FINAL REPORT

"Preparation for the revision of EU-SILC :

Testing of rolling modules in EU-SILC 2017"

Contract number 07142.2015.003 – 2016.131

Statistics Belgium

MARCH 2018 – *slightly adapted for language in May 2018*

Contents

List of figures

List of tables

1. Introduction

1.1 Background information about this report

As part of the revision and modernization of EU-SILC within the new Framework Regulation on Social Statistics (IESS), some topics will be covered on a regular basis in every 3-year or 6-year EU-SILC rolling modules. In this context it has been decided to use the 2017 ad hoc module for already implementing and testing important future rolling modules variables.

The aim of this grant is to collect data according to the specifications of this new agreement in the field of health, labor, over-indebtedness, consumption and wealth in order to test the corresponding variables with the goal to have very well prepared rolling modules for the revised EU-SILC. Belgium has collected information on consumption and wealth, as it is stated in the Annex of the ESS Agreement "Allocation of the ESS agreement topics across countries".

1.2 Objectives and presentation of this report

The interim report, sent in February 2017, reported about the questionnaire construction, description of the pilot tests performed and their results and the placement of the module in the questionnaire. The purpose of this final report is to give an overview of the actions undertaken from then on, more specifically the fieldwork, data processing and analysis.

1.3 The specific objectives taken on by Statistics Belgium

Statistics Belgium volunteered to test the modules concerning consumption and wealth. In this report we will start with a discussion of the fieldwork and data processing of SILC 2017. Next, we will elaborate on the results of the quality assessment based on item non-response, and a validation with other surveys. This report concludes with some concrete advice regarding the variables, the answer modalities and the guidelines.

2. SILC 2017 fieldwork and data processing

As was already explained in the interim report, we started the work for this grant with a pilot study on the new questionnaire for the module variables. As Eurostat leaves it up to the Member States to determine the level of analyses of the majority of the variables concerned in the consumption and wealth modules, we tested on a small scale versions of the questions on both the individual and the household level. Alternatively, for some variables we also included within one level (individual or household) a condense and a more elaborate version of the questions. Based on these results, and after discussion with the relevant experts at Statbel, the questionnaire was constructed (see annex 1). Fieldwork started in March 2017 and took on until October 2017, using CAPI interviewing. In this part of the report, we will elaborate on the fieldwork, more specifically on the data collection level of the variables concerned as described in the interim report, on the final sample size and response, on the variable construction process and on respondents' assessment of the survey.

2.1 Data-collection level of the module

Below Table 1 presents an overview of the variables in both modules and the levels of question formulation. An overview of the specific wording of the questions is joined in annex 1. As can be seen all variables are asked at the household level, except food outside home and public transport. Experiences with the concrete fieldwork, including the results discussed below, confirms that this is a good level of analysis. However, we propose Eurostat to compare this with other Member States' using the other level of analysis, as to get a good comparison of descriptive statistics on information collected at the household level, and information collected at the personal level but aggregated to the household level. Based on our results we are not capable of doing that. For the two variables collected at the individual level (food outside home and public transport) this can, and will, be done below with a comparison of the HBS results.

Module on Consumption		Module on Wealth	
HC010T4 - Food at home (typical week)	Household level	HV010T4 - Value of main residence (current)	Household level
PC010T4 - Food outside home (typical week)	Individual level 16+	HV020T4 - Possession of second (more) residence(s) (current)	Household level
PC020T4 - Public transport (typical week)	Individual level 16+	HV030T4 - Possession of deposits (current)	Household level
HC040T4 - Private transport (typical week)	Household level	HV040T4 - Value of deposits (current)	Household level
HC050T4 - Regular savings (typical month)	Household level	HV050T4 - Possession of bonds, shares publicly traded or mutual funds (current)	Household level
		HV060T4 - Value of bonds, shares publicly traded or mutual funds (current)	Household level

Table 1: Overview level of the questions

2.2 Sample size

For SILC 2017 10.219 household were sampled, with a response for 6.053 households (59%). More specifically, there is a 83% response rate for the panel households and a 38% response rate for the new households. These households together constitute of 14.028 household members. For 11.352 a personal questionnaire was foreseen, of which 71 are missing because of refusal or absence and the unavailability of a proxy.

2.3 Variable construction process

This section of the report briefly describes the variable construction process for all variables in the consumption and wealth module.

Food at home:
- 1 question about the amount at household level
- Only positive values allowed
- Values higher than 10.000 euro are blocked and changed to missing

Food outside home:
- 3 filter questions (yes/no) for subcategories at the individual level
- 3 questions about the amounts for those with yes on the filter question at the individual level
- All amounts are summed up
- If the total amount is higher than 99.999 euro then value is blocked and changed to missing. This includes the extremely high values, as well as those with 'refusal' and/or 'don't know' on at least one of the three amounts

Public transport:
- 1 filter questions (yes/no) at the individual level
- Question about the amount for those with yes on the filter question at the individual level
- Amounts are blocked at 99.999 euro and changed to missing
- Zero values are as well changed to missing with flag -2[1]

Private transport:
- 1 question about the amount at household level
- Only positive values allowed
- Values higher than 999.999 euro are blocked and changed to missing

Regular savings:
- 1 question about the amount at household level
- Only positive values allowed

---

[1] After the preliminary transmission, this was changed. In the final transmission zero values were kept.

- Values higher than 999.999 euro are blocked and changed to missing

Value of main residence:
- 1 question about the amount at household level
- Only positive values allowed
- Values higher than 99.999.999 euro are blocked and changed to missing
- Additional check based on owner/tenant variable to set possible values of tenants on missing

Possession of second (or more) residences:
- 1 question at household level

Possession of deposits:
- 1 question at household level

Value of deposits:
- 1 question at household level
- Only positive values allowed
- Values higher than 999.999 euro are blocked and changed to missing
- Additional check to set possible amounts of those without deposit on missing

Possession of bonds:
- 1 question at household level

Value of bonds:
- 1 question at household level
- Only positive values allowed
- Values higher than 999.999 euro are blocked and changed to missing
- Additional check to set possible amounts of those without bonds on missing

2.4　Households' assessment of the survey

The heaviness of the module could have been measured with the duration respondents needed to answer the questions. However, we do not have this kind of information at the level of the individual module questions. Only the duration of the complete household questionnaire and individual questionnaire was recorded. On average, it took 24 minutes to complete the household questionnaire – that is 4 minutes more than for SILC 2016 and SILC 2015. In SILC 2016 there was quite a heavy module in the household questionnaire about access to services. For SILC 2017 most of the module questions were added as well in the household questionnaire. As such, the longer duration of the interview causes a higher respondent burden because of the module.

The average time for the individual questionnaire is for SILC 2017 14 minutes, while it was 12 minutes in 2016 and 10 minutes in 2015. Again, we replaced some individual level variables of SILC 2016 with those of SILC 2017, so the longer duration might as well be an indication of a

higher burden on respondents. Additionally, in SILC 2015 all except one of the module variables was added to the individual questionnaire, while in SILC 2017 only two variables were asked at that level. The difference of 4 minutes confirms again the heavy burden of the module. However, most households did not complain about the length of the interview – frequencies are in the same line as the previous years (Table 2). As can be seen the new households are slightly more negative than the panel households – although that was also the case in the previous years with other modules.

	2017	2016	2015
Too long	225 (3,7%)	189 (3,2%)	301 (5%)
	New HH: 98 (4,66%)	New HH: 73 (4,3%)	New HH: 124 (6,9%)
	Old HH: 127 (3,21%)	Old HH: 116 (2,8%)	Old HH: 177 (4,2%)
Neither too long/short	5.693 (94%)	5.623 (95,1%)	5.631 (93,7%)
	New HH: 1.949 (92,8%)	New HH: 1.579 (93,7%)	New HH: 1.666 (92%)
	Old HH: 3.744 (94,7%)	Old HH: 4.044 (95,7%)	Old HH: 3.965 (94,5%)
Too short	137 (2,26%)	96 (1,6%)	77 (1,28%)
	New HH: 54 (2,6%)	New HH: 33 (2%)	New HH: 21 (1,2%)
	Old HH: 83 (2,1%)	Old HH: 63 (1,5%)	Old HH: 56 (1.3%)

Table 2: Assessment of duration of the interview

Regarding the difficulty of answering the questionnaire in general, households also evaluated that at the end of the interview. As can be seen in Table 3, the results are comparable to the results of the previous years. Again, the new households experienced more difficulties than the old households, but this was the same in the previous years as well.

	2017	2016	2015
Very difficult	21 (0,4%)	37 (0,6%)	34 (0,57%)
	New HH: 11 (0,5%)	New HH: 10 (0,6%)	New HH: 11 (0,6%)
	Old HH: 10 (0,3%)	Old HH: 27 (0,6%)	Old HH: 23 (0,6%)
Difficult	195 (3,2%)	201 (3,4%)	210 (3,49%)
	New HH: 95 (4,5%)	New HH: 76 (4,5%)	New HH: 67 (3,7%)
	Old HH: 100 (2,5%)	Old HH: 125 (3%)	Old HH: 143 (3,4%)
Neither difficult/ nor easy	2.904 (48%)	2.559 (43,3%)	2.625 (43,68%)
	New HH: 990 (47,1%)	New HH: 720 (42,7%)	New HH: 729 (40,3%)
	Old HH: 1.914 (48,4%)	Old HH: 1.839 (43,5%)	Old HH: 1.896 (45,2%)
Easy	2.590 (27,8%)	2.663 (45%)	2.821 (46,95%)
	New HH: 895 (42,6%)	New HH: 765 (45,4%)	New HH:894 (49,4%)
	Old HH: 1.695 (42,9%)	Old HH: 1.898 (44,9%)	Old HH: 1.927 (45,9%)
Very easy	344 (5,7%)	448 (7,6%)	319 (5,31%)
	New HH: 109 (5,2%)	New HH: 115 (6,8%)	New HH: 110 (6,1%)
	Old HH: 235 (5,9%)	Old HH: 333 (7,9%)	Old HH: 209 (5%)

Table 3: Assessment of difficulty of the interview

2.5   Remarks given during the interview

Interviewers were able to add comments to all questions – both comments of the interviewer and the respondents are possible. These comments show some important points to consider for the future. This section provides an overview of these remarks.

2.5.1   Module on consumption

For '**food at home**' there were 27 of such comments, of which 10 pointed to the difference between the household size and the number of persons consuming food. On the one hand, some households indicate that they consume (some to many) meals for free at other places. Most often these are persons eating at their parents' place. On the other hand, households also indicate that they provide (some to many times) meals for free for their (grand)children or other acquaintances. As such, the amount spend to food at home does not completely reflect the consumption of food at home, but in both cases, it does reflect the amount spend on food to consume at home. In some cases, this is lower because they do not have to provide food each day, in some cases this is higher because they provide for more than the household members. There are also 7 remarks on the payment of the groceries, some respondents indicate that they do consume food at home, and that they prepare their own meals, but that other people pay for the groceries: children, parents, the church or the food bank. Last, there are some remarks indicating that the amount given is only an estimation. However, no remarks were given on the level of data collection.

From the interviewers themselves we received feedback that it is difficult for households to make the selection of their food at home based on the COICOP classification given in the guidelines. The question was formulated as "Can you tell us how much your household spends on food (some examples) and non-alcoholic beverages (some examples)? It only concerns food and non-alcoholic beverages for consumption at home." In the pilot study we tested this version (1 question) with an alternative version separating between food and non-alcoholic beverages. Most respondents favored the single question as they start from their grocery tickets in their cognitive answering process – subtracting the non-food products and alcoholic beverages. Using two separate questions increases the burden. Interviewers indicate that in a first step, it is already difficult to subtract all non-food products – as often Belgians buy both food and non-food at the same time in the same place. In a second step, alcoholic beverages should be subtracted as well, which is again not easy. The more details regarding the products included and excluded, the higher the difficulty for and burden on respondents. Some interviewers indicate that this complexity of this variable let to item non-response.

'**Food outside home**' was measured using 3 filter questions, and 3 questions asking for the amount at the individual level. Altogether, there were 32 remarks on these questions. Some remarks were small and refer to very specific situations, like difficulties to distinguish between

at home and outside home for people owning a restaurant, people being able to eat without charges at other places (cf. above for food at home). However, 11 persons commented on the reference period, having difficulties reporting an amount for a typical week, as they only go to restaurants on for example a monthly basis. 6 additional comments referred to restaurant visits and delivery meals at the household level – however, they indicated that they split up the amount among the household members.

'**Public transport**' was also measured at the individual level with one filter question and a follow-up question on the amount. We received 74 remarks for these two questions. The grant majority of them were respondents explaining their employer pays their public transport, that they have only a minimum amount because of social reductions (e.g. students, persons in retirement), or how they organize the payment of public transport (e.g. with a subscription, a 10-ride card). However, some respondents indicated as well that the reference period of a typical week was not easy (cf. above for food outside home) because some payments occur only annually (subscription) or once in a few months (10-ride card). As it is clearly clarified in the guidelines that in such cases an average should be taken, the remarks of the respondents reflect the burden of having to calculate and recalculate during the interview This also increases the chance that respondent will do a wild guess.

The next variable, '**private transport**', was collected at the household level. After the questions about material deprivation (e.g. car, bike, motorcycle), the household respondents were asked about the cost of private transport. 82 household respondents gave us an additional remark. For more than 15 of them it referred to the details of the calculations (what costs were taken into account) – showing that it is a high burden and requires much of both the respondent and the interviewer to sum everything up, often on an annual basis, and then divide it into 52. Additionally, more than 25 of them explained their low values as they have a company car, or they have vehicles but do not use them for one of the other reason, or some parts of the costs are covered by others. Other respondents indicated that it was a difficult question – some of them clarified that it is difficult to answer at the household level because they are not familiar with costs of other household members

Again, we received the remark of difficult reference period, and the difficulties of averaging everything out. There were also 5 respondents indicating that it is difficult for them to distinguish between private and work-related use of the vehicle, for example self-employed – and the cost they carry as a private person versus as a business. Additionally, some persons also point to the fact that they do not have to pay for the use of the company car, but that at the end, they do need to pay some taxes. At least, there are some respondents referring to the shared use of a car (cf. sharing economy). The questions is however if this is public or private transport.

The last variable in the consumption module is '**regular savings**', this was asked with a single question at the household level. In total, we received 22 remarks for this question. Most of them refer to the details of the calculations, or the indication that they are unable to save on a regular basis. Other remarks clarify that the amount saved really depends from month to month, or that it is useless to save.

Overall, interviewers indicate that the reference period of a typical week is extremely difficult for the respondents. It requires a lot of calculations on an annual basis – as in one single question, a lot of expenses should be considered – and afterwards, is should be recalculated on a weekly basis. Interviewers indicate that at first respondents are willing to do that, but as there are many questions requiring many expenses to consider, their motivation declines to use a calculator and really take every expense into account. This results either in an quick estimation, or in item-non response (cf. below).

2.5.2 Module on wealth

The first variable of the module Wealth is '**value of the main residence**'. This was asked with one single question. 23 households added a remark to these questions, of which 6 refer to a description of the residence as to provide more detailed information. Another 5 households indicate that they are afraid to answer, because they really do not have a clue. At last there is 1 household that clarifies that the house is currently under renovation, and thus that an estimation of the current state does not reflect what it would be after the renovation. Only 1 household explicitly says that the question is too sensitive to answer.

For '**possession of second (or more) residence(s)**' again a single question was asked. We received only 6 comments, mostly referring to more detailed information such as the type of property or the location.

The next variable in the Wealth module is '**possession of deposits**', where no remarks were given. However, we obtained 53 remarks on the variable '**value of deposits**' that was collected at the household level. Where the remarks were kind of friendly for the previous variables discussed, this is not really the case here. 14 just indicate 'no answer', 3 indicate their refusal to give the amount, while 3 others explicitly point to the sensitivity of the question. Others indicate that they do not know because other persons oversee the household's financial management. One remark comes from the interviewer expressing his/her doubts about the respondent's answer. There are still two important types of comments to consider for the future. First, 8 households indicate having a negative value on their accounts, however, this is not allowed for the variable. Second, 4 households indicate explicitly that they did not take specific household members into account because they do not know their value. A specific case here are the savings of children of divorced parents, where one of the parents indicates that the savings managed by the other partner are not taken into account.

'**Possession of bonds**', however, poses of all variables the most problems regarding the sensitivity of the question. Of the 7 comments, 3 explicitly indicate that the question is intrusive, 1 indicates 'enough' and another one indicates 'more than ….. (specific amount filled in)'. Already at the question of 'do you poses bonds', people think about the amount and experience the question to be sensitive. With the variable 'value of bonds', 2 additional respondents point to the intrusiveness of the question. The other 6 remarks refer to values already given in previous questions or the indication that the amount is only an estimation.

3. Item non-response analysis

A first step in the analysis of the variables consists of assessing the item non-response. This part of the report describes on a high level of detail the item non-response of the module variables. At the level of the households, we will work with 6.053 households. At the individual level, we will work with 11.281 individuals when discussing the level of the questions, and 11.352 when discussing the level of the variables. For latter the 71 partial missings are included, while this is not the case for the former. To obtain the overall overview, variables are clustered below within their module.

3.1 Item non-response errors in the module on consumption

First, the variables of the module on consumption are discussed. Where necessary, filter questions are treated separately – to provide as much information as possible.

3.1.1 Food at home
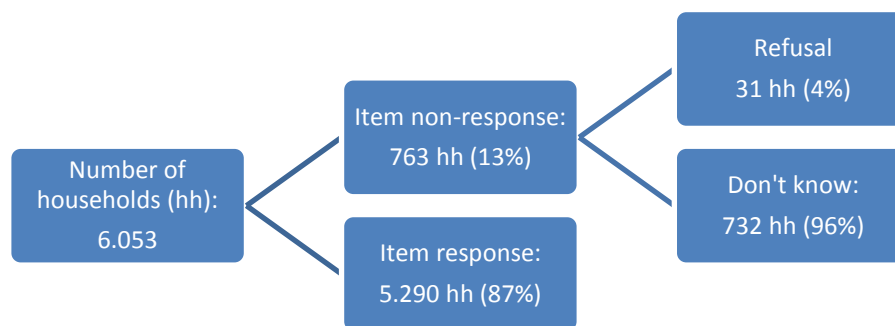
Figure 1: Item non-response food at home

For food at home, 13% of the households do not answer how much they spend in a typical week (Figure 1). Most them (96%) are missing because they don't know the amount of their weekly expenses, 4% are missing because they refuse to answer. Based on this low percentage of refusals (in relation to the percentage of 'don't' know') we believe the information is not too sensitive to ask about. However, 13% item non-response is extremely high, and much higher than what Belgium is used to for household level variables. As we don't believe this information to be too sensitive, the high level of item non-response will probably be caused by the variable's complexity, as was already pointed to above (cf. remarks given during the interview). In the pilot study the current formulation of the questionnaire was the simplest one; the alternative was to ask for food and non-alcoholic beverages separately. Even at the pilot study respondents indicated that the questions were very difficult. We can indeed not expect to get completely reliable information with one single question where HBS sets up a complex design to obtain similar information.

15

3.1.2 Food outside home

Variable 'Food outside home' is operationalized with several questions ('sub variables') in the Belgian SILC questionnaire:

- Eating outside home (yes/no + amount if yes),
- Take away or delivery meals (yes/no + amount if yes),
- Beverage outside home (yes/no + amount if yes).

It is important to analyze the item non-response for these 'sub variables' to understand how we should determine the item non-response for the variable 'Food outside home'.

Figure 2: Item non-response food outside home – eating outside home

Most of the respondents (41%+57%=98%) answer (yes/no) on the question about eating outside home in a typical week (Figure 2). 2% does not give a (yes/no)-answer. The item non-response consists mainly of people don't knowing the answers (96%), the others (4%) refuse to answer whether they eat outside home in a typical week. It only concerns a minority of the respondents, but it seems strange not to know whether you eat outside home in a typical week or not; it might be that this typical week is too difficult to conceptualize – as eating outside home might either be a monthly (or two-monthly) thing than a weekly thing. The SILC 065 document (2017 operation) indicates that in case of difficulties with a typical week, then the first week before the end of the reference period should be chosen. This is however, no solution in case of a monthly event.

16
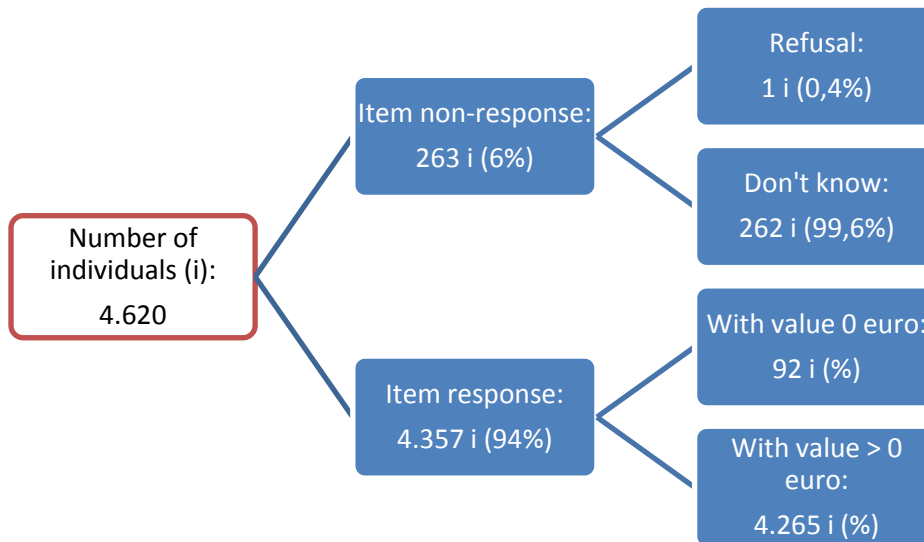
Figure 3: Item non-response food outside home – spending on eating outside home

6% of the individuals who mentioned to eat outside home in a typical week do not mention an amount (Figure 3). Only one refuses to give up an amount, all others don't know. Interviewers indicate that the don't knows should also be interpreted as don't knowing the answer by heart and refusing to do the calculations – as was already explained above. Taking both questions (filter and follow-up) together, there are 537 households with item non-response (9%).
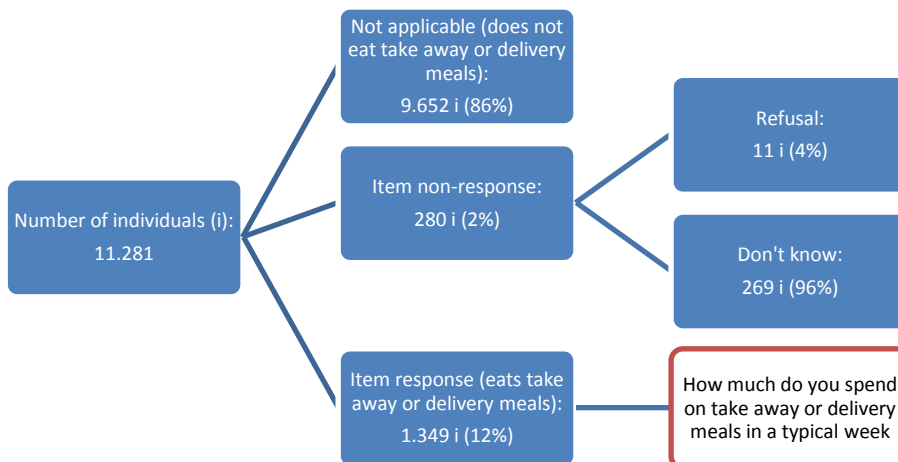
Figure 4: Item non-response food outside home – delivery meals

Most of the respondents (12%+86%=98%) answer (yes/no) on the question about eating take away or delivery meals in a typical week (Figure 4). 2% does not give a (yes/no)-answer. The item non-response consists mainly of people don't knowing the answer (96%), the others (4%) refuse to report whether they eat take away or delivery meals in a typical week. Again, it seems strange not to know whether you eat take away or delivery meals in a typical week or not.
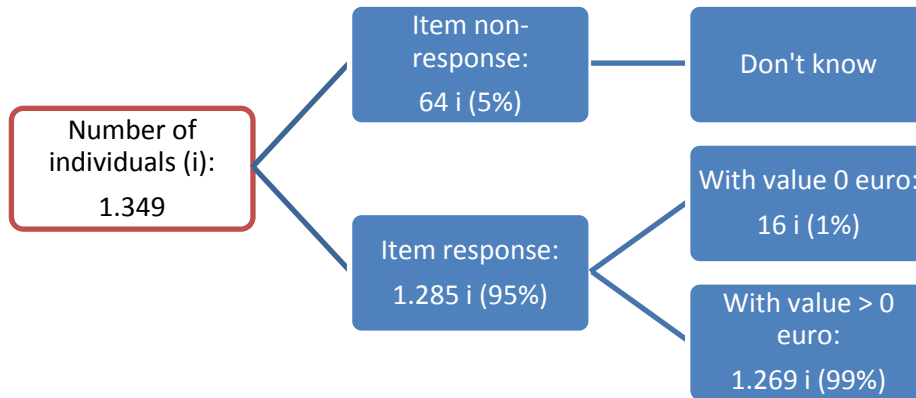
Figure 5: Item non-response food outside home – spending delivery meals

5% of the individuals who mentioned to eat take away or delivery meals in a typical week refuse to give up an amount (Figure 5). 1% (16 individuals) of the individuals mentioning to eat take away or delivery meals in a typical week report to spend €0 on those meals in a typical week. This will be further investigated in the chapter on data-analysis.
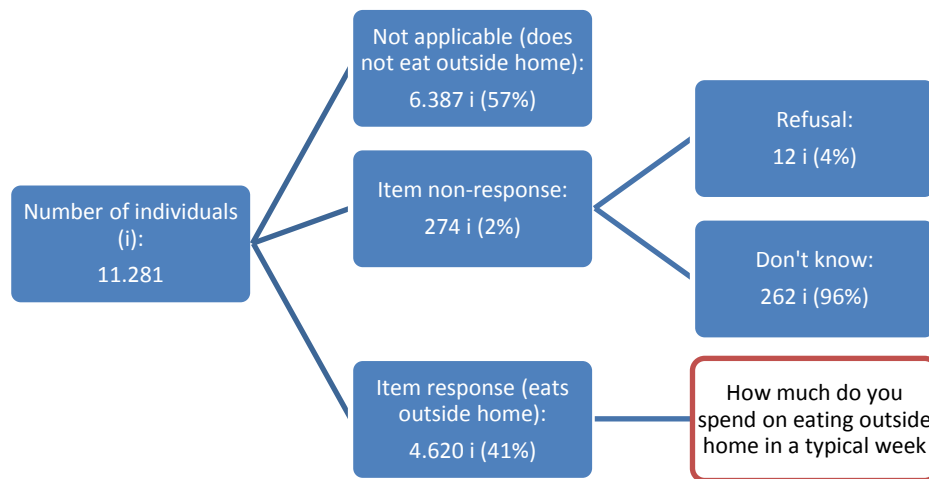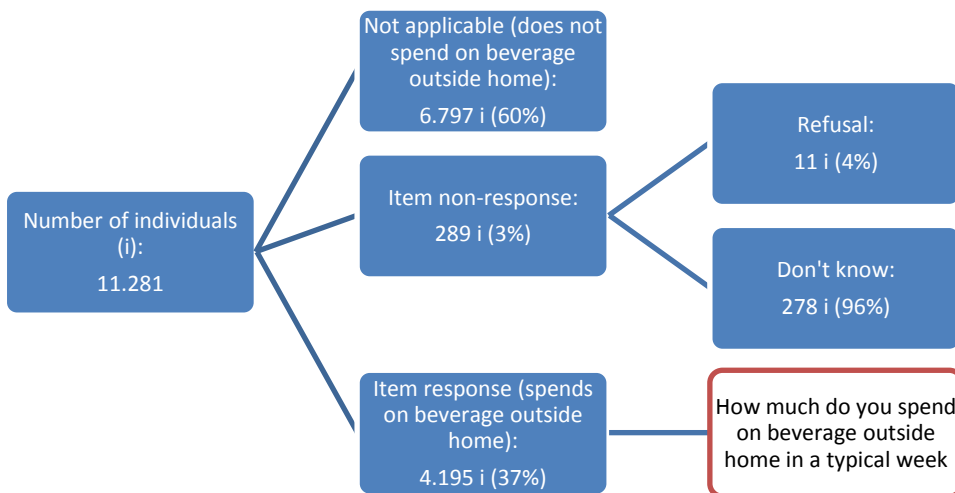

Figure 6: Item non-response food outside home – beverages

Most of the respondents (37%+60%=97%) answer (yes/no) on the question about spending on beverages outside home in a typical week (Figure 6). 3% does not give a (yes/no)-answer. The item non-response consists mainly of people not knowing the answer (96%), the others (4%) refuse to report whether they spend on beverages outside home in a typical week.
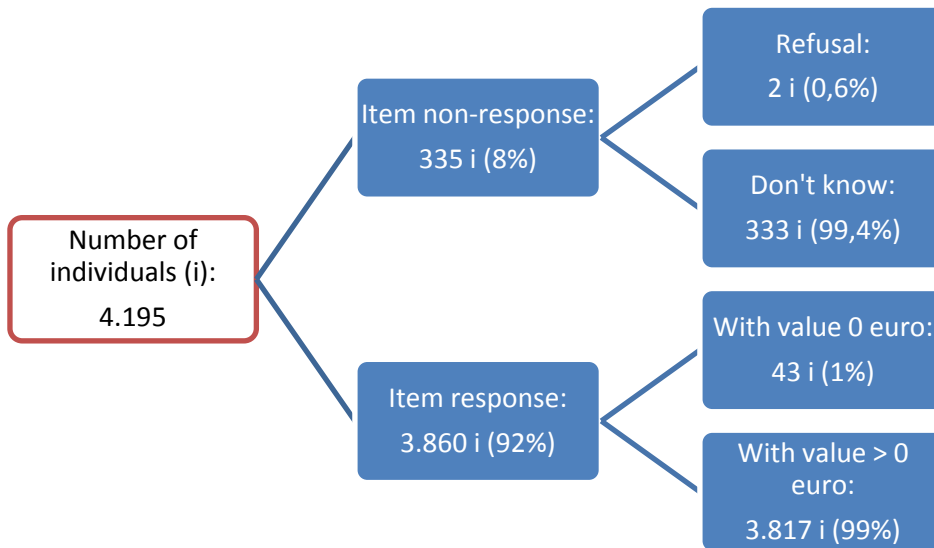
Figure 7: Item non-response food outside home – spending beverages

8% of the individuals who mentioned to spend on beverages outside home in a typical week do not mention an amount (Figure 7). The majority of that group (99,4%) does not know the amount. Together with the filter questions, this brings the item non-response to 624 respondents (10%), again quite high. 1% (43 individuals) of the individuals mentioning to spend on beverages outside home in a typical week report to spend €0 on those beverages in a typical week. This will be further investigated in the chapter on data-analysis.

Now that we looked at the item non-response for each of the 'sub-variables', we will determine the item non-response of the variable 'Food outside home' (Figure 8). All amounts were summed up. Respondents with a refusal or 'don't know' answer on one of these amounts got the -1 flag – even if they provided us with an amount for one or both of the other subcategories. This implies an extremely high non-response rate (flag -1) of 48%.
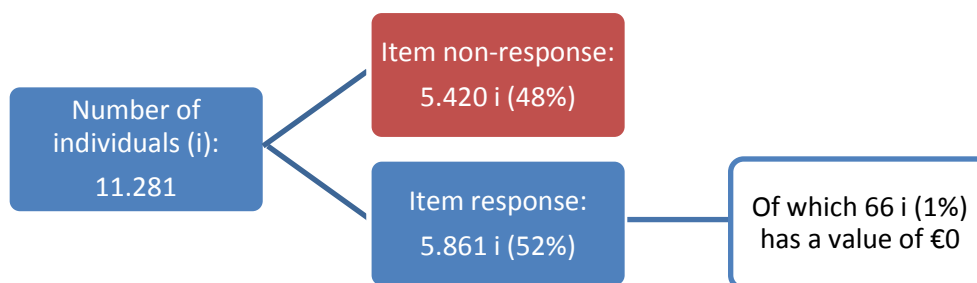
Figure 8: Item non-response food outside home

Alternatively, we could attribute an amount of '0' to all persons answering 'don't know' (but not to the refusals). Doing this results in a mean of 35,69 euro a week, originating from a response rate of 58%. Still, there is a very high non-response rate. Close analysis of respondents with item non-response reveals more than 4.500 respondents indicating that they do not consume food outside home, do not consume take away or delivery meals, and do not consume drinks outside home. In fact, this variable is not applicable for them. In the case where we would use a new flag for them, the results would look as follows:

- Not applicable (does not consume outside home): 4.685 (42%)
- Item non-response (refusal and don't know): 735 (7%)
- Item response: 5.861 (52%)

Including the 71 partial missings would give the results as presented in Figure 9. We believe that this 7% non-response is a more correct approach for determining the item non-response of the variable 'Food outside home', even though it is still high – which might be caused by the difficulties with the reference period discussed above.
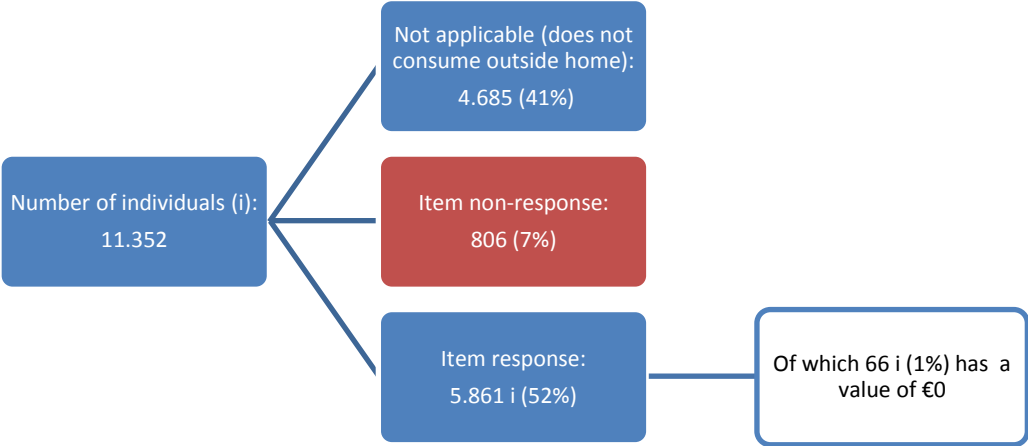
Figure 9: Item non-response food outside home – with additional flag

3.1.3 Public transport

To obtain the variable 'Public transport' a yes/no question was asked, and for the ones responding 'yes' we asked about the amount of costs in a typical week. Again, it is important to analyze the item non-response for the 'sub variables' to understand how we should determine the item non-response for the variable 'Public transport'.
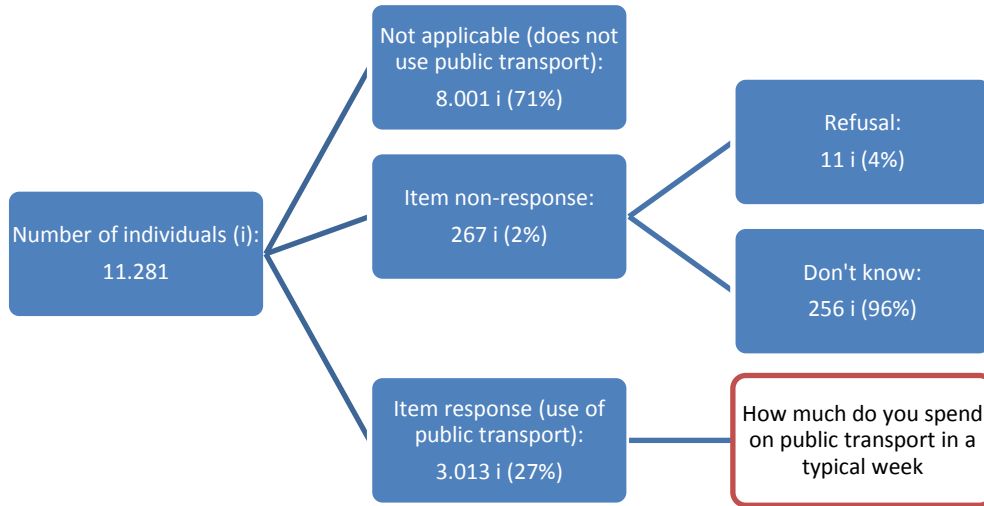
Figure 10: Item non-response public transport – use

Most of the respondents (27%+71%=98%) answer (yes/no) on the question about using public transport in a typical week (Figure 10). 2% does not give a (yes/no)-answer. The item non-response consists mainly of people not knowing the answer (96%), the others (4%) refuse to report whether they spend on public transport in a typical week.
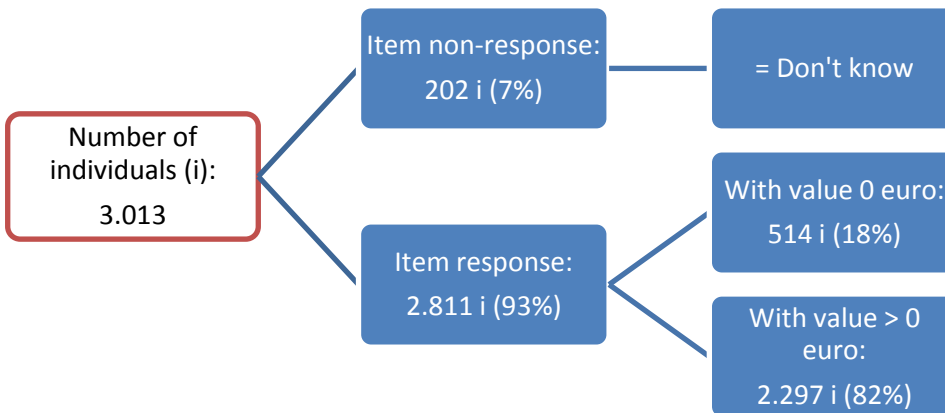

Figure 11: Item non-response public transport – expenses

7% of the individuals who mentioned to spend on public transport in a typical week refuse to mention an amount (Figure 11). 18% of the individuals that use public transport in a typical week report to spend €0. These are especially those respondents whose employer pays for public transport.

Now that we looked at the item non-response for the 'sub-variables' we will determine the item non-response of the variable 'Public transport'. A first approach of the item non-response would imply the sum of the individuals with item non-response in the sub-variables (Figure 12).
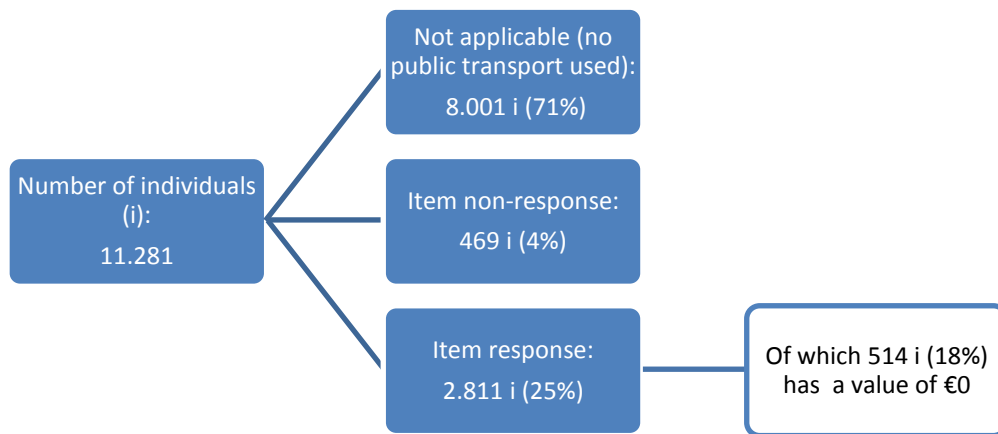
Figure 12: Item non-response public transport with no expenditure

However, when constructing the variable 'Public transport' we had to exclude the 514 users of public transport mentioning €0 (no expenditure) – following the guidelines of SILC 065, and including the 71 partial missings, gives us the result as presented in Figure 13.
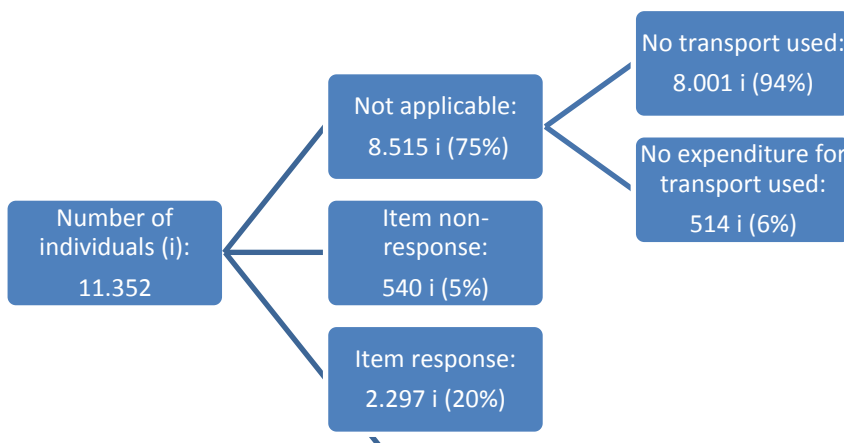
Figure 13: Item non-response public transport without no expenditure

Item non-response (5%) does not differ between the two approaches (zero values in- or excluded). But we would advise to look deeper in on this for the construction. The guidelines do include 'no expenditure' as a possible value, but no code is given, and amounts go from 1 to 999.999,99; while for private transport 'no expenditure' is also possible, but the value range starts at 0.

3.1.4 Private transport
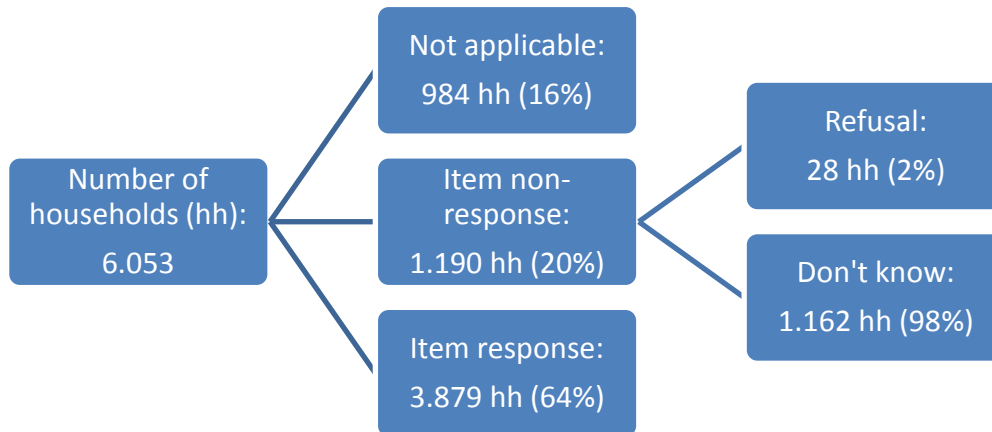
Figure 14: Item non-response private transport – use

The use of private transport does not apply to 16% of the households in the survey (Figure 14). We will look at the item non-response for the 5.069 households where it is applicable (Figure 15).
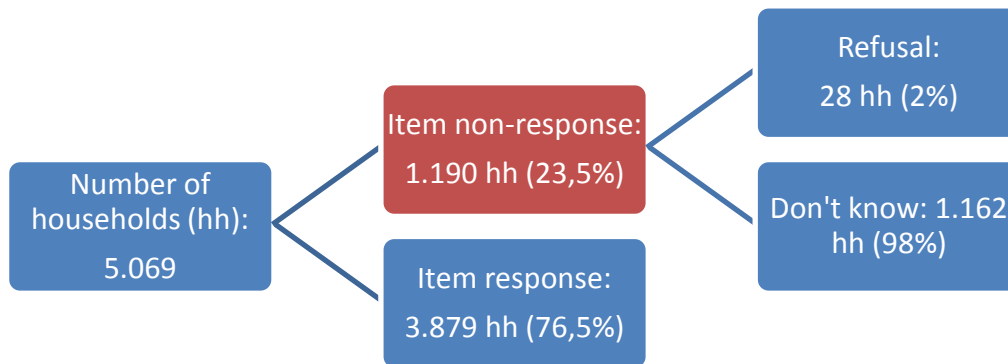
Figure 15: Item non-response private transport

Almost a quarter of the households (23,5%) does not have an answer on the private transport costs in a typical week. 2% of them because they refuse and 98% because they don't know. We have feedback from the fieldwork that the refusals often concern respondents refusing to do the calculations, while in fact they have the information there. This confirms again the complexity of this specific variable. Which such a high non-response the reliability of this variable is questionable.

3.1.5 Regular savings
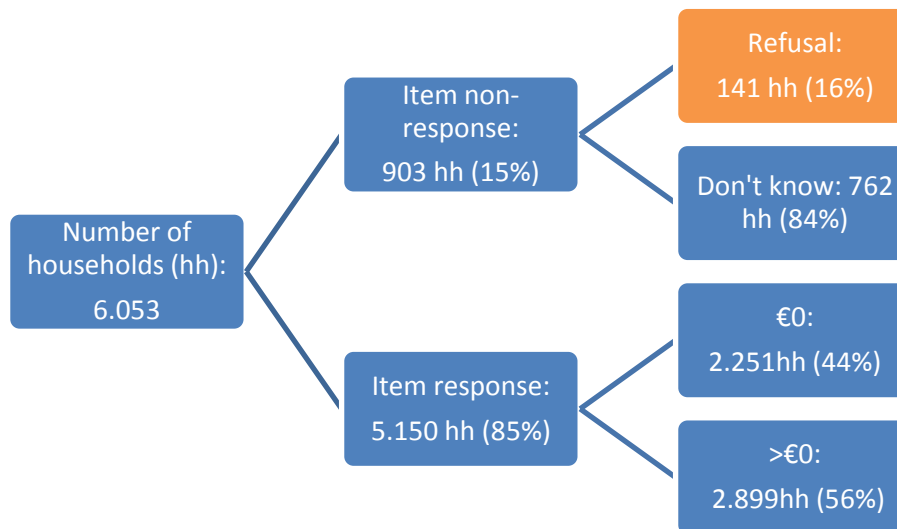
Figure 16: Item non-response regular savings

15% of the households do not answer how much the household saves in a typical month (Figure 16). 84% of them because they don't know, 16% of them because they refuse. 85% of the households respond to the question on how much the household saves in a typical month. 43% of them save €0. Some characteristics of the responding households are looked at in the chapter on data validation. Again, the item non-response is extremely high – jeopardizing data quality. Additionally, the SILC 065 indicates values starting at 1, but also allows 'no savings'. This is confusing and should be clarified.

3.1.6 Conclusion for module on consumption

As becomes clear the item non-response for the variables in the consumption module are quite high, and for regular savings, unacceptable high. Each variable on its own is very complex for respondents. The totality of all difficult and complex questions together in this module places an unacceptable high burden on respondents. Data quality shows that it is suffering.

3.2 Item non-response errors in the Module on wealth

In second place, the variables of the module on wealth are discussed. Where necessary, filter questions are treated separately – to provide as much information as possible.

3.2.1 Value of main residence
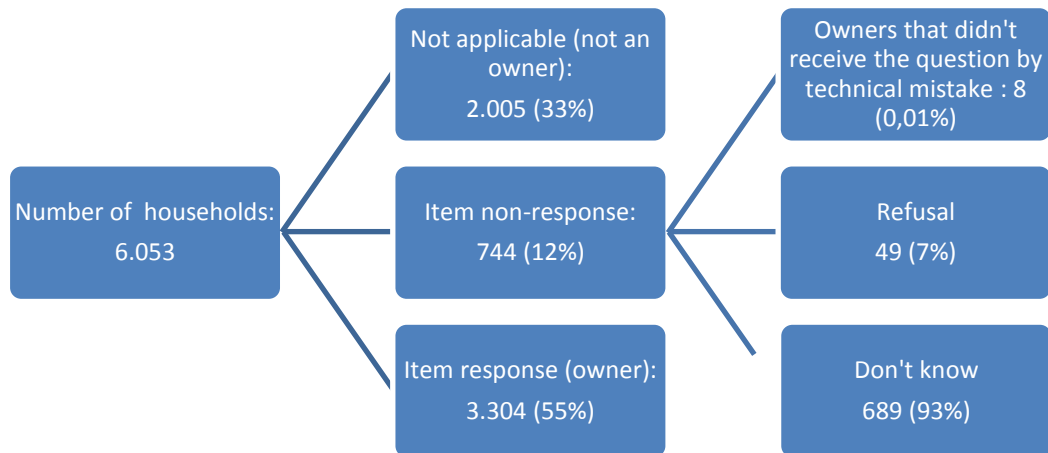
Figure 17: Item non-response value of main residence

The value of the main residence does not apply to 33% of the households because they are not owner of the main residence (Figure 17). We will look at the item non-response for the 4.048 households where it is applicable (Figure 18).
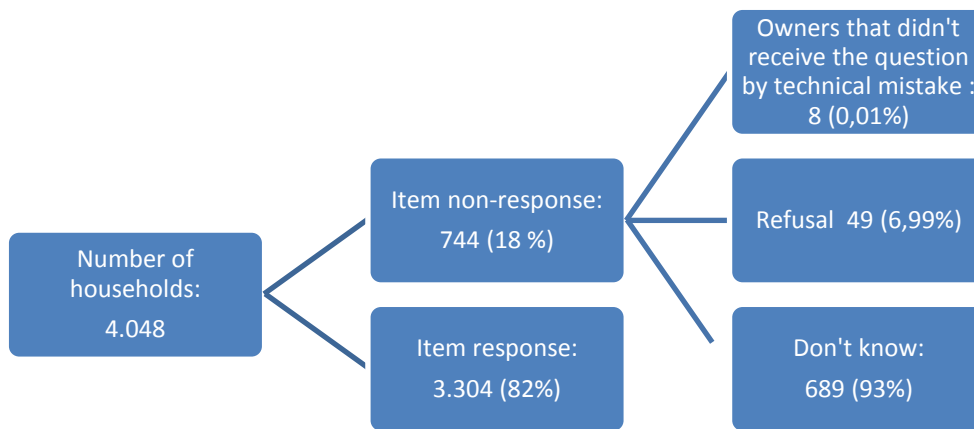
Figure 18: Item non-response mail residence for owners

18% of the households do not give the value of the main residence; 93% of them because they don't know, and the minority (7%) of them because they refuse. There are 5 households mentioning a value of €0 and some households mention very low amounts. This problem will be further analyzed in the chapter on data validation.

3.2.2 Possession of second (more) residence(s)
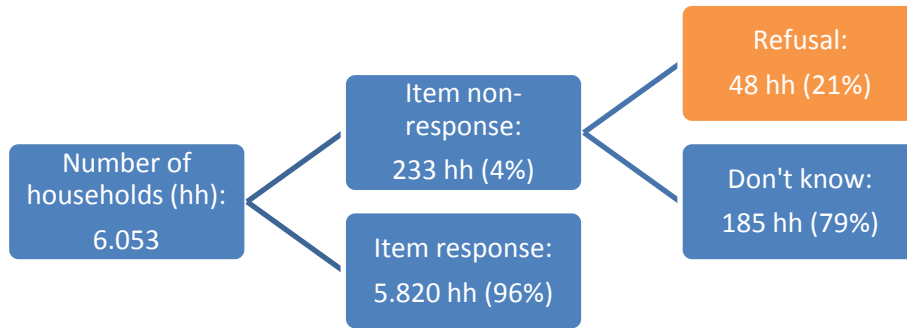
Figure 19: Item non-response possession of second residence

There is an item non-response of 4% on the possession of second (more) residence(s) on household level (Figure 19). Of those 233 households there are 185 households that do not know whether their household possess a second (more) residence(s) and 48 that refuse to answer to the question.

3.2.3 Possession of deposits
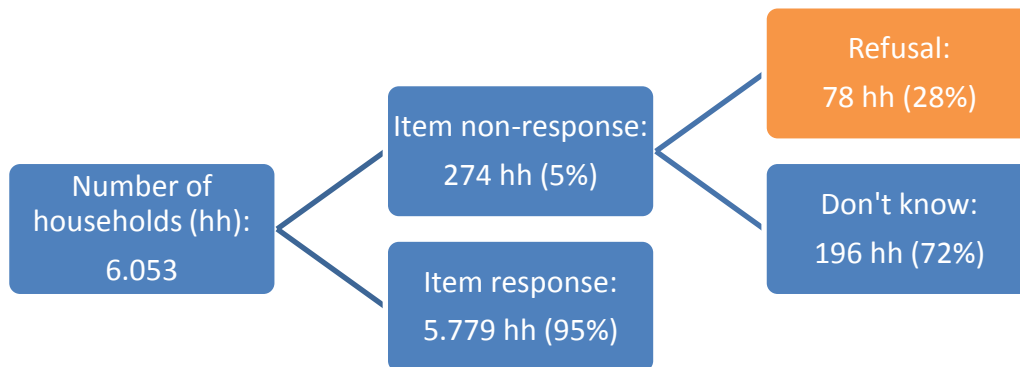
Figure 20: Item non-response possession of deposits

There is a 5% item non-response on the possession of deposits on household level (Figure 20). Of those 274 households there are 196 households that do not know whether they possess deposits and 78 that refuse to answer. Again, it points to the difficulty of the household level, as was already discussed above. Not all household respondents are fully informed about all other household members.

3.2.4 Value of deposits
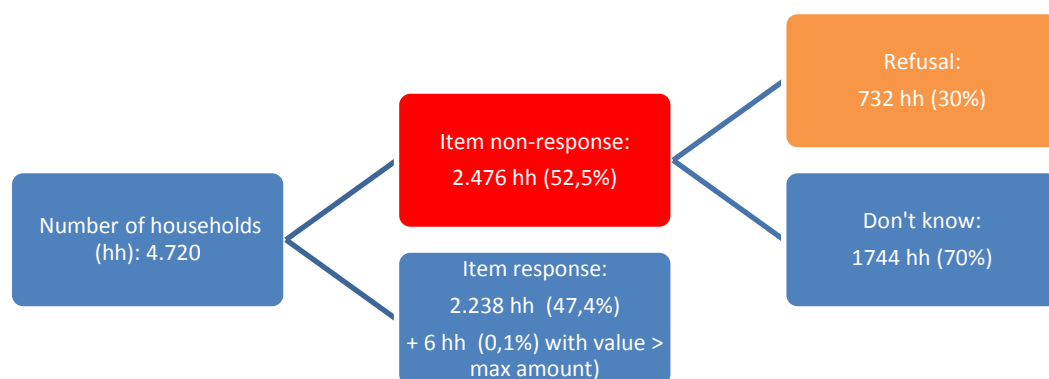
Figure 21: Item non-response value of deposits

4.720 (82%) of the 5.779 households mentioned to possess deposits and were asked about the total amount on the deposits on household level (Figure 21). When the value of the deposits is asked we get a very high item non-response: 52,5% (70% of those non-respondents don't know the answer and 30% refuses to answer). Taking the non-respondents from the 'possession of deposits' into account, there is a non-response for 2.750 households (45%). This is extremely high and jeopardizes data quality fundamentally. From the remarks above it was already clear that not all household respondents all fully informed about their household members, but also – and more importantly – that this is both a difficult and sensitive question. All valid explanations for the high item non-response.

	Households with value of deposits > €999.999,99					
	HH1	**HH2**	**HH3**	**HH4**	**HH5**	**HH6**
Value of main residence (in €)	300.000	2.500.000	Not applicable (renter)	500.000	200.000	275.000
Possession of 2nd/more residence(s)	yes	no	no	yes	yes	yes
Values of deposits (in €)	1.000.000	1.200.000	1.800.000	2.000.000	2.000.000	3.000.000
Values of bonds, etc (in €)	Not applicable (no bonds, etc)	Not applicable (no bonds, etc)	Not applicable (no bonds, etc)	Not applicable (no bonds, etc)	1.200.000	10.000

Table 4: Analysis of high values of deposits

27

It should be noted that 6 households in our survey mention an amount higher than accepted following the Manuel 'DocSILC065 (2017 operations)' where the value of deposits must be between €0 – 999999.99. That is why we did not include these households in the current analyses. When analyzing some characteristics of these 6 households we believe the values seem plausible. We looked at the household type, and for each member of the households at the ages, education level and self-defined current economic status. At household level they all declare to make ends meet (fairly or very) easily, have the capacity to face unexpected financial expenses, to afford paying for one week annual holiday away from home, to afford a meal with meat, chicken, fish (or vegetarian equivalent) every second day, to have a telephone, color TV, computer, washing machine, car and they have no arrears (HS011, HS021, HS031). We also compared the value of deposits with the other variables in the module on wealth (Table 4). As we believe in the plausibility of these values, we think it is a loss of information to exclude these households. Why is the upper limit set on €999999,99? Are they considered as outliers that would deviate the mean (and standard deviation) too much? Than why not choose to keep but top off the value at €999.999,99? It can also be chosen not to use an upper limit. Of course, the outliers have to be checked (more than other variables) on their plausibility.

3.2.5 Possession of bonds, shares publicly traded or mutual funds
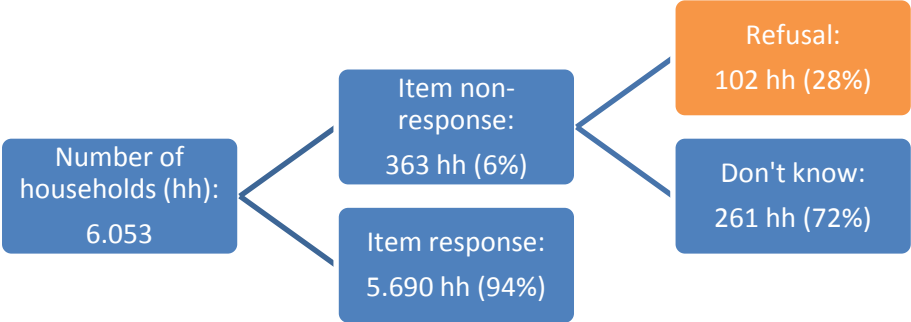
Figure 22: Item non-response possession of bonds

There is a 6% item non-response on the possession of bonds, shares publicly traded or mutual funds on household level (Figure 22). Of those 363 households there are 261 that do not know whether they possess that and 102 that refuse to answer.

3.2.6 Value of bonds, shares publicly traded or mutual funds
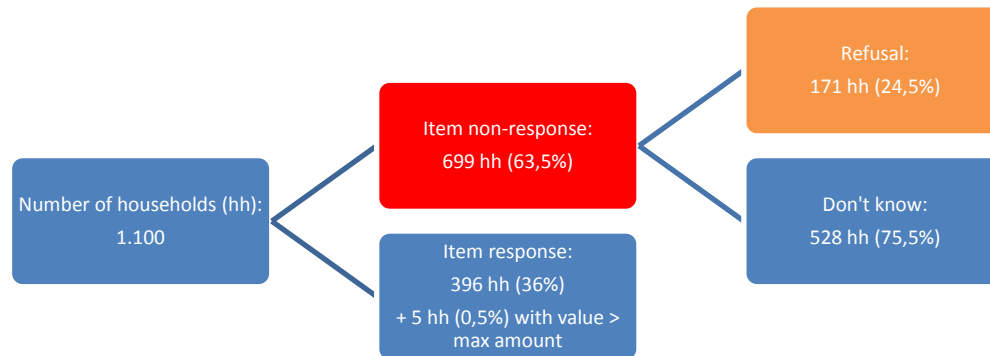
Figure 23: Item non-response value of bonds

1.100 (18% of the 5.690) households mentioned to possess bonds, shares publicly traded or mutual funds. They were asked about the total amount bonds, shares publicly traded or mutual funds on household level. When the value of that is asked we get a very high item non-response: 63,5% (Figure 23). 75,5% of those non-respondents don't know the answer and 24,5% refuses to answer. Again, this shows that the question is both too difficult and too sensitive. What is the value of a variable where only 396 households have a valid response – not even sure it is reliable? This again jeopardizes data quality and reliability.

It should be noted that 5 households in the survey mention an amount higher than accepted following the Manuel 'DocSILC065 (2017 operations)' where the value of bonds, shares publicly traded or mutual funds must be between €0 – 999999.99. That is why we did not include these households in the current analysis. When analyzing some characteristics of these 5 households we believe the values do seem plausible. We looked at the household type, and for each member the ages, education level and self-defined current economic status. At household level they all declare to make ends meet fairly easily or very easily, have the capacity to face unexpected financial expenses, to have a telephone, color TV, computer, washing machine, car (or do not have because they do not want to) and they have no arrears (HS011, HS021, HS031). For one household with a value of €1.000.000 for bonds, €110.000 for deposits, €425.000 for main residence and saving €140 in a typical month, it is mentioned that they have no ability to keep the home adequately warm in winter, no capacity to pay for one week annual holiday away from home and to afford a meal with meat, chicken, fish (or vegetarian equivalent) every second day (which we would believe is not plausible in comparison with their other answers). But the other households do have these abilities and capacities. We also compared the value of bonds, shares publicly traded or mutual funds with the other variables in the module on wealth (Table 5).

	Households with value of bonds, etc > €999.999,99				
	HH1	HH2	HH3	HH4	HH5
Value of main residence (in €)	425.000	450.000	200.000	400.000	650.000
Possession of 2nd/more residence(s)	no	no	yes	yes	no
Values of deposits (in €)	110.000	8.000	2.000.000	20.000	100.000
Values of bonds, etc (in €)	1.000.000	1.200.000	1.200.000	1.200.000	1.500.000

Table 5: Analysis of high values of bonds

As we believe in the plausibility of these values, we think it is a loss of information to exclude these households. Why is the upper limit set on €999999,99? Are they considered as outliers that would deviate the mean (and standard deviation) too much? Than why not choose to keep but top off the value at €999.999,99? It can also be chosen not to use an upper limit. Of course, the outliers must be checked (more than other variables) on their plausibility.

3.2.7 Conclusion for module on wealth

Overall, the item non-response for the module on wealth is higher than used to for household variables in SILC on the one hand, and the consumption module on the other hand. This is a clear indication that the wealth module poses too many problems. Both for the respondents – it is a heavy burden to answer, the questions are difficult and extremely sensitive – as well as at the data quality level – there are too many refusals, and answers given seem not always reliable.

4. Data validation

The previous chapters of this report discussed the problems encountered during field work and problems with data quality caused by item non-response. The fact that Belgium organizes HBS each two years gives a unique possibility for comparison, as there is only 1 year in between both. This chapter discusses the descriptive statistics of the module variables, and the comparison between SILC on the one hand and on the other hand HBS and HFCS. First, the three data sources are briefly explained.

4.1   Main features of SILC, HBS and HFCS

The validity of the module data is checked by comparing the data with the household budget survey (HBS)[2] as expert on consumption, and the Household Budget and Consumption Survey (HFCS)[3] as expert on wealth. In annex 2 some survey characteristics are listed to compare across the three sources.

HBS and HFCS present results at the household level. As was discussed in 1.1 we asked all module variables at the household level, except food outside home and public transport. Those two variables have to be aggregated from individual to household level for comparison reasons, which has some limitations in data analysis. At individual level we collect for each current household member aged 16 and over, whereas at household level we collect for all household members. When aggregating the individual level (only 16+) to the household level we miss information on the children below 16 year old. This jeopardizes the comparison with HBS and HFCS for those two variables.

4.2   Module on consumption

As a first step for each variable in the consumption module, the descriptive statistics are calculated, with and without weights to get both an insight into the data as well as the generalizations to the population. In a second step, this is checked with the external information from HBS.

4.2.1   Food at home

Food at home has at the unweighted household level a mean of 125 euro and a median of 100 euro in a typical week, as is the mode. The range is extremely high: 10.000 euro, as the lowest value is 0 euro, and the highest value is 10.000 euro. As such, this results in a standard deviation of 167 euro. The interquartile range is 80 – which seems plausible, taken all different household types and sizes into account. The guidelines allow households mentioning €0 as a value for food at home, as is the case for 6 households. Although we do not believe these

---

2   Organized by Statbel: https://statbel.fgov.be/en/themes/housing/household-budget-survey-hbs
3   Organized by the National bank: https://www.nbb.be/en/publications-and-research/study-financial-behavior-households-household-finance-and-consumption

values to be plausible, they were kept for the analysis. Even for persons consuming all days of the week outside home (for free or with expenses), we believe that there is at least a minimum of food and beverages to consume at home. We looked at some general characteristics of these households (household type, material deprivation, HS120, HH021) and its members (age, educational status, economic status, region), as well as at their values for the other variables on consumption and on wealth. We were unable to find a specific household profile. Still, we believe that these values are not correct, and should in fact be missing. There are a few other households who also mention low amounts (between €1 and €10). They were analyzed as well, but again no specific household profile could be found. The same goes for the other end of the distribution, where there are households with extremely high weekly expenses, approximately 1% of the households reports values of 400 euros or higher.

In Table 6 the weighted results are presented, and compared with HBS results; zero values are excluded to make the comparison with HBS as good as possible. The median and mean values differ substantially between both sources. The higher values in SILC are most likely the result of an overestimation when asking for a global amount for a typical week. However, since we advise using a single questions for almost all variables in the module as the only way to keep the module acceptable for respondents (cf. pilot reported in interim report), this is the best you can get with this approach. We believe that SILC is not a good source for measuring consumption for food at home.

	SILC	HBS
Variable / description	HC010T4 Food at home	Expenses for COICOP01: Food and non-alcoholic beverages
Reference year	2017	2016
Reference period	A typical week	Year divided by 52
Median	€100	€75,5
Mean	€124	€88,2
Base	All hh, N=5.284	All hh mentioning a expense for COICOP01, N=4.490
Weighted	yes	yes

Table 6: Comparison with HBS for food at home

As already discussed above, feedback from the fieldwork confirms that respondents have a hard time estimating expenses for consumption of food and non-alcoholic beverages. It is likely that they take their costs from the grocery, bakery, butcher, supermarket, etc. where alcoholic beverages as well as non-food consumption are also included. Also, weekly amounts may be estimated higher than monthly or annual. The distribution of the values also shows that respondents estimate their weekly expenses with round numbers (e.g. 80, 100, 150, …), which might also lead to an overestimation. The approach in the Household budget survey where a diary is kept for the expenses in the different classifications of individual consumption

according to purpose is a much better– more correct and detailed – way to obtain realistic results. This is out of scope of what is feasible in a rolling module for SILC. Additionally, using the COICOP classification products gives falsely the impression that expenses were carefully evaluated to be in or out, as is in HBS.

4.2.2   Food outside home

Food outside home has at the unweighted individual level a mean of 39,5 euro, a median of 30 euro in a typical week, and a mode of 20 euro. The range is high: 1.550 euro. As such, this results in a standard deviation of 44,9 euro. Weighted the mean remains 39,5 euro, and the median 30 euro, both at the individual level. There are 66 zero values, but these are plausible, and included in the analysis.

In Table 7 the weighted results are presented at the household level, and compared with HBS results. For SILC, all amounts were summed up within the household. However, in the mean and median calculations, the households with a zero value are excluded, as this is also the case in HBS. Median and mean values differ substantially between both sources, and are higher in SILC than in HBS. This is again an overestimation – even though three separate questions were used to measure food outside home as accurate as possible. When comparing information on food outside home from SILC with the counterparts in HBS we should note that the latter has the information for all the members of the household (whereas for SILC the aggregated data misses information on the children of minus 16 year old). Consequently, SILC mean and median might still be underestimated. Again this shows the poor quality of the SILC data.

	SILC	HBS
Variable / description	PC010T4 Food outside home	Expenses for COICOP111: Catering services
Reference year	2017	2016
Reference period	A typical year	Year divided by 52
Median	€44	€23,7
Mean	€65,6	€39,8
Base	All hh, N=3.563 (aggregation from individual level 16+)	All hh mentioning a expense for COICOP0111, N=3.792
Weighted	yes	yes

Table 7: Comparison with HBS for food outside home

Again, feedback from the fieldwork learns us that respondents have difficulty estimating these expenses per week because of the monthly, two-monthly … frequency of spending on food outside home. Here, the effect of responding with round amounts is also present. Both problems might have caused the overestimation.

4.2.3 Public transport

Public transport has at the unweighted individual level a mean of 10,2 euro, a median of 5 euro in a typical week, and a mode of 1 euro. The range is high: 999,9 euro. As such, this results in a high standard deviation of 31,8 euro. Weighted the mean remains 10,2 euro, and the median 5 euro. The 514 individuals spending no amount on public transport – for example because it is paid by their employer – are not included in this analysis, as they are coded as non-applicable as discussed above. Including these values would lower the unweighted mean to 8,3 euro, median to 4 euro, and mode to 0 euro.

The median and mean values between SILC and HBS are comparable when using adjusted COICOP categories in HBS, and excluding zero values in both calculations. Table 8 presents the weighted results at the household level. The median and mean values differ substantially between both sources, and are again overestimated in SILC.

	SILC	HBS	
Variable / description	PC020T4 Public transport	Expenses for COICOP 0731-0732-0735: Personal Transport services by railway, road, combined passenger transport	Expenses for COICOP 073: Personal Transport services
Reference year	2017	2016	
Reference period	A typical week	Year divided by 52	
Median	€7	€6,2	€9,3
Mean	€13,6	€12,4	€23,5
Base	All hh, N=1.659 (aggregation from individual level 16+)	All hh mentioning a expense for COICOP 0731-0732-0735, N=1.156	All hh mentioning a expense for COICOP 073, N=1.358
Weighted	yes	yes	

Table 8: Comparison with HBS for public transport

The two questions were: "During a typical week, do you use public transport (such as train, tram, subway, bus, taxi, …)?" (Yes/No) If yes: "Could you tell me how much you spend per week on public transport?" (Amount per week: … €). Respondents are to include COICOP group 07.3 (railway, road, air, sea and inland waterway, combined passenger transport and other purchased transport services). Answering the questions on public transport, we do not believe respondents take air or sea and inland waterway into account for a typical week, as these are transports more often used for holidays in Belgium. Of course there are people using them in a typical week for professional means, but then the expenses are not on their personal account. When in HBS all 07.3 transport is selected, it also includes these types of transport for holidays. As such, there is an important difference in reference period between both surveys. So, without air, sea and inland waterway and other purchased transport services, the HBS descriptive

statistics are only half of what they are. Comparing all 07.3 public transportations, SILC results are heavily underestimated. Excluding air, water and other, SILC results are slightly higher than HBS, but in fact incomparable.

4.2.4 Private transport

Private transport has at the unweighted household level a mean of 65,7 euro, a median of 40 and mode of 0 euro in a typical week. The range is high: 10.000 euro. As such, this results in a high standard deviation of 187,5 euro. Weighted the mean is 67,1 euro, and the median remains 40 euro. This analysis includes the zero values. Without them, there is a unweighted mean of 82,3 and median of 50, and a weighted mean of 83,4, while the median remains 50.

Table 9 presents the comparison with HBS, where zero values are excluded. The results are in line with the food variables results. Again, SILC mean and median is substantially higher than HBS mean and median. This overestimation might be caused by the reference period for a typical week, asking difficult calculations from respondents. The questions was formulated as: "You have specified that your household has [specified number of] car(s), [specified number of] bike(s) and [specified number of] moped(s) or motorcycle(s). Could you tell me how much you spend per week on private transport? This includes gas expenses, insurance, maintenance and spare parts, parking tickets, ..." (Amount per week: … €). COICOP 12.5.4 (insurance) should also be included – 12.5.4.1 refers to motor vehicle insurance and should indeed be included, while 12.5.4.2 refers to travel insurance which should not be included as the reference period is a typical week. Excluding 12.5.4.2 in HBS lowers the descriptive statistics, however, SILC results remain highly overestimated.

	SILC	HBS	
Variable / description	HC040T4 Private transport	Expenses for COICOP 0721-0722-07230-072420: Operations of personal transport and COICOP 012541: Motor vehicle insurance	Expenses for COICOP 0721-0722-07230-072420: Operations of personal transport and COICOP 01254: Insurance connected with transport
Reference year	2017	2016	
Reference period	A typical week	Year divided by 52	
Median	€50	€35,8	€37,4
Mean	€83	€50,8	€55,3
Base	All hh using private transport, N=3.879 (non-response excluded)	All hh mentioning a expense for COICOP0721-0722-07230-072420 and COICOP12541, N=3.943	All hh mentioning a expense for COICOP0721-0722-07230-072420 and COICOP1254, N=3.983
Weighted	yes	yes	

Table 9: Comparison with HBS for private transport

4.2.5 Regular savings

Regular savings has at the unweighted household level a mean of 370,5 euro, a median of 75 and mode of 0 euro in a typical month. The range is high: 180.000 euro. As such, this results in a high standard deviation of 3.460 euro. Weighted the mean is 361,6 euro, and the median remains 80 euro. Table 10 presents the comparison with HBS, where zero values are excluded to make the comparison possible. However, HBS statistics are calculated on only 321 respondents and are therefore unreliable to compare with.

	SILC	HBS
Variable / description	HC050T4 Regular savings in typical month	Deposits on saving accounts: No adequate comparison because of low N *
Reference year	2017	2016
Reference period	A typical week	Year divided by 52
Median	€300	€221*
Mean	€638	€100*
Base	All hh, N=2.899	All hh mentioning a expense for 'Belgian' COICOP129990, N=321
Weighted	yes	yes

Table 10: Comparison with HBS for regular savings

4.2.6 Conclusion for module on consumption

The data validation for the consumption module is not that positive for SILC. Analysis shows that almost all consumption variables are overestimated in SILC.

4.3 Module on wealth

As a first step for each variable in the wealth module, the descriptive statistics are calculated, with and without weights to get both an insight into the data as well as the generalizations to the population. In a second step, this is checked with the external information from HBS and HFCS.

4.3.1 Value of main residence

Value of the main residence has at the unweighted household level a mean of 323.147 euro, a median of 275.000 and mode of 300.000 euro. The range is high: 650.000 euro. As such, this results in a high standard deviation of 335.308 euro. Weighted the mean is 319.128 euro, and the median remains 275.000 euro.

A comparison between SILC and HFCS median shows that again SILC has a higher value (Table 12). Contrary to the comparison between SILC and HBS for the consumption module above, were the overestimation of SILC could be explained by the methods (i.e. 1 interview question in SILC and a diary in HBS), there is no fundamental difference in mode between SILC and HFCS. However, knowing that HFCS oversamples the more wealthy households, this is an unexpected result.

	SILC	HFCS
Variable / description	HV010T4 Value of main residence	Value of main residence
Reference year	2017	2014
Reference period	Current	Current
Median	€275.000	€249.700
Mean	€319.127,55	
Base	All owners, N=4.048	All owners, N= 1.567
Weighted	yes	yes

Table 11: Comparison with HFCS for value of the main residence

Additionally, we observe 5 households reporting a value of 0, and 7 additional household reporting extremely low values (smaller than 550 euro). These cases were analyzed in depth, but no specific profile emerged, and the quality of these answers is uncertain. To get some insight into the quality of this variable, we supplemented the main question with a clarifying question asking the household respondent which information was used to obtain the value. The unweighted frequencies are presented in Table 12. In contrast to what we thought approximately 50% of the households have a (more or less) reliable source to base their estimation on. However, for more than one third it is a wild guess.

Source of the information	N (%)
Estimation of a professional real estate agent	1.121 (11,29%)
Value of a comparable property nearby	1.675 (16,86%)
Knowledge of the housing market	1.168 (11,76%)
Calculated based on the price paid or previous estimation	1.089 (10,96%)
Wild guess	3.812 (38,38%)
Other	349 (3,51%)
Refusal	108 (1,09%)
I don't know	611 (6,15%)

Table 12: Frequency table of source of information for value of the main residence

Most of the answers in the category 'other' also refer to people without a clue, giving a wild guess. Together with the 38,38%, this gives an extremely high fraction of guesses, and we do not believe this variable to be reliable at all. Others refer to their own professional experience, to the fact that the residence is recently sold, to estimations by public notary after the passing of their partner or to estimations in light of an insurance policy.

4.3.2 Possession of second (more) residence(s)

15% of the households mention to be in possession of second (more) residence(s), 85% mentioned not to be. The comparison with HBS and HFCS is not self-evident (Table 13), as the HFCS' scope is much more limited than in SILC. HFCS is interested in a second dwelling owned by the household to be used by the household, and not investment property. For SILC this variable should capture if households own properties other than the main residence, all properties should be taken into account. The oversampling of the wealthy in HFCS is an

additional difficulty. Therefore HFCS and SILC results are incomparable. The comparison between SILC and HBS is not straightforward either, as there is no such variable in HBS. However, there is a variable measuring the expenses for a second residence, which can be used to derive this information, which is not ideal and might be an underestimation of this proportion. The extremely low percentages in HBS compared to SILC and HFCS confirms this assumption. Again, both surveys are incomparable for this variable.

	SILC	HFCS	HBS
Variable / description	HV020T4 Possession of second (more) residence(s)	Possession of other real estate property	Derive an estimate of 'Possession of second residence' from the variable "Expenses for Second residence": No adequate comparison available because of differences in scope*
Reference year Reference period	2017 Current	2014 / 2010 Current	2016 Current (Year divided by 52)
Percent	15%	18,5% / 16,4%	4,66%*
Base Weighted	All hh, N=5.820 yes	All hh, N=2.238 yes	4490 yes

Table 13: Comparison with HBS and HFCS for possession of second (more) residence

There is also an overlap with the current SILC variable HY040 (income from rental). In Belgium we first ask households whether they (or one of their members) own a second property, and then we ask for the rent received. However, the reference period there is the income reference period, while for HV020T4 the reference period is current – except for countries using administrative data, then the income reference period is possible as well. For respondents it is annoying to receive kind of the same question twice, only with a difference in reference period.

4.3.3 Possession of deposits

82% of the households mention to be in possession of deposits, 18% mentioned not to be. This percentage is lower in SILC than in HFCS (Table 14). Again, the oversampling in HFCS should be taken into account. More importantly, as the aim of the HFCS is to obtain results on the distribution of wealth among households, the respondents explicitly know in advance that questions will concern there wealth. We believe responses will be more accurate than in the module of SILC where this is not the case. The comparison with HBS is not possible as 23% of the households did not report any bank charges – this is composed of 12% giving a zero value and 11% answering 'don't know'. We cannot derive 89% with a possession of deposits, as it is unclear for the 11% whether they do not have a deposit, or whether they just did not know the amount of the charges.

	SILC	HFCS	HBS
Variable / description	HV030T4 Possession of deposits (including sight accounts, saving accounts, time deposits, etc)	Possession of deposits (including sight accounts, saving accounts, time deposits, etc)	Derive an estimate for possession of deposits from "Bank charges": No adequate comparison available*
Reference year	2017	2014	2016
Reference period	Current	Current	
Percent	82%	97%	
Base	All hh, N=5.779	All hh, N=2.238	
Weighted	yes	yes	

Table 14: Comparison with HBS and HFCS for possession of deposits

4.3.4 Value of deposits

Value of deposits has at the unweighted household level a mean of 30.947 euro, a median of 5.104 and mode of 0 euro. The range is high: 700.000 euro. As such, this results in a high standard deviation of 64.803 euro. Weighted the mean is 30.844 euro, and the median is 5.000 euro.

The comparison with HFCS is presented in Table 15. However, in HFCS saving and sight accounts are reported separately. Taking this into account, it is still obvious that the HFCS median is substantially higher than SILC median. However, it is unclear to what extent zero values are included in HFCS, while in SILC they are. Omitting these values in SILC doubles the median to 10.000 euro, which is still lower than HFCS. Again, this might be caused by the fact that respondents know beforehand that the survey is about wealth and the oversampling of the wealthy, while in SILC this is not the case. Additionally, HFCS uses several questions while in SILC only 1 was asked.

	SILC	HFCS
Variable / description	HV040T4 Value of deposits	Value of deposits: value of sight accounts, value of saving accounts
Reference year	2017	2014
Reference period	Current	Current
Median	€5.000	€16.000 (saving account) €1.800 (sight account)
Mean	€30.884	
Base	All hh having deposits, N=2.238	All hh, N=2.238
Weighted	yes	yes

Table 15: Comparison with HFCS for value of deposits

4.3.5 Possession of bonds

19% of the households mention to be in possession of bonds, shares publicly traded or mutual funds, 81% mention not to be. A comparison between SILC and HFCS is presented in Table 16. In HFCS 7% of the households reports having bonds, 11% report having shares and 21% report

having mutual funds. There are some households with more than one type of financial product, so these percentages cannot be summed. Still, 21% of the households report having mutual funds, while in SILC there is 19% reporting having at least one of the financial products. Even though the comparison is not that straightforward, there seems to be an underestimation in SILC. The same three explanations as for deposits might apply here: oversampling, knowing the theme in advance and separate questions.

	SILC	HFCS
Variable / description	HV050T4 Possession of bonds, shares publicly traded or mutual funds	Possession of bonds, Possession of shares (publicly traded), Possession of mutual funds
Reference year	2017	2014
Reference period	Current	Current
Percent	19%	7.8% (bonds) 11% (shares) 21% (mutual funds)
Base	All hh, N=5.690	All hh, N=2.238
Weighted	yes	yes

Table 16: Comparison with HFCS for possession of bonds

4.3.6 Value of bonds, shares publicly traded or mutual funds

Value of bonds, shares or mutual funds has at the unweighted household level a mean of 78.216 euro, a median of 24.500 and mode of 10.000 euro. The range is high: 900.000 euro. As such, this results in a high standard deviation of 64.803 euro. Weighted the mean is 74.023 euro, and the median is 24.000 euro. The 11 households with a zero value are included here.

	SILC	HFCS
Variable / description	HV060T4 Value of bonds, shares publicly traded or mutual funds	Value of bonds, Value of shares (publicly traded), Value of mutual funds
Reference year	2017	2014
Reference period	Current	Current
Median	€24.000	€12.400 (bonds) €10.000 (shares) €28.800 (mutual funds)
Mean	€74.023	
Base	All of the 19% of hh having bonds etc., N=396	All hh possessing bonds, shares, mutual funds
Weighted	yes	yes

Table 17: Comparison with HFCS for value of bonds

Table 17 presents the comparison for value of bonds between SILC and HFCS. Again, in HFCS there are three separate variables for bonds, shares and mutual funds. A comparison is not straightforward, but the median of mutual funds in HFCS is higher than the SILC median for all

financial products together. Based on this, it seems that SILC underestimates households' wealth – for the same three reasons discussed above.

4.3.7   Conclusion for module on wealth

The data validation for the wealth module is not that optimistic for SILC. Analysis shows that all wealth variables are underestimated in SILC.

5. Conclusions and advice

With the goal to have good, reliable and validated variables to include in the SILC rolling modules with the implementation of IESS, Statbel tested the first proposition of consumption and wealth variables in SILC 2017. This report described in detail our experiences with the fieldwork and the quality of the final variables. In this last section we will bring all findings together and formulate concrete advice for the future SILC.

5.1 Conclusions module on consumption

The item non-response for the consumption module is overall high, and unacceptable high for private transport and regular savings (Table 18). Each variable on its own is very complex for respondents. The totality of all these difficult and complex questions together in this module places an unacceptable high burden on respondents. Data quality shows that it is suffering.

Item non-response module on Consumption	%
HC010T4 - Food at home	13
PC010T4 - Food outside home	7
PC020T4 - Public transport	5
HC040T4 - Private transport	23,5
HC050T4 - Regular savings	15

Table 18: Overview item non-response consumption

The data validation for the consumption module is not that positive for SILC. Analysis shows that almost all consumption variables are overestimated in SILC. There are several reasons for this overestimation:

- asking too much information per variable (respondents tend to include more than the set of COICOP classes aimed at),
- weekly amounts may be estimated higher than monthly or annual,
- respondents estimate their weekly expenses with round numbers (e.g. 80, 100, 150, …).

5.2 Conclusions module on wealth

Overall, the item non-response for the module on wealth is higher than used to for household variables in SILC on the one hand, and the consumption module on the other hand (Table 19). This is a clear indication that the wealth module poses too many problems. Both for the respondents – it is a heavy burden to answer, the questions are difficult and extremely sensitive – as well as at the data quality level – there are too many refusals, and answers given seem not always reliable.

Item non-response module on Wealth	%
HV010T4 – Value of main residence	18
HV020T4 – Possession of second (more) residence(s)	4
HV030T4 – Possession of deposits	5
HV040T4 – Value of deposits	52,5
HV050T4 – Possession of bonds, shares publicly traded or mutual funds	6
HV060T4 – Value of bonds, shares publicly traded or mutual funds	63,5

Table 19: Overview item non-response wealth

Also, the data validation for the wealth module is again not that positive for SILC. Analysis shows that all wealth variables are underestimated in SILC. There are several reasons for this underestimation:

- the sensitive topic,
- and asking too much information per variable (where – in contrary to a less sensitive topic such as consumption - respondents are less prepared or willing to give detailed information on their wealth).

5.3   General conclusions and concrete advice

Taking everything together, we can conclude that:

- the burden on respondents is higher than the burden for the SILC 2016 ad hoc module. Using single questions for almost all variables is the only way to keep the module acceptable for respondents.
- under no circumstances variables can be added.
- given that the questions are often open to wider interpretation, specific situations have to be explained in guidelines (which reduces interviewer bias), but at the same time balance has to be maintained with a burden on interviewer/household.
- when analyzing each variable we often concluded that the reliability is low. What is the value of these module variables when they are not reliable, i.e. when many respondents do not provide answers, and when the answers given are proven to be over- or underestimated when compared to the expert sources? How can these variables be used in a statistical matching process when they do not match with the sources to be matched? Indeed, there is a difference in year of reference between SILC and HFCS, but the difference between SILC and HBS is only one year.

Furthermore, the analysis brings **specific advice for each variable**:

Advice Food at home (HC010T4):

- Drop variable – what is the value of a variable that is clearly overestimated?

- If variable is kept:
  o   note that household level is feasible.

43

- o reduce the complexity of the variable by using a typical month as reference period and including the alcoholic beverages (especially as they are included in food outside home) so the variable includes simply all food and beverage expenses to consume at home.
- o do not use the COICOP classification; just use 'food and beverages to consume at home'. Using it gives a false impression of being complete and very accurate. Instead of using COICOP classifications guidelines for interviewers can clarify that it refers to food and beverages paid by the household to consume at the households' residence – i.e. possibly including more than only the household members.

Advice Food outside home (PC010T4):

- Drop variable – what is the value of a variable that is clearly overestimated?

- If variable is kept:
  - o note that personal level is feasible – but a high burden. However, an important difference between the household and the personal level should be taken into account, as the household level includes also the children aged less than 16, while the personal level does not. If the personal level is chosen, aggregations at the household are not always that convenient.
  - o add the 'non applicable' flag.
  - o reduce the complexity of the variable by changing the reference period to a typical month.
  - o if typical week is kept as reference period, clarify that averages should be calculated in case of events that happen with a monthly, two-monthly, … frequency.

Advice Public transport (PC020T4):

- Drop variable – what is the value of a variable that is clearly underestimated?

- If variable is kept:
  - o note that personal level is feasible.
  - o reduce the complexity of the variable by changing the reference period to a typical month.
  - o clarify the value 'no expenditure' in the guidelines.
  - o allow a zero value.
  - o exclude air and water transport, or i.e. do not use the COICOP classification

Advice Private transport (HC040T4):

- Drop variable – what is the value of a variable that is clearly overestimated?

- If variable is kept:
  - note that household level is feasible, but difficult for households in specific situations (e.g. children with own vehicles and responsibilities thereof, partners with not-shared vehicles, etc.).
  - reduce the complexity of the variable by changing the reference period to a typical month.
  - clarify the values in the guidelines 'no expenditure', while in fact this is equal to the zero value that is allowed
  - exclude travel insurance, or i.e. do not use the COICOP classification
  - Guidelines should clarify what to do with the taxes paid on the 'free use' of company cars, the use of cars in the sharing economy.

Advice Regular Savings (HC050T4):

- Drop variable because of low reliability.

- If variable is kept:
  - note that household level is feasible.
  - clarify the values in the guidelines 'no saving'.
  - allow zero values.
  - consider a dichotomous variable 'yes able to save in a typical month' and 'no not able to save in a typical month'.

Advice Value of main residence (HV010T4):

- Drop variable because of low reliability.

- If it is kept, the guidelines should clarify that it concerns the value of the main residence in its current state, in cases of renovation the current state should be taken into account: not yet renovated, partially renovated or already renovated.

Advice Possession of second residence (HV020T4):

- Keep possession of second residence without the value.
- Change the reference period from 'current' to 'income reference period' for everybody (not only for countries using administrative data) as to avoid overlap with the variable HY040 (Income from rental).

Advice Possession of deposits (HV030T4):

- Note that there is an underestimation.
- Possible to keep, even if value is dropped (cf. possession of second residence without the value)

Advice Value of deposits (HV040T4):

- Drop variable because underestimated (very sensitive).

- If question is kept:
  - note that household level is feasible.
  - consider using a categorical variable.
  - allow negative values.
  - allow higher values.

.

Advice Possession of bonds (HV050T4):

- Note that there is an underestimation (very sensitive).
- Possible to keep, even if value is dropped (cf. possession of second residence without the value)

Advice Value of bonds (HV060T4):

- Drop variable because underestimated (too sensitive).

Annex 1: Overview of the questions

VARIABLE	UNIT	MODE OF COLLECTION	QUESTION(S)
HC010T4: Food at home (MoC1)	Household - all current household members	Personal interview with the household respondent	The next question relates to the budget your household spends on <u>groceries</u> during a typical week. Could you tell me how much your household does spend per week to purchase food (such as bread, meat, fish, cheese, eggs, fruit, vegetables, …) and non-alcoholic beverages (such as coffee, tea, milk, mineral waters, soft drinks, fruit and vegetable juices, ...)? It concerns the purchase of food and non-alcoholic beverages for consumption at home. Amount (per week): ………………….. €
PC010T4: Food outside home (MoC2aa)	Individual level - each current household member aged 16 and over or selected respondent (where applies)	Personal interview (proxy as an exception)	Now I'd like to ask you a few questions about the budget you spend on <u>food and drinks outside the home</u> during a typical week. During a typical week, do you eat outdoors (at a restaurant, a student canteen, a cafeteria, a snack bar, a sandwich bar, a chip shop, ...)? 1. Yes 2. No
PC010T4: Food outside home (MoC2ab)			Could you tell me how much you spend per week on outdoor eating? Amount (per week): ………………….. €

PC010T4: Food outside home (MoC2ba)	Individual level - each current household member aged 16 and over or selected respondent (where applies)	Personal interview (proxy as an exception)	During a typical week, do you make use of takeaway or delivery meals? 1. Yes 2. No
PC010T4: Food outside home (MoC2bb)			Could you tell me how much you spend per week on takeaway or delivery meals? Amount (per week): ………………….. €
PC010T4: Food outside home (MoC2ca)	Individual level - each current household member aged 16 and over or selected respondent (where applies)	Personal interview (proxy as an exception)	During a typical week, are you going for a drink in a bar, a coffee shop, a teahouse, a canteen or another catering establishment? 1. Yes 2. No
PC010T4: Food outside home (MoC2cb)			Could you tell me how much you spend per week on outdoor drinking (alcoholic and non-alcoholic beverages)? Amount (per week): ………………….. €
PC020T4: Public transport (MoC3a)	Individual level - each current household member aged 16 and over or selected respondent (where applies)	Personal interview (proxy as an exception)	During a typical week, do you use public transport (such as train, tram, subway, bus, taxi, …)? 1. Yes 2. No
PC020T4: Public transport (MoC3b)			Could you tell me how much you spend per week on public transport? Amount (per week): ………………….. €

HC040T4: Private transport (MoC4)	Household - all current household members	Personal interview with the household respondent	You have specified that your household has [specified number of] car(s), [specified number of] bike(s) and [specified number of] moped(s) or motorcycle(s). Could you tell me how much you spend per week on <u>private transport</u>? This includes gas expenses, insurance, maintenance and spare parts, parking tickets, ... Amount (per week): ………………….. €
HC050T4: Regular savings (in a typical month) (MoC5)	Household - all current household members	Personal interview with the household respondent	The next question relates to the <u>saving behavior</u> of your household during a typical month. Could you tell me how much your household set aside each month? If the amount your household set aside differs from month to month, please give an average amount. Amount (per month): …………………. €
HV010T4: Value of main residence (MoW1a)	Household - all current household members	Personal interview with the household respondent	About how much would you expect to get for your home if you sell it now? If house is shared with someone else, please consider the price of the entire property and not just your household's share. Amount: ………………….. €
HV010T4: Value of main residence (MoW1b)	Household - all current household members	Personal interview with the household respondent	Can you tell me what this estimate is based on? 1. Professional valulation - estate agent/surveyor etc. 2. Price of neighbouring/similar property 3. Knowledge of local market 4. Calculation based on purchase price or earlier valuation 5. Guess 6. Other --> Describe

HV020T4: Possession of second (more) residence(s) (MoW2)	Household - all current household members	Personal interview with the household respondent	Are you or another person from your household owner of another dwelling, building or ground except the dwelling you currently live in? 1. Yes 2. No
HV030T4: Possession of deposits (MoW3a)	Household - all current household members	Personal interview with the household respondent	Does anyone in your household have a sight account, any saving accounts, time deposits, certificates of deposit or other such deposits? Sight accounts are called also current accounts, draft accounts or checking accounts. 1. Yes 2. No
HV040T4: Value of deposits (MoW3b)	Household - all current household members	Personal interview with the household respondent	In total, how much is in all these accounts now? Amount: ………………….. €
HV050T4: Possession of bonds, shares publicly traded or mutual funds (MoW4a)	Household - all current household members	Personal interview with the household respondent	Does anyone in your household have or own corporate or government bonds, bills or notes or stock shares in any publicly traded companies or any investments in mutual funds, money market mutual funds or hedge funds? Pension savings funds should not be taken into account. 1. Yes 2. No
HV050T4: Value of bonds, shares publicly traded or mutual funds (MoW4b)	Household - all current household members	Personal interview with the household respondent	What is the current market value of your household's investments in all bonds/debt securities, in all shares and each type of fund? Amount: ………………….. €

Annex 2: Comparison SILC, HBS and HFCS

	SILC	HFCS	HBS
Number of households	6.053	2.238*	4.490
Reference year	2017	2014	2016

* oversampling of wealthy households (based on administrative information on local level).

Household structure	SILC	HFCS	HBS

Household size:

	SILC	HFCS	HBS
1	32,69	33,8	33,84
2	33,54	31,6	31,26
3	13,27	15,1	15,49
4	13,45	12,6	13,56
5 and more	7,05	6,8	5,75

Household status:

	SILC		HFCS		HBS	
Owner – outright	35		38,4		27	
Owner – by mortgage	33	68	31,9	70,3	46	73
Renter or other	32		29,7		27	