels changements de votre situation depuis 2002. Depuis le début de l'année 2002, votre situation a-t-elle changé ?'. On demande alors pour chaque mois de l'année 2002 de donner l'activité principale (salarié, indépendant, étudiant, chômeur, ...) ainsi que le type de revenu reçu à cette occasion (salaire, allocations de chômage, aucun revenu, ...).

Un autre type d'incohérence qui fait l'objet d'une analyse également est le fait pour une même personne de répondre avoir reçu en 2002 des allocations de remplacement et des salaires pendant une période supérieure à 12 mois.
Après ces traitements, les fichiers de données sont prêts pour l'imputation.

II. Deuxième phase de traitement (imputation)

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.

Table 2: Percentage of imputation over the total number of observations per (target) variable

Income Component		Question in the Belgian questionnaire		Percentage imputed cases (total number of observations)	Method
Code	Description	Code	Description		
HY040	Income from rental of a property or land	H28	Rental of a part of the house	14.7 (34)	Median of the given amounts
HY040	Income from rental of a property or land	H59,H60	Rental of property or land other than own house	11.8 (450) 2.9 (450)	Hot deck (imputation of a randomly drawn given amount) (for the intervals, middle of the class)
HY040	Income from rental of a property or land			14.8 (472)	
HY050	Family/children related allowances	H70_A	Child allowance	2.5 (2196)	Regression with number of children and age of the oldest child as auxiliary variables
HY050	Family/children related allowances	H72	Birth grant	11.2 (169)	Median of the given amounts
HY050	Family/children related allowances	(I88)	Income maintenance benefit in the event of childbirth	22.2 (126) 11.1 (126) 1.6 (126)	1) Imputation IVE: auxiliary variable salary of last year (2002) or salary of last month and no change since 2002 2) If not possible: hot deck based on the given amounts 3) Two cases for which the net amount was given and gross computed with a regression net-gross

HY050	Family/children related allowances	(I88)	Parental leave benefit	38.5 (39)	Imputation of legal amounts
HY050	Family/children related allowances			5.8 (2223)	
HY060	Social assistance	H56A, H56B		12.6 (159)	Hot deck
HY070	Housing allowance	H33, H34		15.2 (46)	Hot deck
HY080	Regular inter-household cash transfer received	H66_B	Alimony and child support received	1.9 (265)	Median
HY080	Regular inter-household cash transfer received	H68_A	Regular cash support	8.6 (174)	Hot deck
HY080	Regular inter-household cash transfer received			4.5 (420)	
HY090	Interests, dividends, etc.	H79, H80		48.1 (1013)	IVE (auxiliary variables: sort assets (bank accounts, bonds,…), tenure status, …) Ranges of values: IVE with bounds
HY110	Income received by people aged < 16	H54		20.7 (29)	Median
HY130	Regular inter-household cash transfer paid	H62_B	Alimony and child support paid	1.1 (275)	Median
HY130	Regular inter-household cash transfer paid	H64_A	Regular cash support	8.7 (309)	Hot deck
HY130	Regular inter-household cash transfer paid			9 (561)	

HY140	Tax on income and social contributions	I94	Repayments for tax adjustment	13.3 (2940)	IVE (auxiliary variables: number of children, yearly gross and net wages in 2002, current economic status, status in employment, more than one job or not, interests on mortgage, limits as log(minimum of the given repayment values) and log (maximum of the given values)
HY140	Tax on income and social contributions	I96	Receipts for tax adjustment	8.4 (3168)	IVE (auxiliary variables: number of children, yearly gross and net wages in 2002, current economic status, status in employment, more than one job or not, interests on mortgage, limits as log(min of the given receipts values) and log (max of the given values)
HY140	Tax on income and social contributions			50 (5555)	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	I35-A, I35-C	Wages salaries	3.7 (5788) 11.6 (5788) 5.8 (5788)	1) Net is given, imputation based on a net-gross regression 2) Imputation of last month wages (if no change since 2002) 3) If not possible, IVE (auxiliary variables: age, gender, ISCED level, PL030, PL040, PL070, PL072, PL110, PL130, PL140, PL150, PL160, last month wages, …)
PY010	Employee cash income	(i42a_a_ne)	Pay for overtime	56.2 (395)	IVE (auxiliary variables: age, gender, ISCED level, monthly and yearly wages in 2002 (gross and net), current economic status, status in employment, NACE)
PY010	Employee cash income	(i42_b_ne)	Commissions	34.3 (67)	Median
PY010	Employee cash income	(i42_c_ne)	Tips	42.4 (59)	Median
PY010	Employee cash income	(i42_d_ne)	Additional payments based on productivity	29.4 (119)	Median

PY010	Employee cash income	(i42_e_ne)	End of the year payments	13.4 (3314)	IVE (auxiliary variables: same as pay for overtime)
PY010	Employee cash income	(i42_f_ne)	Thirteenth month payment	14.6 (779)	IVE (auxiliary variables: same as pay for overtime)
PY010	Employee cash income	(i42_g_ne)	Fourteenth month payment	9.9 (61)	Median
PY010	Employee cash income	(i42_h_ne)	Holiday payments	11.9 (4362)	IVE (auxiliary variables: same as pay for overtime)
PY010	Employee cash income	(i42_i_ne)	Profit sharing	35.3 (133)	Median according to imputation classes based on PL040
PY010	Employee cash income	(i42_j_ne)	Shares	48.2 (56)	Median according to imputation classes based on PL040
PY010	Employee cash income	(i42_k_ne)	Allowances paid for working in remote locations	41.5 (82)	Median
PY010	Employee cash income	(i42_l_ne)	Other additional payments	12.8 (149)	Median
PY010	Employee cash income	I67_b, I67_c	Income for jobs other than main job : wages and salaries	45 (67)	1) monthly income is given and number of months in 2002 2) net income is given and regression net-gross 3) median according to status in employment for second job
PY010	Employee cash income			43 (5785)	
PY050	cash benefits or losses from self-employment	I67_b, I67_c	Income for jobs other than main job : self-employed	45 (33)	1) monthly income is given and number of months in 2002 2) net income is given and regression net-gross 3) median according to status in employment for the second job

PY050	cash benefits or losses from self-employment	I57, (I64_b)	Income for main job : self-employed	79.8 (784)	1) only one variable could be filled in: gross or net. First, application of gross-net model or regression net-gross 2)Subjective value given for question I64_b was used 3)IVE software with auxiliary variables: gender, age (in classes), status in employment (self-employed with or without employees), I51 (work together with other member of the household), I49 (activities in industry, agriculture, private practices), I54 (own business with or without associates)
PY050	Cash benefits or losses from self-employment			79.8 (784)	
PY080	Pension from Individual private plans	I81	Pension Fund (Fonds de pension)	9.7 (72)	Median in imputation classes based on former status in employment
PY080	Pension from Individual private plans	I78	Group insurance (Assurance-groupe)	22.2 (18)	Median
PY080	Pension from Individual private plans	I82	Savings for ones old day (Epargne-pension)	30 (30)	Hot deck
PY080	Pension from Individual private plans	I85	Life insurance (Assurance-vie)	12.5 (8)	One case : median
PY080	Pension from Individual private plans			13.7 (117)	
PY090	Unemployment benefits	i72_a	Subsistence income for persons entering the labour market	28.6 (35)	1) Number of months : calendar (1 case) 2) Values : median imputed for gross, gross-net model

PY090	Unemployment benefits	(i72_b)	Full unemployment benefits	16 (1129)	1) number of months : calendar (4 cases), random drawn according to the frequency table of the given values (26 cases) 2) regression with as independent variables : recoded i3 and situation of the person in the household (one-person household, partner has income or not, dependent children or not, ….)
PY090	Unemployment benefits	I72_c	Partial unemployment benefits	15.9 (132)	1) number of months : calendar (2 cases), median (1) 2) Values : gross imputed : median
PY090	Unemployment benefits	I72_d	Other financial assistance (Allocation de garantie de revenus)	18.2 (22)	1) number of months : see above 2) Values : gross imputed : median
PY090	Unemployment benefits	(I72_e)	Other financial assistance (Allocation du fonds de sécurité d'existence)	37.9 (58)	1) number of months : random drawn according to the frequency table of the given values (16 cases) 2) Values : gross imputed : median
PY090	Unemployment benefits	(I72_f)	Vocational training allowance	25 (24)	1) number of months : median (2 cases) 2) Values : gross imputed : median
PY090	Unemployment benefits	(I72_g)	Allowance for child care/ mobility payments	6.5 (46)	1) number of months : median (1 case) 2) Values : gross imputed : median
PY090	Unemployment benefits	(I72_h)	Other cash benefits	21.4 (28)	1) number of months : median (1 case) 2) Values : gross imputed : median
PY090	Unemployment benefits	I73	Early retirement benefits	15 (208)	1) net is given (regresion net-gross) 2) last month value is given 3) hot deck in classes based on gender, age and former economic status

PY090	Unemployment benefits			18.9 (1586)	
PY100	Old age benefits	(I74_b)	Old age pensions	19.8 (1692)	1) net is given (4.1) 2) net is imputed based on last month household total income (question H51) (12.4) 3) IVE auxiliary variables : household net income (H51, h52), former status in employment (i8_a), activity sector (i8_b), NACE (i9), Supervision of otherone's work (i11), size of the firm (i12), income of partner, age, gender (3.7)
PY100	Old age benefits	(I74_c)	Other financial assistance to old aged people[5]	29.2 (24)	Median in classes based on household situation : one-person household or living with partner
PY100	Old age benefits	(I74_d)	Other financial assistance to old aged people[6]	23.8 (21)	Net (1 case) Median
PY100	Old age benefits	(I74_e)	Type of old age benefits not given	23.3 (120)	1) net is given, regression net-gross (10) 2) net is imputed based on actual situation (last month income or household income) (13.3)
PY100	Old age benefits			17.4 (2160)	
PY110	Survivor's benefits[7]	(I74_a)		20.9 (460) among which 4 (460) 10.9 (460) 5.9 (460)	 1) net is given 2) last month value is given (or H51 household total income is given and can be used) 3) hot deck

[5] Revenus garantis aux personnes âgées

[6] Complément au revenu garanti aux personnes âgées

[7] Individuals could answer 'yes' to the filter of question I74_a and be more than 65 years. After imputation, the values of the benefits were classified as old-age benefits (342 cases). The variables PY100G and PY110N are thus filled in for 118 cases among which 17.8 % were imputed.

PY120	Sickness benefits	(I87)	Paid sick leave (temporary inability to work due to sickness)	24.2 (203)	1) net is given (2 cases) 2) last month value (2 cases) 3) IVE : auxiliary variables : age, status in employment, sector, period (number of months given), has ever worked or not, gross wages for last year (i35_b) (remaining cases : 20.2 %)
PY120	Sickness benefits	(I87)	Paid sick leave (temporary inability to work due to professional sickness or injury)	17.1 (82)	Hot deck
PY120	Sickness benefits	(I87)	Other sickness benefits	36.4 (33)	Hot deck
PY120	Sickness benefits			24 (312)	
PY130	Disability benefits	I87	Disability pension	20.3 (418)	
				2.9	1) net is given (12 cases)
				10	2) last month value (42 cases)
				1	3) last month household income can be used (4 cases)
				6.4	3) IVE (27 cases) :auxiliary variables (gender, household situation (living alone or not), age, age below legal pension age or not, status in employment, sector, has ever worked or not)
PY130	Disability benefits	(I87)	Integration income for the handicapped	26.6 (75)	
				2.7	1) net is given (2 cases)
				2.7	2) last month income is given (2 cases)
				5.3	3) actual household income can be used (4 cases)
				16	3)hot deck according to variables age, gender, has ever worked, health problems (12 cases)
PY130	Disability benefits			21.1 (483)	

PY140	Education-related allowances	H74	Grants, scholarship and other educational help to pupils (of secondary schools)	6.2 (208)	
PY140	Education-related allowances	H76	Grants, scholarship and other educational help to students (of colleges)	9.6 (94)	
PY140	Education-related allowances			22.9(288)	
PY200G	Gross monthly earnings for 2003	I18-gross I19-net I20-scale			
				1) 2.8 (4899)	1) No gross, but scale was given: imputation of mean of interval
				2) 9.2 (4899)	2) No gross, but net was given: imputation with regression model using net amount
				3) 4.1 (4899)	3) No gross, neither net was given: imputation on basis of yearly income 2002 (checked for monthly income, change of job, …)

When it is not described, the *number of months* has always been imputed based on the calendar; further, if no calendar was usable, random draw according to the frequency table of the given numbers took place (Note that the calendar questions were only asked to persons who admit that something changed in their situation since 2002, not everyone did answer it).

Imputation of non-responses concerned only gross amounts when gross and net values were collected. One method of imputation for gross values was of course to use the net amounts when given. The percentages of imputation and the methods that are presented in the table concern *only gross variables, except for HY090* (Interests, dividends, etc.) where gross was not recorded.

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 RB210 (basic activity status) of the individual. When the answer is missing or 4 (other inactive person), it is chosen not to impute any income. The method for

imputation differs with the categories: imputation based on a regression for the wages (no difference between employee and employer, independent variables are age and gender), imputation of a sub-category median for the unemployment and retirement incomes. Net incomes were computed with a gross to net model, based on the imputed gross incomes.

The method used for obtaining income target variables in the required form

From net to gross

Gross income at component level had to be provided to Eurostat. Since in the Belgian survey the respondent provided sometimes only net income, imputation of gross incomes on the basis of collected net incomes had to be carried out. The overall procedure used for this is described here.

Missing values on gross income variables (PY010G, PY020G, ... and components) were, if collected, imputed on the basis of the corresponding net variables (PY010N, PY020N, ... and components). The implementation of this imputation procedure was quasi-similar for almost all (income) variables on which it was applied. The procedure implied 6-steps:

1. Identification of the 'reference cases' (both gross and net collected) and identification of the cases to be imputed (net collected – gross missing).
2. Calculation of the gross/net ratio for the reference cases. Cases with an extreme value on this ratio were excluded from further use in the procedure.
3. Curve estimation of the relation (regression model) between gross and net income. The best fitting model (linear, logarithmic, quadratic, exponential) was being implemented.
4. Implementation of the regression model for the reference cases to identify outliers.
5. Re-implementation of the regression model for the reference cases after removal of the outliers.
6. Actual imputation step: missing (gross) values are imputed on the basis of
    a) net values and
    b) the estimates for the relation between gross and net income assessed in the steps above.

From gross to net

Note that for several income variables, we have a gross value and not a net value. We have imputed a net value following the Belgian rules of taxation used to obtain a net from the gross value. Here is a general explanation of the Gross Net Model.

First, we have transformed the gross value in a monthly gross value.

Second, according to the type of income, the social security rules of taxation were applied.

Third, we have transformed the monthly gross value in a yearly gross value.

Four, according to the type of income, the tax withholding and the advance tax payment were taken into account.

Following these steps, we have obtained the yearly net values

III. Autres variables: company car, imputed rent

La possibilité d'utiliser une **voiture de société** à des fins personnelles est considérée comme un revenu dont la valeur doit être imputée dans les revenus non monétaires des employés. Le gain pour les individus d'utiliser une voiture de société à des fins personnelles n'était pas recueilli directement auprès des répondants. Le montant est calculé en appliquant les règles fiscales.

Le **loyer imputé** se réfère à la valeur de loyer qui doit être imputée pour tous les ménages qui rapportent qu'ils ne paient pas de loyer, soit parce qu'ils sont propriétaires, soit parce qu'ils vivent dans un logement qu'ils louent à un prix inférieur à celui du marché, soit encore parce qu'ils occupent un logement gratuitement. Cette variable n'est pas encore calculée. Elle sera inclue dans la liste des variables au plus tard pour SILC 2007, mais l'information a été recueillie en vue de permettre une première analyse interne des résultats et faciliter le choix de la méthode.

IV. Création des target variables

Avant la livraison des fichiers de données à Eurostat, de nombreux traitements des données sont encore nécessaires, notamment la création de nouvelles variables, comme par exemple, le revenu total brut du ménage, calculé comme la somme des composantes du revenu personnel brut de tous les membres du ménage plus les composantes du revenu brut au niveau du ménage moins le remboursement d'intérêts hypothécaires.
Un autre variable concerne les poids (voir annex 4 sur l'extrapolation).

V. Validation

Parmi les dernières phases dans le traitement des données, on trouve la validation, le contrôle de vraisemblance selon programme d'Eurostat sous SAS et la rédaction du rapport de qualité.
La validation des données est une procédure qui consiste à revoir les données avant de les produire officiellement pour s'assurer que des données grossièrement erronées n'apparaissent plus et pour identifier les données de moindre qualité.
Par exemple, la Commission ne publie pas les estimations si elles sont basées sur moins de 20 observations ou si le taux de non-réponse pour le poste concerné est supérieur à 50%. Les données sont publiées avec un 'flag' si l'estimation est basée sur 20 à 49 observations ou si le taux de non-réponse pour le poste concerné dépasse 20% tout en étant inférieur ou égal à 50%. Les données sont publiées par la Commission selon les modalités normales lorsqu'elles sont basées sur au moins 50 observations et que le taux de non-réponse pour le poste ne dépasse pas 20%.
Dans le même ordre d'idée, il est nécessaire de quantifier le travail d'imputation, nommer les variables pour lesquelles on a utilisé des méthodes d'imputation et préciser quelles méthodes, indiquer les valeurs imputées (flags) et donner des taux d'imputation.
Le rapport de qualité demande une estimation de toutes les erreurs possibles : couverture, mesure, non-réponse, procédure, ...

VI. Création des indicateurs de cohésion sociale

Un des but final de l'enquête SILC est la production d'indicateurs de pauvreté et d'inclusion sociale qui soient comparables au niveau européen. Les données finales doivent permettre de calculer ces indicateurs, dont l'un des plus connus est le pourcentage d'individus vivant dans des ménages qui ont un revenu inférieur à 60 % du revenu médian national (at-risk-of-poverty rate).

Annex 4. Extrapolation for SILC 2003

In order to extrapolate the respondents' answers (6199 responding and acceptable households, 15154 individuals, we followed Eurostat's instructions (EU-SILC 134-04).

♦ **Step 1: design weights**

These weights (inverses of inclusion probabilities) were easy to compute: since our two-stage sampling design was self-weighting within each stratum (= province = NUTS2), this weight equals $\frac{total\ number\ hh\ in\ stratum}{number\ hh\ drawn\ in\ stratum}$ (the circumstance that the number of "November hh" could differ from the number of "September hh" is irrelevant here).

	N	Minimum	Maximum	Mean	Std. Deviation
Design weight	6199	181,55	435,94	330,4121	75,85014

♦ **Step 2: non-response weights**

The idea is clear:

1. Division of the sample households into groups or 'weighting classes', showing the number of households selected and the number of households with interview acceptance equal 1 in each class.

2. Computation of the response rate for each category in the classification.

3. Assigning a uniform weight to all households in a category, in inverse proportion to the response rate of the category.

...but the implementation (defining *classes*) not that easy

Between

a. geographic location (NUTS2 for us),
b. type of place of residence (urban-mixed-rural classification),
c. various socio-economic characteristics of the areas.,
d. tenure status,
e. dwelling type
f. household size (1, 2, 3 and 4+, according to the National Population Register)
g. household type,
h. other characteristics which may be related to the level of household living conditions

we took only a, b and f (d being available only for respondents –more about that soon!, c, e, g & h too fuzzy). The crossing of a and b produced 24 (<11*x*3 : some cells were empty!) geographic zones, which, combined with the 4 hh sizes, yielded 96 classes.

	N	Minimum	Maximum	Mean	Std. Deviation
Nonresponse weight	6199	354,45	2420,36	710,9437	207,09730

♦ **Step 3: Adjustment ... to external source**

We took 28 dummy individual calibration variables (2 sexes x 14 age classes), 4 dummy hh variables (1, 2, 3 and 4+, according to the interviewer's findings, which in

a few cases differed from the NPR[8]). The totals correspond to the NPR figures (private hh only) as of 31 December 2003.
After a lengthy discussion, we decided to add a 33rd dummy variable (owner (or rent-free)/tenant), using as calibration totals the results of the 2001 Census. We thought that the goal (giving a higher weight to owners, under-represented among the respondents) outweighed the drawbacks (different sources for respondent hh SILC 2003 and Census 2001[9], higher spread of weights[10]). The impact on the results was marginal (increasing the mean equivalised income by €80).
Recall that all hh members have to be given the same weight.
The computations were performed with g-Calib under SPSS, under a scheme that guarantees positive weights.

	N	Minimum	Maximum	Mean	Std. Deviation
Weight (without tenure)	6199	10,22	6725,21	707,9103	395,18494
Final weight (with tenure)	**6199**	**8,57**	**6734,45**	**707,9103**	**400,19015**

♦ **Step 4 : Variance estimation**
The estimated (at-risk-of-)poverty rate is 15,8; we estimate the standard deviation thereof 0,48%, corresponding to a 95%confidence interval 14,86 – 16,74.....

[8] Correlation 0,98; the NPR seems to underestimate slightly the size.

	N	Minimum	Maximum	Mean	Std. Deviation
Size according to NPR	6199	1	18	2,43	1,374
Size according to interview	6199	1	19	2,45	1,376

[9] This problem will rise again in the "coming" years (Extrapolation of SILC 2004...)
[10] And (probably) design effect; note that the (unweighted!) means of weights coincide (6199 x 707,9103 = 4388336 = number of private hh) .

PART 2: FIRST RESULTS

In this part we present the first results of the current cross-sectional SILC 2003. We emphasize the provisional nature of these results, however, as Eurostat is currently in the process of validating the data.

♦ Background.

'At the Nice European Council in December 2000, Heads of State and Government reconfirmed and implemented their March 2000 (Lisbon) decision that the fight against poverty and social exclusion would be best achieved by means of the open method of coordination.
Key elements of this approach are the definition of commonly agreed objectives for the European Union (EU) as a whole (eg. The 'social policy agenda'), the development of appropriate national action plans to meet these objectives (eg. 2001-03, 2003-05), and the periodic reporting and monitoring of progress made.
It is in this latter context that the Laeken European Council in December 2001 endorsed best practice criteria for indicator design, and a first set of 18 common statistical indicators for social inclusion, which will allow monitoring in a comparable way of Member States' progress towards the agreed EU objectives. These indicators need to be considered as a consistent whole reflecting a balanced representation of EU social concerns. They cover four important dimensions of social inclusion (financial poverty, employment, health and education), which highlight the "multidimensionality" of the phenomenon of social exclusion. These indicators are often referred to as "the Laeken indicators" ' (EU-SILC 131-rev/04)[11].

Atkinson e.a. (2000) propose a three level classification of indicators for use in the EU monitoring process. From this point of view some of the 'Laeken indicators' are level 1 (lead) –indicators, others are level 2 (complementary to level 1). Level 3 consists of indicators included in the member states' National Action Plans on Social Inclusion (NAPinc). These indicators focus on specificities in particular areas.

♦ Definitions and strategy

For the (current) cross-sectional component of EU-SILC 2003 the six participating members calculated a subset of the Laeken indicators (six indicators), equivalised disposable income and the gender pay gap.

In addition to these indicators Belgium already calculated several other indicators measuring non-monetary dimensions of poverty and social inclusion in particular. Several of these indicators are included in the Belgian 'National Action Plan on Social Inclusion' (NAPincl).

[11] EUROPEAN COMMISSION EUROSTAT, *Common Cross-sectional EU indicators based on EU-SILC; the gender pay gap*, Eurostat, EU-SILC 131-rev/04

Table 1: Indicators calculated for the current cross-sectional component of EU-SILC.

Primary indicators
At-risk-of-poverty rate (after social transfers)
Inequality of income distribution S80/S20 income quintile share ratio
Relative median at-risk-of-poverty gap, by age and gender
Secondary indicators
Dispersion around the at-risk-of-poverty threshold
At-risk-of-poverty rate before social transfers by age and gender
Inequality of income distribution: Gini coefficient
Other indicators
Equivalised disposable income
The gender pay gap
Non-monetary indicators*
Ability to make ends meet
Indicators related to the dwelling: quality, lack of comfort, space
Capacity to face unexpected financial expenses
Capacity to afford a meal with meat, chicken ....
Capacity to afford an annual holiday
Health related indicators
Possession of durables
...

*The majority of the non-monetary indicators are included in the NAPincl.

In **annex 1** we give extensive information on both calculation and interpretation of the indicators. We also explain the breakdowns we used. We disaggregated several of the indicators we calculated by the so-called 'Laeken-breakdowns'; gender, age, most frequent activity status, household type, accommodation tenure status and work intensity.

Combining monetary and non-monetary indicators of poverty has proven a particular successful strategy to identify particularly deprived individuals and households (Atkinson, e.a. 2002)[12].
We adopted this strategy and cross-classified several non-monetary indicators with income and indicators of financial poverty.
These cross-tabulations also serve an important other purpose. They are an elementary check of the (internal) validity of the data as they allow identification of contradictory outcomes on related monetary and non-monetary indicators

Before presenting the results of the indicators we start our outline of the results below with a short look at some relevant features of the survey itself and the distributions of some key-demographic and economic variables.

[12] ATKINSON T., CANTILLON B., MARLIER E., AND NOLAN B., *Social Indicators: the EU and Social Inclusion*, Oxford University Press, Oxford, 2002, p. 240

♦ **Key characteristics: achieved sample size, response rate and distributions on principle demographic and economic variables.**

In **table 2** we list a number of important figures characterizing the survey for both households and individuals.

Table 2 : survey characteristics and key demographic and economic variables

Households	
Sample characteristics	
In sample [13]	11254
Achieved	6199
Response rate	45,9 %
Household type	
1 person	32,9 %
2 adults, both <65 years	17,6 %
2 adults, at least one 65+ years	9,3 %
Other no dependent children	3,6 %
Single parent	9,4 %
2 adults, 1 dependent child	9,2 %
2 adults, 2 dependent children	9,5 %
2 adults, 3+ dependent children	5,6 %
Other with dependent children	2,6 %
Individuals (16+)	
Sample characteristics	
In sample	11873
Achieved	12004
Response rate	45,4 %
Sex	
Males	50,9 %
Females	49,1 %
Age	
0 - 14	20,6 %
15 - 64	66,6 %
65+	12,7 %

[13] Number of eligible households at contacted addresses.

Table 2 continued

Education (highest ISCED level attained)	
Pre-primary education	6,9 %
Primary education	14,4 %
Lower secondary education	16,4 %
(upper) secondary education	32,9 %
Post-secondary non tertiary education	3,3 %
First stage of secondary education	25,5 %
Second stage of tertiary education	0,5 %
Self defined current economic status	
Full time job	39,1 %
Part-time job	10,7 %
Unemployed	9,1 %
Pupil, student	9,5 %
In (early) retirement	19,5 %
Permanent disabled	3,1 %
Fulfilling domestic task	7,2 %
Other inactive persons	1,8 %

♦ **Results for the monetary indicators: equivalized disposable income, the Common cross-sectional EU indicators and the unadjusted gender pay gap**

In **annex2-table 3** an overview of the common cross-sectional EU indicators, mean equivalised income and the unadjusted gender pay gap is presented. Also the standard errors for these indicators are reported (**table 3B**).

Our results indicate that 15,8% of the population is at risk of poverty. This means they are living in a household where the disposable income (after equivalisation for household size) is below 60% of the national median income at individual level (below 9098,64 Euro).

Overall women and older persons (65+) seem to be more at risk of poverty than men and persons under the age of 65.
Having a job is an important buffer against poverty. The poverty rate of the working is several times lower than the rate for the inactive.
Persons living on their own and single parents in particular are especially vulnerable to poverty.
Tenants are more at risk than home-owners for poverty, or vice versa, poorer people are more likely to rent their dwellings.

The income quintile share ratio (S80/S20) examines the ratio of the sum of (equivalised disposable) income received by the 20% of the country's population with the highest income (top interquintile interval) to that received by the 20% of the country's population with the lowest income (lowest interquintile interval). Our results indicate that the top interquintile income interval has 5 times the income of the lowest interquintile interval. By definition this ratio only considers changes in the top and bottom income quintiles.
In the calculation of the Gini-coefficient the complete income distribution is considered. The Gini-coefficient is defined as the relationship of cumulative shares of the population arranged according to the level of equivalised disposable income, to the cumulative share of the equivalised total disposable income received by them. A Gini of 100% implies complete inequality, a Gini of 0% perfect income equality within society. The Gini stands at 30% in our data.

The relative median risk of poverty gap is the difference between the median equivalised disposable income of persons below the at-risk of poverty threshold and the at-risk of poverty threshold itself, expressed as a percentage of the at-risk-of-poverty threshold. For our data this indicator equals 23,1% overall. This means that half of the persons at risk for poverty disposed of an income at of above 76,9 % of the at risk of poverty threshold.

In **annex3-table 4** we compare the results of SILC with the results of ECHP – wave 2 up to 8. For the majority of the indicators marked differences are observed between SILC and ECHP. We emphasis however that comparison between SILC and ECHP is not straightforward (see annex 5). Therefore caution in the interpretation of these differences is necessary.

♦ **Results for the non-monetary indicators**

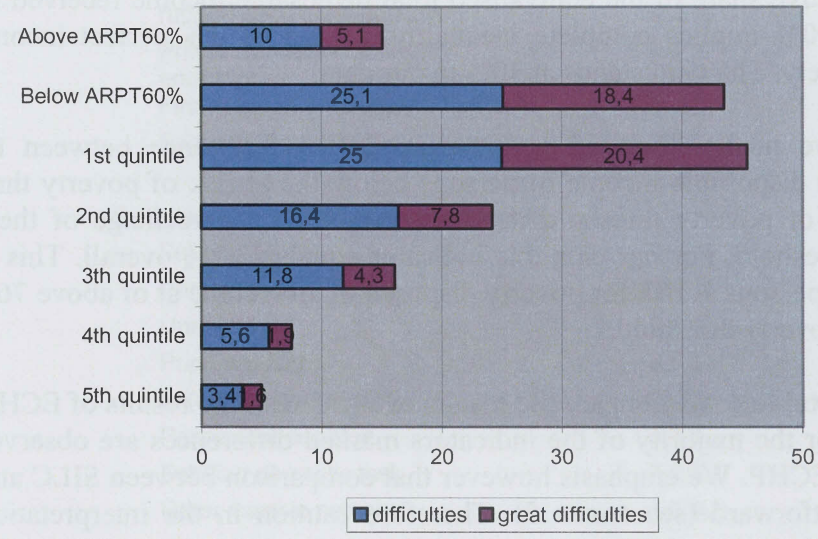
In **annex 4 – table 5** we present the result for the non-monetary indicators and variables. In table 5 we give for each of these indicators overall results (column 'total') and results cross-tabulated by income (below or above 60% threshold and the income quintiles).

As we already explained combining non-monetary variables and income facilitates identification of particularly poor persons. The concept of 'consistent poverty' is of particularly interest for this matter. 'Consistent poverty' refers to people both falling below a predefined income threshold (the 60% threshold of median income in this case) and reporting an enforced lack of a common set of items (Atkinson, et.al. 2002; p.121). The combination of several of the non-monetary indicators and income in table 4 allow an idea of the level consistent poverty in the survey.

An indication of the level of persons in the survey experiencing extreme poverty (but not necessarily in consistent poverty) can be found in figure 1. The figure shows the result of cross-tabulation of the ability to make ends meet and the position of one's disposable income relative to 60% threshold (ARPT60) and to the total income distribution (the quintiles).
Overall, almost one fifth of the persons in the survey (19,6%) lives in a household of which the reference persons indicates difficulties to make ends meet. For 7,2%

(included in the former overall %) the reference person even indicates great difficulties.
43,5 % (25,1% + 18,4%) of the persons below the 60% threshold lives in households experiencing difficulties to make end meet. They are, by any standard, particularly deprived.
With regard to the validity and quality of the data we find the consistently decreasing number of persons reporting difficulties in the upper quintiles reassuring.

Figure 1 : % individuals living in households where the reference person indicates (great) difficulties to make ends meet.
Results for individuals below and above ARTP60% and the income quintiles.



♦ Validity and quality of data issues

In order to anticipate on a number of concerns about both the external and internal validity of the data we added two more annexes.

In the first one - **annex 5** - we compare SILC on key-variables with several other statistics (ECHP, LFS and Social economic survey).

In **annex 6** we scrutinized the results for the self-employed and the lowest income households to anticipate on concerns about the validity of the results for these subgroups.

Annex 1. Information on interpretation and calculation of indicators

Indicator or variable	Calculation and/or interpretation
Equivalised disposable income	For each person, the 'equivalised disposable income is defined as his/her total household disposable income divided by equivalised household size. The total disposable income of a household is calculated by adding together the personal income received by all of household members plus income received at household level, once corrected by 'within-household non-response inflation factor to compensate the non-response in individual questionnaires'. The equivalised household size is defined according to the modified OECD scale (which gives a weight of 1.0 to the first adult, 0.5 to other household members aged 14 or over and 0.3 to household members aged less than 14).
At-risk-of-poverty rate 'ARPR' (after social transfers)	The percentage of persons, over the total population, with an equivalised disposable income below the 'at-risk-of-poverty threshold'. The at-risk-of poverty threshold is set at 60% of the national median equivalised disposable income. The value of at-risk of poverty threshold in national currency will be converted into EURO (if necessary) and into PPS. The conversion of national currency values into Euro and PPS will be done using official exchange rates and PPS published by Eurostat: *New Cronos, Theme 2, Domain 'Price', Collection 'PPP', Table 'PPPSNA95'.*
At-risk-of poverty rate (after social transfers) broken down according to certain variables	The at-risk-of-poverty rate (after social transfers) will be broken down by the following variables: age and gender, most frequent activity status during the income reference period, household type, accommodation tenure status, work intensity The 'at-risk-of poverty rate (after social transfers) broken down by according to these variables is calculated as the percentage of persons in each category with an equivalised disposable income below the 'at-risk-of-poverty threshold' over the total population in the same category (i.e. for each breakdown, the equivalised disposable income of each person is compared with the at-risk-of-poverty threshold calculated for the total population.) The most frequent activity status is the status that individuals declare to have occupied for

continued	
	more than half the total number of months for which information on any status in the variables of activity status during the income reference period are available. The 'at-risk-of-poverty rate (after social transfers) broken down by work intensity categories and broad household types is calculated as the percentage of persons in work intensity and household type (over the total population in the same group) with an equivalised disposable income below the 'at-risk-of-poverty threshold'. The work intensity (WI) of the household refers to the number of months that all working age household members have been working during the income reference period as a proportion of the total number of months that could be theoretically be worked within the household.
Inequality of income distribution S80/S20 income quintile share ratio	Ratio of the sum of equivalised disposable income received by the 20% of the country's population with the highest equivalised disposable income (top interquintile interval) to that received by the 20% of the country's population with the lowest equivalised disposable income (lowest interquintile interval).
Relative median at-risk-of-poverty gap by age and gender	The difference between the median equivalised disposable income of persons below the at risk of poverty threshold and the at-risk of poverty threshold itself, expressed as a percentage of the at-risk-of-poverty threshold.
Dispersion around the at-risk-of-poverty threshold	The percentage of persons, over the total population, with an equivalised disposable income below 40%, 50% and 70% of the national median equivalised disposable income.
At-risk-of-poverty rate before social transfers except old-age and survivors' benefits	The 'at-risk-of-poverty rate before social transfers except old-age and survivors' benefits' shows the percentage of persons (over the total population) having an equivalised disposable income before social transfers except old-age and survivors' benefits below the national 'at risk-of-poverty threshold'.

At-risk-of-poverty rate before social transfers including old-age and survivors' benefits	The 'at-risk-of-poverty rate before social transfers including old-age and survivors' benefits' shows the percentage of persons (over the total population) having an equivalised disposable income before social transfers except old-age and survivors' benefits below the national 'at risk-of-poverty threshold'.
Inequality of income distribution: Gini coefficient	The Gini coefficient is defined as the relationship of cumulative shares of the population arranged according to the level of equivalised disposable income, to the cumulative share of the equivalised total disposable income received by them.
Gender pay gap in unadjusted form	The 'gender pay gap in unadjusted form' is the difference between men's and women's average gross hourly earnings as a percentage of men's average gross hourly earnings. The gender pay gap is given as the difference between average gross earnings of male paid employees and of female paid employees as a percentage of average gross earnings of male paid employees. The population consists of all paid employees aged 16 – 64 that are 'at work > x hours per week'.
Difficulties to make ends meet	The household respondent's assessment of the level of difficulty experienced by the household in making ends meet. A household may have different source of income and more than one household member may contribute to it. Thinking of the household's total monthly income, the idea is with which level of difficulty the household is able to pay its usual expenses. We calculate the percentage of the individuals in the households that declare to have difficult or very difficult experience to make ends meet.
Non-capacity to afford a meal with meat, chicken, fish (or vegetarian equivalent) every second day	Whether, according to the household questionnaire respondent, the household can afford a meal with meat, chicken or fish (or equivalent vegetarian) every second day, if wanted. We calculate the percentage of the reference persons of the households that hasn't got the capacity to afford a meal with meat, chicken, fish (or vegetarian equivalent) every second day.

Non-capacity to face unexpected financial expenses	Whether, according to the household respondent, the household can face itself unexpected financial expenses. We calculate the percentage of the reference persons of the households that hasn't got the capacity to face unexpected financial expenses.
Non - capacity to afford paying for one-week annual holiday away from home.	It is about the ability to pay, regardless of whether the household actually want the item. The household has to have the capacity to pay for seven days the accommodation, and this for all the household members. The amount that the household is able to pay is inappropriate to evaluate the capacity of a household to afford paying for one-week annual holiday. We calculate the percentage of the reference persons of the households that hasn't got the capacity to afford paying for one-week annual holiday away from home.
Possession of car, (mobile) telephone, washing machine, television, computer, internet	Whether the household has the item or whether the household does not have the item because it cannot afford it (enforced lack) or for other reasons. 'Enforced lack' implies that the item is something that the household would like to have, but cannot afford. Possession of the item does not necessarily imply ownership: the item may be rented, leased or provided on loan. The household should be considered to possess the item if any member possesses it.
Lack of bath or shower in dwelling	Whether the dwelling has proper room with a bath or a shower. We calculate the percentage of the individuals in the households that doesn't have a proper room with a bath or shower.
Lack of indoor flushing toilet for sole use of household	If the household has an indoor flushing toilet in the dwelling and this for the sole use of household. We calculate the percentage of the individuals in the households that doesn't have a flushing toilet for sole use of household.
Leaking roof, damp walls/floors/foundation, or rot in window frames or floors	Whether, in the judgment of the household respondent, the dwelling has a problem with a leaking roof, dampness in the walls, floors or foundation or rot in window frames and doors. We calculate the percentage of the individuals in the households that declare to have problems with a leaking roof, dampness in the walls, floors or foundation or rot in window frames and doors.

Short of space	A household has short of space if there is less than one room[14] for each member of the household. We calculate the percentage of the individuals in the households that declare to have short of space.
Unable to keep home adequately warm	This question is about ability to pay to keep the home adequately warm, regardless of whether the household actually needs to keep it adequately warm. We calculate the percentage of the individuals in the households that declare to have problems with keeping the home adequately warm.
At least one problem: lack of elementary comfort, short of space, problems with ability to keep warm	Lack of elementary comfort, short of space, problem with ability to keep warm are variables calculated on household level. These variables are recalculated on a individual level.
Lack of elementary comfort	A dwelling with lack of elementary comfort misses a proper room with bath/shower or hot running water, or an indoor flushing toilet for sole use of the household. We calculate the percentage of the individuals in the households that have lack of elementary comfort.
Unmet need for medical examination for financial reason	The aim of this variable is to capture the person's own assessment of whether he or she needed a medical examination but was not able to due to financial reason (could not afford it). We calculate the percentage of the individuals in the households that reported unmet need for medical examination for financial reasons.
Bad or very bad health	It is subjective indicator of the general health of the person (self-perceived health). It is not intended to measure temporary health problems. It is expected to include the different dimensions of health, i.e. physical, social and emotional function and biomedical signs and symptoms. It omits any reference to an age. It is not time limited. We calculate the percentage of individuals that declare to have a bad or very bad health.
Limitation in activities because of health	The person's self-assessment of whether they are hampered in their daily activity by any ongoing physical or mental health problem, illness or disability. Activity limitations are defined as 'the difficulties the individual experience in performing an activity'. Limitations should be due to a health conditions. This is a subjective indicator. We calculate the percentage of individuals who are limited in activities because of health problems.

[14] A room is defined as a space of a housing unit of at least 4 square meters such as normal bedrooms, dining rooms, living rooms and habitable cellars and attics with a high over 2 meters and accessible from inside the unit.

Annex 2. – Table 3: Equivalized disposable income, the Common cross-sectional EU indicators and the unadjusted gender pay gap

Mean equivalized income.

16786,05 Euro

Risk – of - poverty threshold.

1 person household	9098,64 Euro
2 adults and 2 dependent children	19107,14 Euro

Risk – of - poverty rate by age and gender.

% below ARPT

	Total	females	males
Total	15,8	17,0	14,6
0-15	17,6	18,6	16,7
0-64	15,0	16,0	14,0
16+	15,3	16,5	14,0
16-64	14,1	15,1	13,1
16-24	17,3	17,7	16,9
25-49	12,9	13,9	11,9
50-64	14,3	15,7	12,9
65+	21,9	23,1	20,3

Risk – of - poverty rate by most frequent activity and gender.

% below ARPT

	total	females	males
Total	15,5	16,7	14,2
At work	6,8	5,6	7,7
inactive total	24,1	24,6	23,5
unemployed	35,4	32,9	38,9
Retired	17,8	18,6	16,9
Other inactive	25,7	26,2	24,6

Risk – of - poverty rate by tenure status.

% below ARPT

Total	15,9
Owner or rent-free	11,9
tenant	27,0

Risk – of- poverty rate by household type.

% below ARPT

total no dependent children	16,3
1 person (total)	22,2
2 adults, both < 65 years	11,4
2 adults, at least one 65+ years	20,8
Other no dependent children	6,8
total dependent children	15,7
single parent, at least 1 dependent child	31,2
2 adults, 1 dependent child	10,8
2 adults, 2 dependent children	8,6
2 adults, 3+ dependent children	17,4
other households dependent children	12,9

Risk – of - poverty rate by household type – single households

% below ARPT

Female	24,3
Male	19,8
< 65	21,7
65+	23,2

Risk – of - poverty rate by work intensity

% below ARPT

Household without dependent children	W=0	27,2
	0<W<1	8,9
	W=1	6,3
Household with dependent children	W=0	61,8
	0<W<0,5	31,0
	0,5<W<1	14,1
	W=1	5,9

Dispersion around at – risk—poverty-threshold

% below ARPT

40% of median	5,4
50% of median	9,6
70% of median	25,3

Risk – of - poverty rate by age and gender before all transfers.

% below ARPT

	Total	females	males
Total	41,7	44,9	38,4
0-15	35,0	36,2	33,9
16+	43,7	47,3	39,8
16-64	34,9	37,8	31,9
65+	93,0	92,2	94,1

Risk – of - poverty rate by age and gender before all transfers (including pensions).

% below ARPT

	Total	females	males
Total	30,4	31,6	29,1
0-15	34,6	35,7	33,6
16+	29,1	30,5	27,7
16-64	29,5	30,9	28,1
65+	27,0	28,5	24,8

Relative median risk-of-poverty gap by age and gender.

% below ARPT

	Total	females	males
Total	23,0	21,4	24,3
0-15	24,0	-	-
16+	22,5	21,4	24,1
16-64	25,0	24,6	25,8
65+	16,8	16,8	16,8

S80/S20 quintile share ratio.

5,0

Gini coefficient.

30,0

Annex 2. – Table 3B: Standard errors for equivalized disposable income, the Common cross-sectional EU indicators and the unadjusted gender pay gap

Risk – of - poverty threshold. SE	
1 person household	98,16 Euro

Risk – of - poverty rate by age and gender.
SE prop. below ARPT

	Total	females	Males
Total	0,0048	0,0064	0,0057
0-15	0,0095	0,0123	0,0101
0-64	0,0055	0,0064	0,0058
16+	0,0055	0,0063	0,0060
16-64	0,0053	0,0060	0,0060
16-24	0,0111	0,0135	0,0135
25-49	0,0056	0,0065	0,0067
50-64	0,0108	0,0130	0,0118
65+	0,0170	0,0178	0,0209

Risk – of - poverty rate by most frequent activity and gender.
SE prop. below ARPT

	total	females	Males
Total	0,0048	0,0064	0,0057
At work	0,0043	0,0052	0,0053
unemployed	0,0096	0,0102	0,0116
Retired	0,0220	0,0249	0,0318
Other inactive	0,0128	0,0150	0,0154
total inactive	0,0118	0,0134	0,0172

Risk – of - poverty rate by tenure status.
SE prop. below ARPT

Total	0,0048
Owner or rent-free	0,0058
tenant	0,0127

Risk – of- poverty rate by household type.
SE prop below ARPT

total no dependent children	0,0078
1 person (total)	0,0127
2 adults, both < 65 years	0,0109
2 adults, at least one 65+ years	0,0217
Other no dependent children	0,0206
total dependent children	0,0074
Single parent, at least 1 dependent child	0,0213
2 adults, 1 dependent child	0,0135
2 adults, 2 dependent children	0,0116
2 adults, 3+ dependent children	0,0192
other households dependent children	0,0397

Risk – of - poverty rate by household type – single households
SE prop. below ARPT

Female	0,0174
Male	0,0172
< 65	0,0148
65+	0,0208

Risk – of - poverty rate by work intensity
SE prop. Below ARPT

Household without dependent children	W=0	0,0186
	0<W<1	0,0132
	W=1	0,0088
Household with dependent children	W=0	0,0341
	0<W<0,5	0,0935
	0,5<W<1	0,0146
	W=1	0,0065

Dispersion around at – risk—poverty-threshold

	% below ARPT
40% of median	0,0033
50% of median	0,0043
70% of median	0,0063

Risk – of – poverty rate by age and gender before all transfers.
SE prop. below ARPT

	Total	females	males
Total	0,0058	0,0068	0,0066
0-15	0,0107	0,0143	0,0120
16+	0,0061	0,0070	0,0072
16-64	0,0063	0,0072	0,0077
65+	0,0183	0,0193	0,0230

Risk – of – poverty rate by age and gender before all transfers (including pensions).
SE prop. below ARPT

	Total	females	males
Total	0,0053	0,0062	0,0064
0-15	0,0105	0,0138	0,0120
16+	0,0056	0,0067	0,0068
16-64	0,0065	0,0074	0,0078
65+	0,0091	0,0104	0,0121

S80/S20 quintile share ratio.

0,2027

Gini coefficient.

0,5020

Gender pay gap.

2,5272

Note:

Multiplying these results by 100 and subsequently by 1,96 yields plus-minus 95% confidence levels.

Annex 3. – Table 4: ECHP – SILC Comparison of equivalized disposable income, the Common cross-sectional EU indicators and the unadjusted gender pay gap

2001			ECHP1995	ECHP1996	ECHP1997	ECHP1998	ECHP1999	ECHP 2000	ECHP2001	SILC 2003
Mean equivalised disposable income		EUR	15015	15720	15650	15645	16415	17283	17803	16.786,0
Risk-of-poverty threshold	1 person hh	EUR	8033	8467	8451	8408	8520	8861	9295	9.099,0
	2 adults 2 dep. children	EUR	16869,3	17780,7	17747,1	17656,8	17892	18608,1	19519,5	19.107,0
Risk-of-poverty rate	Total	Total	16	15	14	14	13	13	13	15,8
by age and gender		M	15	14	13	12	11	12	12	14,6
		F	17	17	15	15	14	14	15	17,0
	0-15	Total	16	15	14	13	12	11	12	17,6
	0-64	Total	14	13	12	12	11	11	11	15,0
		M	14	13	12	11	10	10	10	14,0
		F	15	14	13	13	12	12	12	16,0
	16+	Total	16	15	14	14	13	13	14	15,3
		M	15	14	13	12	12	12	12	14,0
		F	17	17	15	15	14	14	15	16,5
	16-64	Total	14	13	12	12	11	11	10	14,1
		M	13	12	11	10	10	9	9	13,1
		F	14	14	13	14	12	12	12	15,1
	16-24	Total	18	17	16	17	16	12	12	17,3
		M	20	19	16	16	16	10	11	16,9
		F	17	16	17	19	16	13	12	17,7
	25-49	Total	11	11	11	10	10	10	10	12,9
		M	10	9	9	8	8	9	8	11,9
		F	12	13	12	12	12	12	11	13,9
	50-64	Total	17	14	13	13	10	10	12	14,3
		M	17	13	12	13	10	9	10	12,9
		F	17	14	13	14	10	11	13	15,7
	65+	Total	25	25	23	22	22	24	26	21,9
		M	24	24	23	22	20	24	24	23,1
		F	26	26	23	22	22	24	26	20,3

Table 4 continued

Risk-of-poverty rate	Total	Total	16	15	14	13	13	13	13	15,5
by most frequent activity		M	14	13	13	11	11	11	11	14,2
by gender		F	17	17	15	15	14	14	15	16,7
(a) Of which:	Total	Total	6	6	5	4	5	5	4	6,8
At work		M	6	6	6	4	5	5	4	7,7
		F	6	6	5	5	4	4	4	5,6
(d) Not at work	Total	Total	25	23	22	21	20	20	22	24,1
		M	26	23	22	21	19	19	21	24,6
		F	24	23	22	21	20	20	22	23,5
	(e) Of which:	Total	34	32	30	31	32	35	32	35,4
	Unemployed	M	47	42	42	38	42	46	40	38,9
		F	27	26	24	26	26	29	27	32,9
	(e) Of which:	Total	21	21	20	18	18	18	21	17,8
	Retired	M	21	21	21	19	18	19	22	16,9
		F	21	21	20	18	18	17	20	18,6
	(f) Of which:	Total	25	22	21	21	19	18	21	25,7
	Other inactive	M	26	22	18	18	15	11	13	24,6
		F	25	22	22	23	21	21	24	26,2
Risk-of-poverty rate	All hh no dep. childr.		17	16	15	14	14	15	16	16,3
by household type	1 person hh	Total	23	24	21	21	19	19	21	22,2
	2 adults no dep. childr.	(both < 65)	12	9	8	8	9	10	8	11,4
	2 adults no dep. childr.	(at least one 65+)	24	23	23	22	22	23	26	20,8
	Other hh no dep. childr.		11	11	9	9	6	11	8	6,8
	All hh with dep. childr.		14	14	13	12	11	11	11	15,7
	Single parent	(at least 1 child)	34	30	31	26	32	27	25	31,2
	2 adults 1 dep. child		10	9	8	6	8	7	7	10,8
	2 adults 2 dep. childr.		11	12	10	8	9	11	11	8,6
	2 adults 3+ dep. childr.		17	16	13	19	11	9	7	17,4
	Other hh with dep. childr.		16	17	18	12	12	11	15	12,9

Table 4 continued

Risk-of-poverty rate	Total		16	15	14	14	13	13	13	15,9
by accommodation	(a) Owner or rent-free		13	12	12	10	10	9	10	11,9
tenure status	(b) Tenant		25	26	23	26	25	27	28	27,0
Risk-of-poverty rate	All hh no dep. childr.	WI = 0	33	28	27	30	22	25	26	27,2
by work intensity of		0 < WI < 1	8	6	7	5	4	6	7	8,9
the household		WI = 1	4	4	5	6	6	4	3	6,3
	All hh with dep. childr.	WI = 0	61	65	58	56	66	68	73	61,8
		0 < WI < 0.5	32	21	39	29	33	27	36	31,0
		0.5 <= WI < 1	10	11	9	11	8	9	7	14,0
		WI = 1	5	6	5	3	3	3	3	5,9
Dispersion around the	(a) 40% of median	Total	4	4	4	3	3	3	2	5,4
risk-of-poverty threshold	(b) 50% of median	Total	9	8	8	8	7	7	6	9,6
	(c) 70% of median	Total	24	23	23	22	22	21	21	25,3
Risk-of-poverty rate	Total	Total	42	42	41	41	40	40	38	41,7
before and after transfers		M	39	38	38	37	36	36	34	38,4
by age and gender		F	45	45	44	44	44	44	42	44,9
(a) before all transfers	0-15	Total	34	34	32	30	30	28	26	35,0
	16-64	Total	34	33	32	32	30	30	28	34,9
		M	32	31	30	29	28	27	25	31,9
		F	36	36	35	35	33	33	31	37,8
	65+	Total	87	88	88	90	89	90	91	93,0
		M	87	86	89	90	88	89	91	94,1
		F	88	89	88	89	89	90	92	92,2
(b) including pensions	Total	Total	27	27	26	25	24	23	23	30,4
		M	26	25	25	24	23	22	21	29,1
		F	28	28	27	27	26	25	25	31,6
	0-15	Total	33	33	31	29	29	26	26	34,6

Table 4 continued

	16-64	Total	25	25	24	24	22	21	21	29,5
		M	24	23	23	22	21	19	18	28,1
		F	26	26	26	26	24	23	23	30,9
	65+	Total	27	28	27	26	26	28	29	27,0
		M	26	27	27	26	25	28	28	24,8
		F	29	29	26	26	27	27	30	28,5
Relative median	Total	Total	21	19	21	20	18	17	15	23,0
risk-of-poverty gap		M	21	17	21	18	17	17	15	24,3
by age and gender		F	21	20	20	20	18	18	15	21,4
	0-15	Total	21	21	22	18	18	16	16	24,0
	16-64	Total	21	18	22	21	18	19	15	25,0
		M	21	17	22	20	17	19	15	25,8
		F	21	18	22	21	18	19	15	24,6
	65+	Total	21	19	17	19	20	16	15	16,8
		M	19	15	17	19	25	15	15	16,8
		F	22	22	17	20	20	17	15	16,8
S80/S20 quintile share ratio			4,5	4,2	4	4	4,2	4,3	4,0	5,0
Gini coefficient			29	28	27	27	29	30	28	30,0

Annex 4. Non-monetary indicators

Table 4: Current cross-sectional non-monetary: Percentage of individuals in households with ...

Indicator	Total	Below ARPT	Above ARPT	1th quintile	2nd quintile	3th quintile	4th quintile	5th quintile
difficulties to make ends meet	19,6	43,5	15,1	45,4	24,2	16,1	7,5	5
non- capacity to face unexpected financial expenses	33,3	58,9	28,4	59,6	44,2	32,4	19,7	10,4
can't afford to receive family or friends every month	12,9	26,9	10,3	26,3	17	11,2	6,7	3,4
can't afford a meal with meat every second day	5,1	13,1	3,6	12,1	5,9	4	2,1	1,7
can't afford paying for one week annual holiday	29,4	51,5	25,2	53,1	38,5	26,9	17,1	11,2
lack of elementary comfort	3,3	7,1	2,6	6,2	5,7	2,6	1,2	1
lack of indoor flushing toilet for sole use of household	1,3	2,7	1	2,4	2,5	0,7	0,3	0,5
lack of bath or shower in dwelling	1,3	2,9	1	2,6	2,1	0,8	0,5	0,5
leaking roof, damp walls/floors/foundation, or rot in window frames or floors	14,9	22,3	13,5	20,5	15,1	15,4	13,0	10,5
less than one room per person	10,2	17,9	8,8	19,1	13	9,2	6,6	3,3
unable to keep home adequately warm	6,4	13,1	5,1	12	6,2	7,1	3,8	2,9
at least one problem: lack of elementary comfort, short of space, problems with ability to keep warm, ...	28,5	44,2	25,6	43,7	32,4	28,3	21,9	16,2
unmet need for medical examination for financial reasons	6,0	14,1	4,4	13,5	7,7	5,7	2,3	0,6
bad or very bad health	7,5	11,4	6,8	12,1	12,1	6,5	4,1	2,8
limitation in activities because of health problems	11,1	13,3	10,6	14,8	15,6	11	8,7	5,4

Annex 5. External validity: SILC – other surveys

First of all we describe operational differences between ECHP and SILC, which could lead to differences in calculated indicators.
Secondly, we check the external validity of the SILC 2003 survey. We will compare the SILC 2003 characteristics with ECHP wave 1 (1994) and wave 8 (2001), Socio Economical Survey (2001), the Labour Force Survey 2003 and the National population register (2003).

♦ Operational differences ECHP –SILC

The EU-SILC replaced the ECHP[15], which was conducted by the Universities of Antwerp and Liege. The last results of the ECHP related to the year 2001. The ECHP survey was also a panel. Although ECHP explored more sociological themes than the SILC survey, ECHP probes mainly income and living condition indicators. The definitions of the two types of studies are broadly comparable, but there are some operational differences. In the ECHP it was for independents impossible to give a negative in-come (debts), for the SILC survey it was possible. Also for pensions there were some differences: for SILC it was collected in a more detailed way than for ECHP. Also for the compound variables (total gross household income, total disposable household income) there are some differences between the two studies. In SILC for the computation of the total disposable household income we took account of the regular tax household cash transfer paid; ECHP didn't. Notwithstanding these differences, the income data from the two sources and the analyses based on them are broadly comparable.
However, the poverty reported in the EU-SILC is not directly comparable with those of the ECHP. Overall the poverty measures recorded in SILC are some 2,8 percentage higher than those for ECHP.
This reflects the sensitivity of indicators of this type to changes in survey methodology. Unfortunately it is not possible to say precisely which methodological changes account for the gap between the 2001 and 2003 results. Ireland (Central Statistics Office, 2005)[16] mentioned the same remark; they mentioned that the deprivation rates recorded in SILC were some 3,5 percentage points higher than those reported in the predecessor of SILC.
Another important difference between SILC and ECHP is the question format; it is to say CAPI[17] versus PAPI[18]. CAPI has several advantages such as a smoother fieldwork (automatic routing) with shorter interview duration, the use of automatic controls when entering the data, which gives better data quality, etc. The Central Statistics Office of Ireland mentions: "*Evidence from other surveys indicates that this tends to result in higher levels of deprivation being reported and it is estimated to have accounted for almost a half of the overall difference observed*"[19].

[15] European Community Household Panel
[16] Central Statistics Office, *EU Survey on Income and Living Conditions (EU-SILC): First Results 2003*, 24 January 2005, Central Statistics Office Ireland, p. 2.
[17] Computer Assisted Personal Interview
[18] Paper Assisted Personal Interview
[19] Central Statistics Office, *EU Survey on Income and Living Conditions (EU-SILC): First Results 2003*, 24 January 2005, Central Statistics Office Ireland, p. 2.

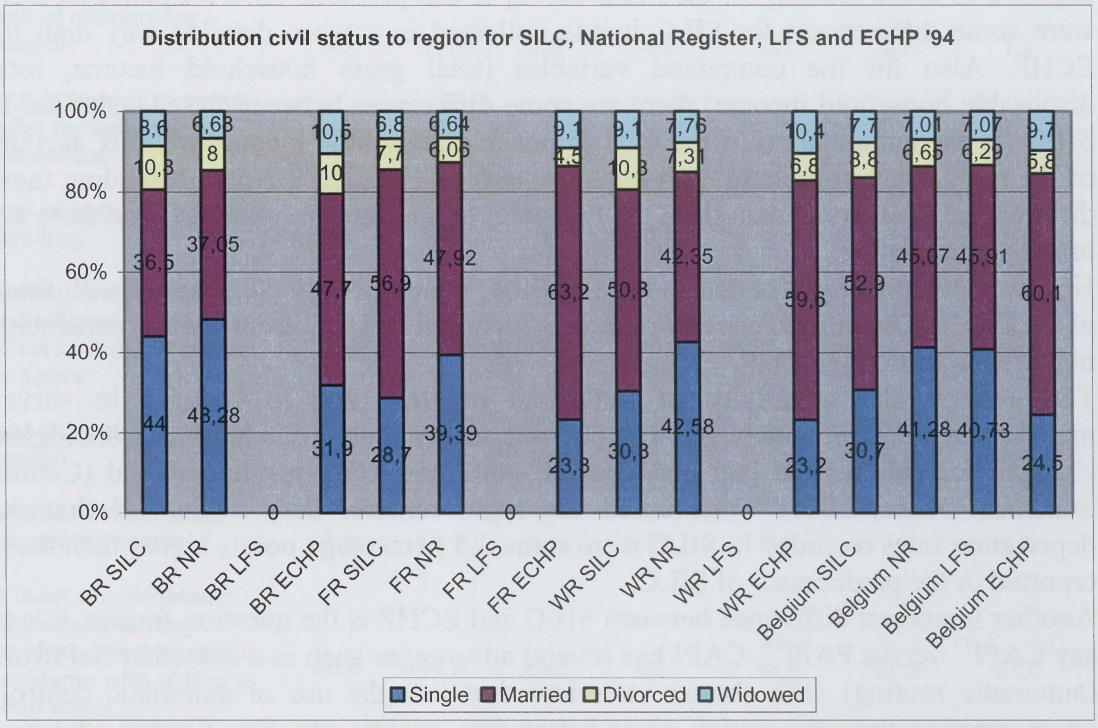
It can be concluded that SILC and ECHP are two surveys that want to measure poverty and living conditions of Belgian households. It is evident that there are some differences between (interview method used, definitions, compositions of indicators,…) these two instruments. These have an influence on the measurement of deprivation and consistent poverty. For this we have to be careful to treat SILC as a successor of ECHP.

♦ **Distribution civil status, sex and age to region, for SILC-LFS and ECHP '94**

If we compare the SILC distribution concerning civil status with the national population register and LFS, which are results for the year 2003, we can see that SILC has 10% less single persons and 7% more married couples, 2% more divorced persons. Concerning widows we see the same percentages.

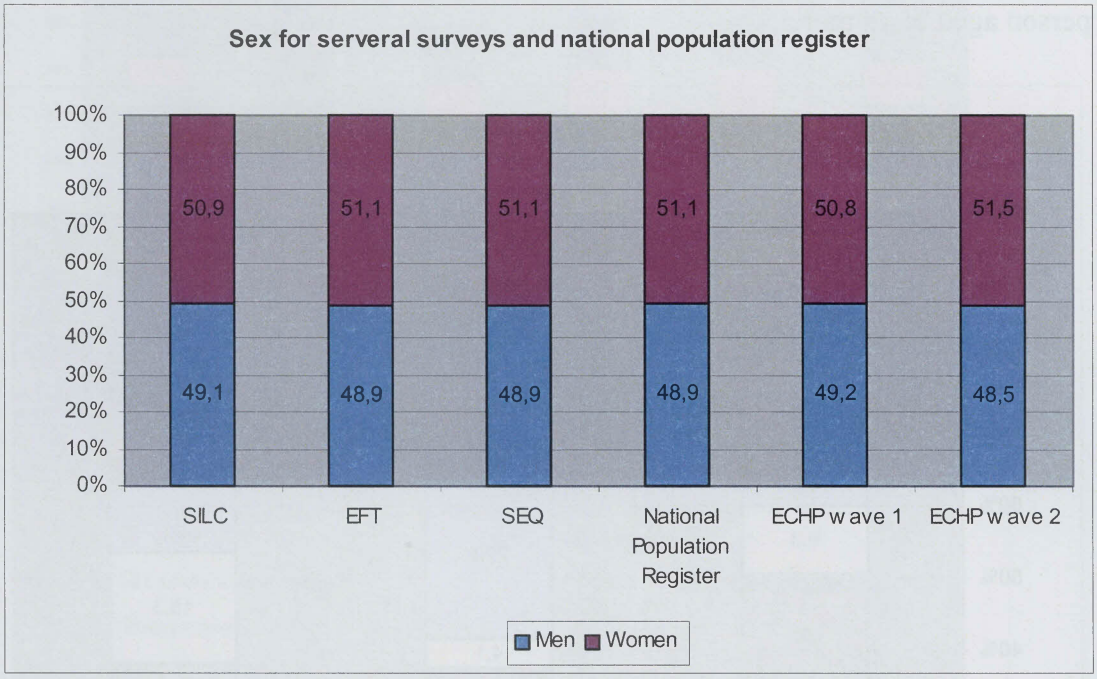
If we compare SILC with ECHP wave 1 we see that ECHP has less single households, less married couples and more divorced and widowed households.

We can argue that the SILC survey is not so representative for civil status. We didn't weight the survey for civil status. We weighted by household type, age distribution and tenure status (as recommended by Eurostat in doc. 123/03).

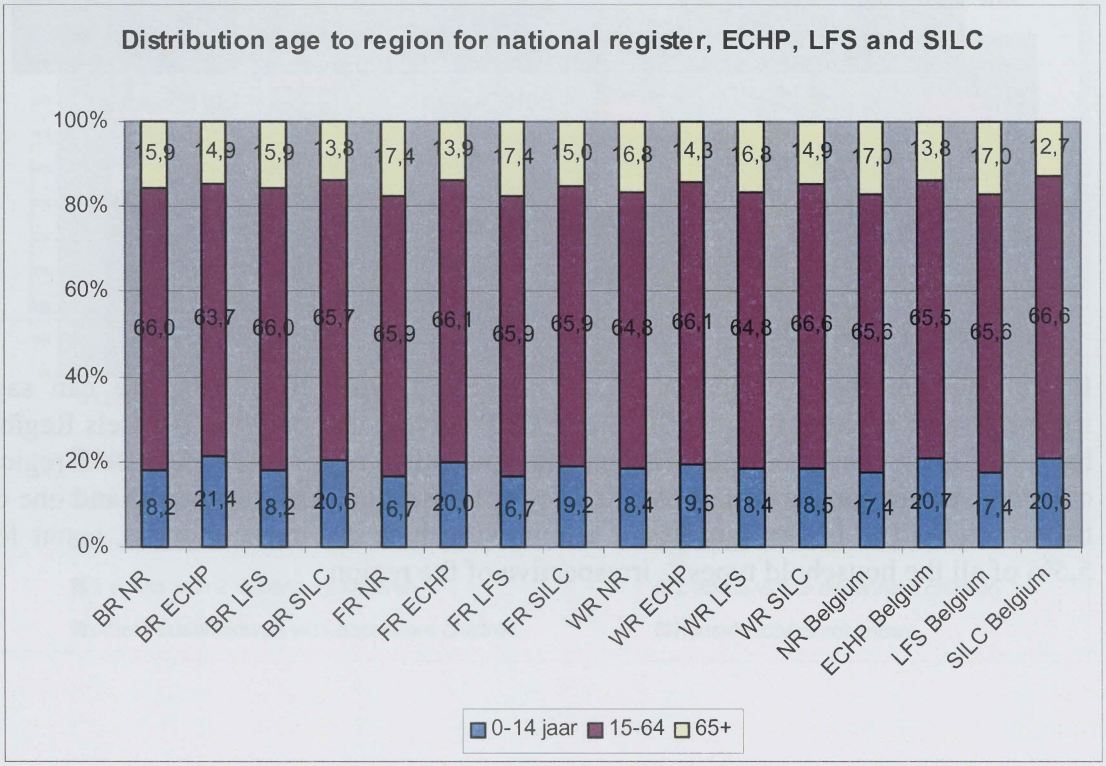


Legend: BR SILC = Brussels' Region SILC, FR SILC = Flemish Region SILC, WR = Walloon Region SILC, Belgium SILC = national data SILC, NR = National population register, LFS =Labour Force Survey, ECHP = ECHP 1994 (wave 1).

The SILC 2003 survey is representative for sex: 51% of the population is female, 49% is male.

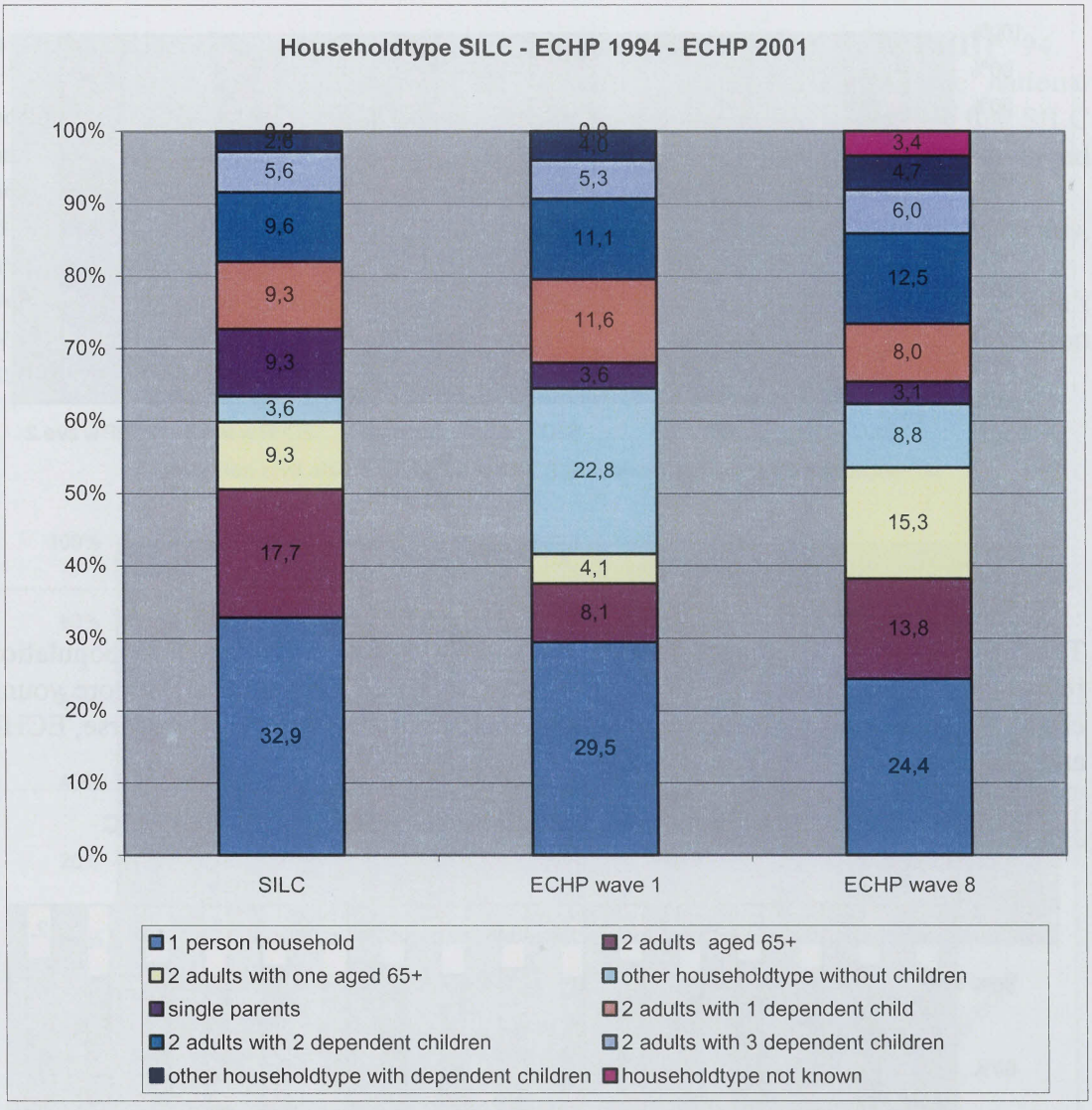


There are no big differences between the 3 surveys and the national population register for age distribution. ECHP 1994 and SILC 2003 have up to 3 % more young children in their survey than LFS or the national population register. In reverse, ECHP and SILC have less elderly.



♦ **Distribution household type for SILC and ECHP**

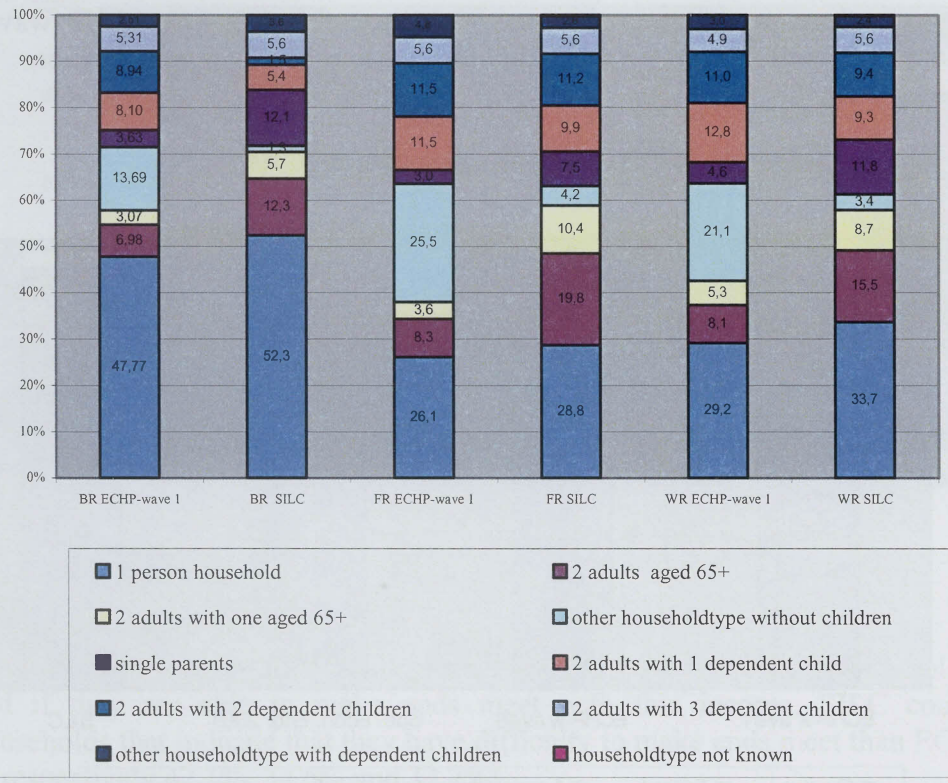
The distribution for the household types for SILC and ECHP (1994 and 2001) is slightly different for the two surveys. In general we can see that the SILC 2003 survey has more one-person households, single parents, and two adults with at least one person aged 65 or more.



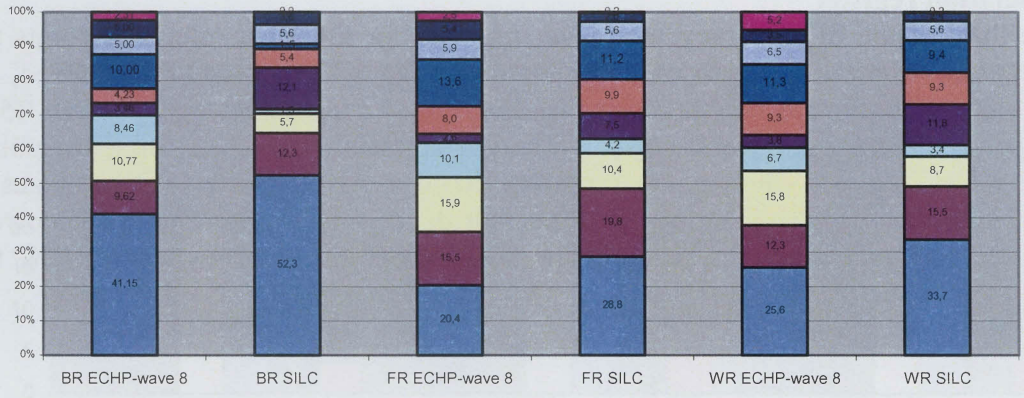
If we compare the distribution of the household types by region, we can say, irrespective of whether it is the SILC or ECHP survey, that only the Brussels Region has a different household pattern than the two other regions. The Brussels region counts more one-person households and fewer households with two adults and one or two children. The bigger families, 2 adults with three ore more children, count for 5,5% of all the household types[20], irrespective of the region.

[20] Only in the Walloon Region of ECHP wave 1 there are less households of two adults with three or more children.

Distribution household type for SILC 2003 and ECHP 1994

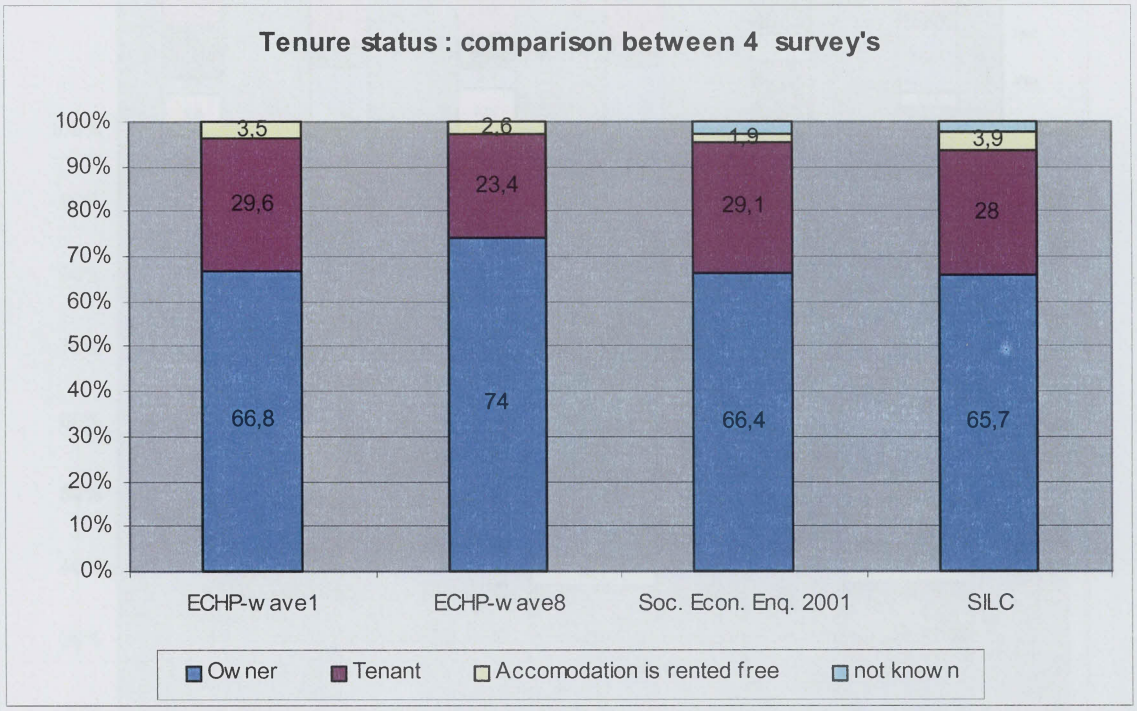


Distribution household type for SILC 2003 and ECHP 2001



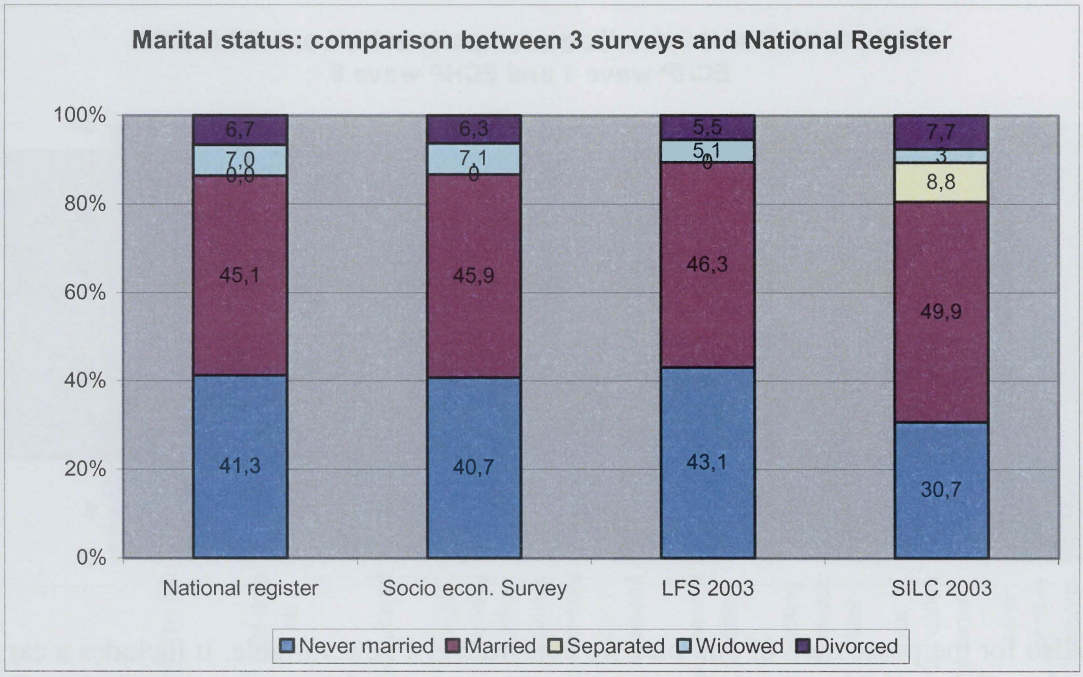
♦ **Distribution of tenure status: comparison between 4 surveys**

The ECHP wave 1, the socio economic survey of 2001 and SILC have similar distributions of tenure status: almost two third are owners, one third is tenant. Of these tenants two to three percent rent the accommodation for free. The ECHP wave 8 counts more owners than the three other surveys: this is probably due to attrition.



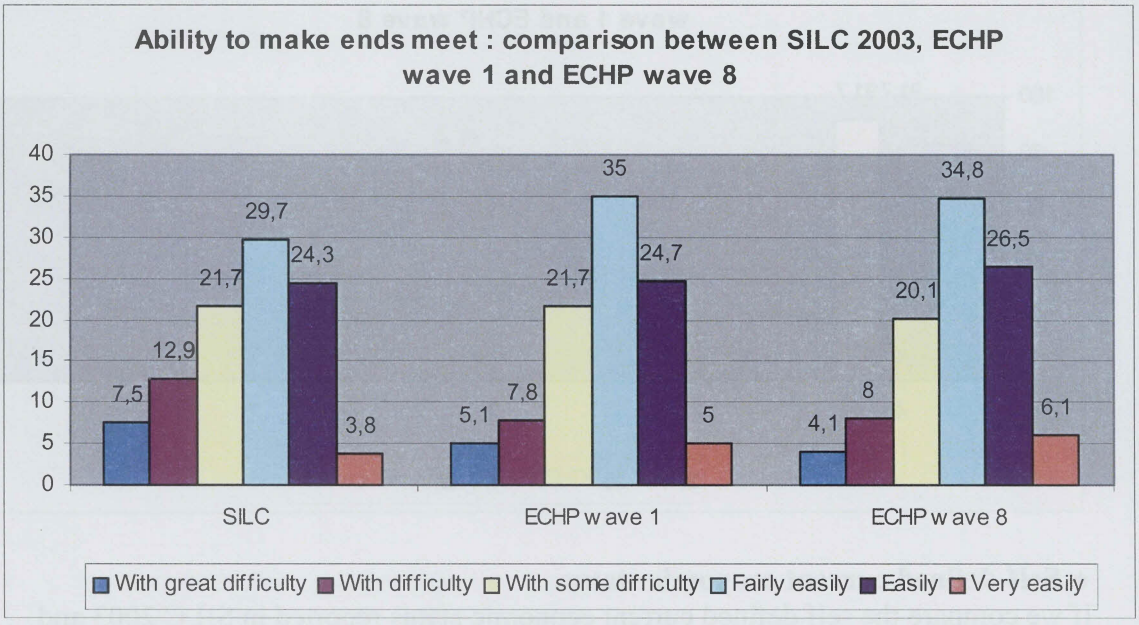
♦ **Distribution of marital status: comparison between 3 surveys and the national population register**

One third of the SILC survey was never married, for the other reference points ten percent more was never married. Fifty percent of the SILC household is married, for the other surveys and the national population register the percentage is about 45%. Respectively 9% and 8% are separated versus divorced. The other reference points counts less divorced couples. For separated couples, we do not have percentages in other sources. SILC counts only three percent of the households who are widowed; the national population register and the socio economical survey counts 7%, the labour force survey 5%.

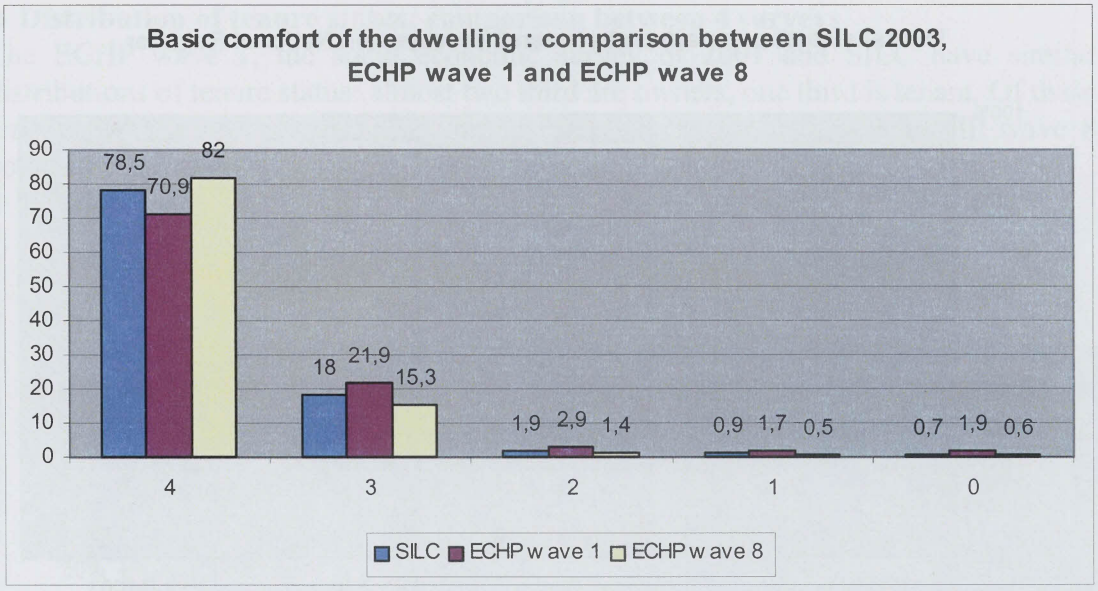


♦ **Non-monetary items**

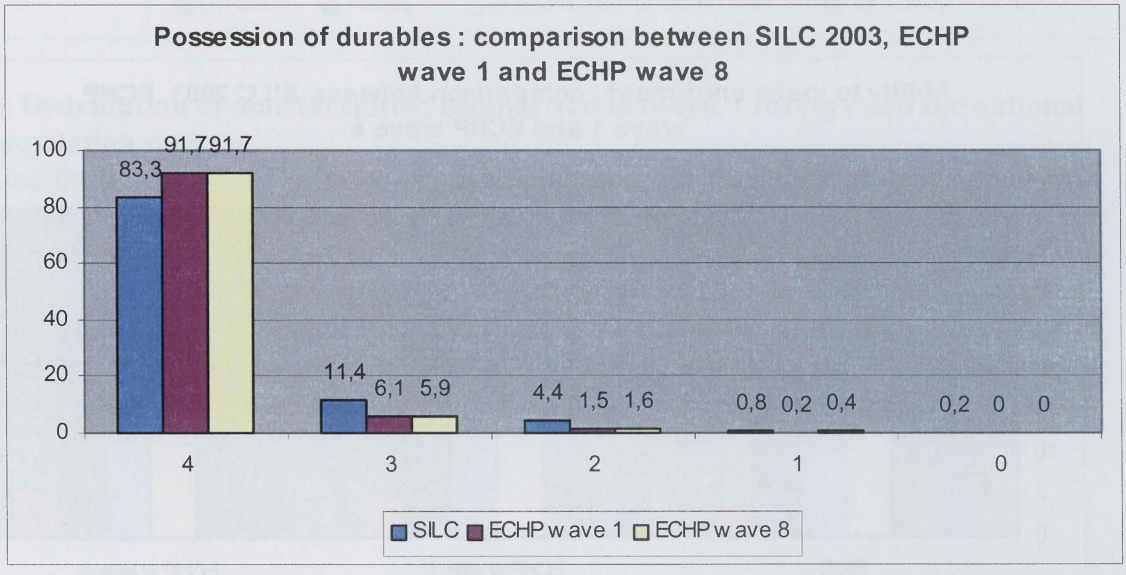
We asked the households to think at their total monthly or weekly household income, and if they are able to make ends meet with this income. SILC counts more households that indicate that they have difficulty to make ends meet than ECHP 1 and 8 (respectively 42,1%, 34,6% and 32,2%).



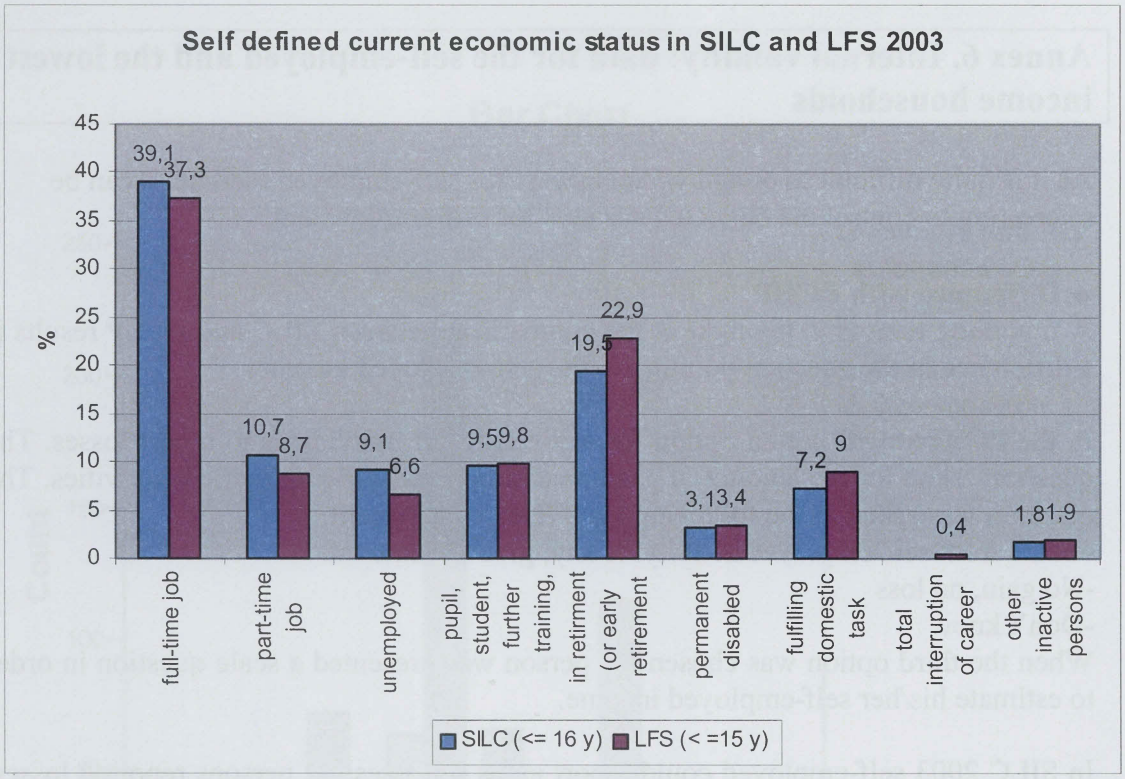
We constructed a variable that measured the basic comfort of a dwelling. This variable includes bath or shower, an indoor flushing toilet, hot running water and central heating. A score of four means that the household has all the basic comfort in their dwelling, a score of zero means that they have no basic comfort at all in their dwelling. ECHP wave 1 counts the lowest percentage households with all the basic comfort; ECHP wave 8 counts the highest percentage.



Also for the possession of durables we constructed a new variable. It includes a car, a colour TV, a telephone and a home computer. A score of four means possession of all durables enumerated, zero means no possession of the durables enumerated.
In the survey ECHP wave 1 and wave 8 92% of the households posses a car, a colour TV, a telephone and a home computer at the same time. For the SILC survey "only" 83% of the household's posses all these durables.



♦ **Self-defined current economic status**
If we compare the self defined current economic status reported in SILC 2003 and LFS 2003 we can't see major differences. Though there is an important difference between the unemployed and the (early) retired persons. SILC counts more unemployed people than LFS. On the other hand LFS counts more retired people.


Self defined current economic status in SILC and LFS 2003
%
45
40
35
30
25
20
15
10
5
0
39,1
37,3
10,7
8,7
9,1
6,6
9,5
9,8
19,5
22,9
3,1
3,4
7,2
9
0,4
1,8
1,9
full-time job
part-time job
unemployed
pupil, student, further training, in retirment (or early retirement
permanent disabled
fulfilling domestic task
total interruption of career
other inactive persons
SILC (<= 16 y)
LFS (< =15 y)

Annex 6. Internal validity: data for the self-employed and the lowest income households

As it is quite difficult to obtain valid answers for self-employed income, it can be interesting to control the collected answers for self-employment.

♦ **Difference with ECHP**

A first thing to note in the light of the comparison between SILC and ECHP results is a difference in the question pointing to the self-employed income.

In the ECHP it was not an option for self-employed individuals to report losses. The question asked for the amount of gains resulting from self-employment activities. The question was asked in the following way: (choose between)
-Gains from self-employed activity. If any, give the amount.
-No gain, no loss
-Don't know
When the third option was chosen the person was presented a scale question in order to estimate his/her self-employed income.

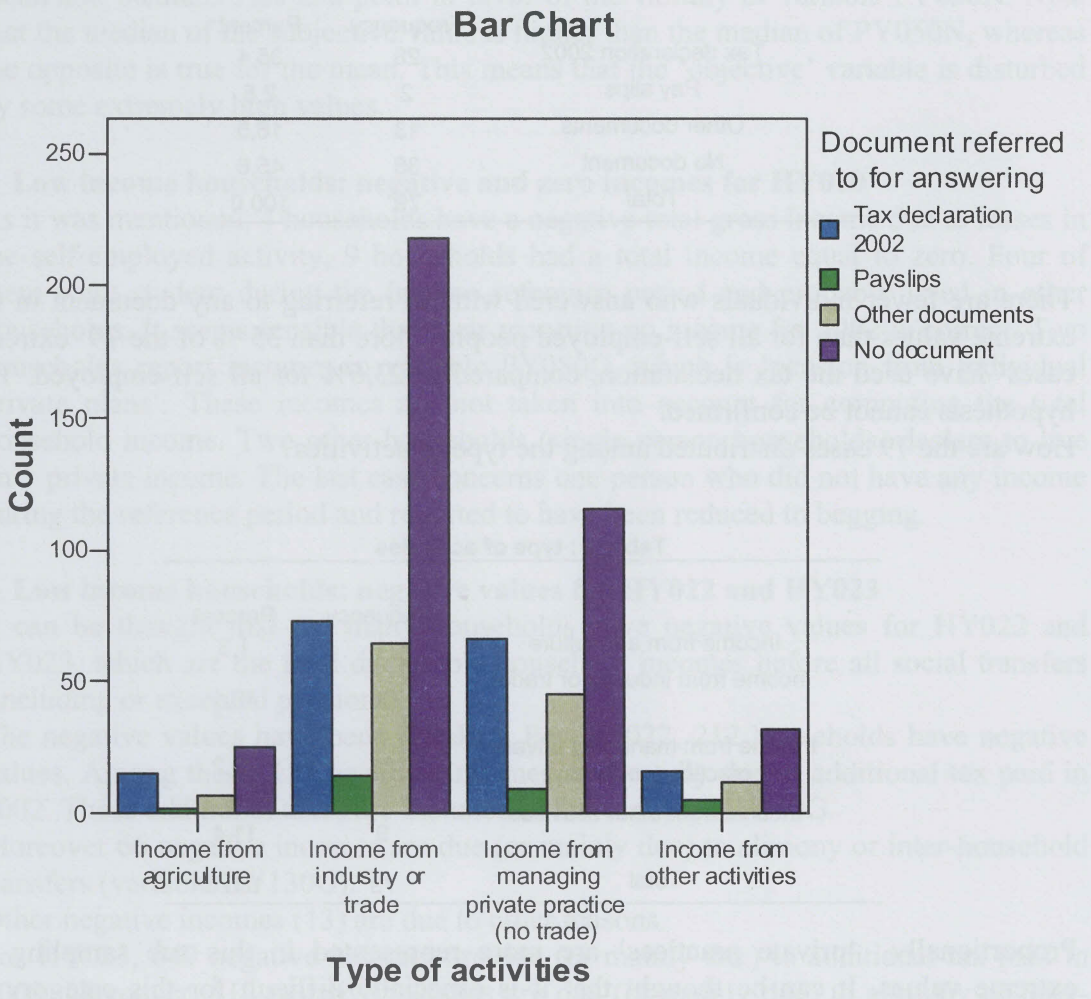
In SILC 2003 self-employed could report gains or losses: 42 persons reported losses. The target variable HY010, which is the total gross income of the household, is negative for 4 households due to self-employed losses.

♦**Control of validity: sources used to report the income**

One question concerning the document that the individual referred to when answering the income questions has been added in the Belgian questionnaire. Here is the crosstab of this variable 'document' and the variable concerning the type of self-employment activity.
Note that more than half of the respondents did not consult any document in order to provide figures about their income. This can be understood since tax declaration for example does not concern the income reference period that the question aimed for.

Table 1: Sources used to report the income by type of self-employment activity

		Tax declaration 2002	Pay slips	Other document	No document	Total /% column
Income from agriculture	Number	15	2	7	25	49
	% row	30,6	4,1	14,3	51	**6,8**
Income from industry or trade	Number	73	14	65	218	370
	% row	19,7	3,8	17,6	58,9	**51,4**
Income from private practice	Number	66	9	45	116	236
	% row	28	3,8	19,1	49,2	**32,8**
Income from other activities	Number	16	5	12	32	65
	% row	24.6	7.7	18.5	49.2	**9**
Total	Number	170	30	129	391	720
	% row	**23,6**	**4,2**	**17,9**	**54,3**	



When analyzing the crosstabs, the pattern is always the same for each sort of activities. Respondents most answered without consulting any document. The relation between both variables 'type of activities' and 'document that was referred to' is not significant (Chi-square test, p-value=0,27).
It seems, however, that people managing private practice used tax declaration more often than other categories, whereas in the industry or trade sectors they mainly answered with no document to refer to.

Using this question 'Document', one hypothesis that can also be checked is the fact that extreme absurd answers come from people who don't refer to any document.

Here follows the distribution of the variable 'documents that individuals referred to' for 79 cases that can be considered as extreme values for self-employed income. This income was lower than percentile 10 or higher than percentile 90.

Table 2 : documents that individuals referred to

	Frequency	Percent
Tax declaration 2002	28	35,4
Pay slips	2	2,5
Other documents	13	16,5
No document	36	45,6
Total	79	100,0

There are fewer individuals who answered without referring to any document in the extreme values than for all self-employed people. More than 35 % of the 79 'extreme cases' have used the tax declaration, compared to 23,6% for all self-employed. The hypothesis cannot be confirmed.
How are the 79 cases distributed among the type of activities?

Table 3 : type of activities

	Frequency	Percent
Income from agriculture	1	1,3
Income from industry or trade	38	48,1
Income from managing private practice (no trade)	31	39,2
Income from other activities	9	11,4
Total	79	100.0

Proportionally 'private practices' are more represented in this sub sampling of extreme values. It can be thought that it is especially difficult for this category to provide income values.

♦ **Control of validity: comparison to subjective value**

Only 109 individuals (among 783 for whom PY050G >0) reported a value for the subjective question added in the Belgian questionnaire: without taking into account how much did your activity bring in for your household (after deduction of all professional costs)? This question intended to verify accounting figures provided by the person. The variable that is to be compared is PY050N, the self-employed income after deduction of all taxes.
In spite of this very small sample size, it is interesting to try and interpret some results.

Table 4 : comparison to subjective value

	N	Minimum	Maximum	Mean	Std. Deviation	Median
Cash benefits or losses from self-employment (net) PY050N	109	-6000.00	229490.01	**9882.6525**	25101.87572	2000
What did your activity bring in?	109	150	50000	**7935.43**	11441.178	2500
Valid N (listwise)	109					

With the exception of some extreme cases where a loss is reported whereas the subjective answer is positive, both variables bring in fairly similar results for their mean and median. This is a point in favor of the validity of variable PY050N. Note that the median of the subjective value is higher than the median of PY050N, whereas the opposite is true for the mean. This means that the 'objective' variable is disturbed by some extremely high values.

♦ **Low income households: negative and zero incomes for HY010**

As it was mentioned, 4 households have a negative total gross income due to losses in the self-employed activity, 9 households had a total income equal to zero. Four of them were student during the income reference period and probably lived in other households. It seems sensible that their reporting no income for 2002 is correct. Two households report incomes in variable PY080G, which is 'pension from individual private plans'. These incomes are not taken into account for computing the total household income. Two other households (single-person households) declare to live on 1 private income. The last case concerns one person who did not have any income during the reference period and reported to have been reduced to begging.

♦ **Low income households: negative values for HY022 and HY023**

It can be thought that too many households have negative values for HY022 and HY023, which are the total disposable household incomes before all social transfers (including or excepted pensions).

The negative values have been checked. For HY022, 212 households have negative values. Among these, 133 negative incomes are broadly due to additional tax paid in 2002. These additional taxes are included in the variable HY140G.

Moreover 66 negative incomes are due (or mainly due) to alimony or inter-household transfers (variable HY130G).

Other negative incomes (13) are due to other reasons.

For HY023, 649 negative incomes are due (or mainly due) to additional tax paid in 2002. Moreover, 112 negative incomes are due (or mainly due) to alimony or inter-household transfers (variable HY130).

The remaining negative incomes are due to other reasons.

It has to be noted that the definition of total disposable income did not take inter-household transfers paid into account in the definition of ECHP, whereas these transfers are deducted (in addition to tax) from total gross income in order to obtain total disposable income in the SILC.

♦ **Low income households: outliers?**

The 13 negative or zero incomes were analyzed before. The positive but low incomes are now to be controlled. The limit for belonging to a low-income household is chosen equal to 6056 euro as equivalized income. This corresponds to 40% of the median of equivalized income.

5,4 % of the individuals report income below the limit. Which are the characteristics of these persons?

Half of them (50,1 %) have ever worked, whereas 62 % of the total population has ever worked.

Table 5: Person has ever worked

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	142009	34,0	49,9	49,9
	No	142408	34,1	50,1	100,0
	Total	284417	68,1	100,0	
Missing	System	133353	31,9		
Total		417770	100,0		

The next variable that is analyzed is the self-defined economic status. The problem is that the period is 'current' and does not correspond to the income reference period (2002). Nonetheless it is clear that the part of unemployed, students and nonactive persons is higher for these low-income households than in general.

Table 6 : Self-defined current economic status

		Low-income households		Population	
		Frequency	Percent	Frequency	Percent
Valid	working full time	109477	26,2	3083445	39,1
	working part time	21412	5,1	848385	10,7
	Unemployed	70454	**16,9**	717353	9,1
	Student	59231	**14,2**	746952	9,5
	in retirement	57813	13,8	1535831	19,5
	permanently disabled	7453	1,8	246098	3,1
	domestic tasks	49600	11,9	572218	7,2
	other inactive	42330	**10,1**	145006	1,8
	Total	417770	100,0	7895288	100,0

In order to compare the situations in 2002, during the income reference period, the number of months spent in full-time or part-time work, retirement, unemployment, studying, ... during the year can be compared.
Based on the variables PL070 to PL090, the following table can be constructed (note that the individuals who declared less than 12 months in one status are not taken into account (few cases) and this leads to totals not equal to 100 %)

Table 7 : Percent individuals who declared 12 months in 2002 in

	Belonging to low-income households	Distribution of total population
Working (full-time and part-time)	26,4	47
Unemployment	14,7	7,1
Retirement	**12,8**	**18,9**
Studying	16	10,1
Inactive status	25,4	12,5

The pattern is the same as in the previous table.
It is to note that retired individuals represent a smaller part of the low-income households than expected based on the distribution of total population. It also seems from the indicators that retired individuals are less poor than in the former ECHP study. This is observed in other countries (Greece and Austria) and Eurostat mentions that an explanation could be that these incomes (pensions) are collected in a more

detailed way than previously. Although private pensions are not counted in SILC, the SILC questionnaire asks more details about pension.