POVERTY AND SOCIAL INCLUSION INDICATORS – METHODOLOGICAL NOTES[1]

The Survey on Income and Living Conditions (SILC) provides four basic files containing target variables based on common concepts and definitions. On the basis of these target variables, additional variables (derived variables) are calculated for each statistical unit-observation, to support the computation of the indicators. Additionally, a relatively important component of the variables is the linking or identification variables, such as the year of the survey, that characterize the whole survey. Auxiliary variables on the other hand, are also computed variables but rather than referring to distinct statistical units-observations, they refer to the whole statistical population.

Household target variables

Household target variables are variables collected from the sampled units and concern the household in relation with the observed phenomenon. These main target variables are listed in table:

HOUSEHOLD DATA (H-FILE)
HB050: MONTH OF HOUSEHOLD INTERVIEW
HB060: YEAR OF HOUSEHOLD INTERVIEW
HH021: TENURE STATUS
HH040: LEAKING ROOF, DAMP WALLS/FLOORS/FOUNDATION, OR ROT IN WINDOW FRAMES OR FLOOR
HH050: ABILITY TO KEEP HOME ADEQUATELY WARM
HH060: CURRENT RENT RELATED TO OCCUPIED DWELLING
HH070: TOTAL HOUSING COST
HH081: BATH OR SHOWER IN DWELLING
HH091: INDOOR FLUSHING TOILET FOR SOLE USE OF HOUSEHOLD
HS011: ARREARS ON MORTGAGE OR RENT PAYMENTS
HS021: ARREARS ON UTILITY BILLS
HS031: ARREARS ON HIRE PURCHASE INSTALMENTS OR OTHER LOAN PAYMENTS
HS040: CAPACITY TO AFFORD PAYING FOR ONE WEEK ANNUAL HOLIDAY AWAY FROM HOME
HS050: CAPACITY TO AFFORD A MEAL WITH MEAT, CHICKEN, FISH (OR VEGETARIAN EQUIVALENT) EVERY SECOND DAY
HS060: CAPACITY TO FACE UNEXPECTED FINANCIAL EXPENSES
HS070: DO YOU HAVE A TELEPHONE (INCLUDING MOBILE PHONE)?
HS080: DO YOU HAVE A COLOUR TV?
HS090: DO YOU HAVE A COMPUTER?
HS100: DO YOU HAVE A WASHING MACHINE?
HS110: DO YOU HAVE A CAR?

---

[1] These methodological notes are extracted from Eurostat` Working paper with the description of the 'Income and living conditions dataset'. Detailed information about the list of social inclusion indicators, definitions and algorithm for their calculation can be found on the following site:
http://ec.europa.eu/eurostat/statistics-explained/index.php/EU_statistics_on_income_and_living_conditions_(EU-SILC)_methodology;

HS120: ABILITY TO MAKE ENDS MEET
HS140: FINANCIAL BURDEN OF THE TOTAL HOUSING COST
HS150: FINANCIAL BURDEN OF THE REPAYMENT OF DEBTS FROM HIRE PURCHASES OR LOANS
HS160: PROBLEMS WITH THE DWELLING: TOO DARK, NOT ENOUGH LIGHT
HS170: NOISE FROM NEIGHBOURS OR FROM THE STREET
HS180: POLLUTION, GRIME OR OTHER ENVIRONMENTAL PROBLEMS
HS190: CRIME VIOLENCE OR VANDALISM IN THE AREA
HY010: TOTAL HOUSEHOLD GROSS INCOME
HY020: TOTAL DISPOSABLE HOUSEHOLD INCOME
HY022: TOTAL DISPOSABLE HOUSEHOLD INCOME BEFORE SOCIAL TRANSFERS OTHER THAN OLD-AGE AND SURVIVOR'S BENEFITS
HY023: TOTAL DISPOSABLE HOUSEHOLD INCOME BEFORE SOCIAL TRANSFERS INCLUDING OLD-AGE AND SURVIVOR'S BENEFITS
HY040G/HY040N: INCOME FROM RENTAL OF A PROPERTY OR LAND
HY090G/HY090N: INTEREST, DIVIDENDS, PROFIT FROM CAPITAL INVESTMENTS IN UNINCORPORATED BUSINESS
HY050G/HY050N: FAMILY/CHILDREN RELATED ALLOWANCES
HY060G/HY060N: SOCIAL EXCLUSION NOT ELSEWHERE CLASSIFIED
HY070G/HY070N: HOUSING ALLOWANCES
HY080G/HY080N: REGULAR INTER-HOUSEHOLD CASH TRANSFER RECEIVED
HY081G/HY081N: ALIMONIES RECEIVED (COMPULSORY + VOLUNTARY)
HY100G/HY100N: INTEREST REPAYMENTS ON MORTGAGE
HY110G/HY110N: INCOME RECEIVED BY PEOPLE AGED UNDER 16
HY120G/HY120N: REGULAR TAXES ON WEALTH
HY130G/HY130N: REGULAR INTER-HOUSEHOLD CASH TRANSFER PAID
HY131G/HY131N: ALIMONIES PAID (COMPULSORY + VOLUNTARY)
HY140G/HY140N: TAX ON INCOME AND SOCIAL CONTRIBUTIONS
HY170G/HY170N: VALUE OF GOODS PRODUCED FOR OWN CONSUMPTION
PERSONAL DATA (P-FILE)
PB010: YEAR OF THE SURVEY
PB020: COUNTRY
PB030: PERSONAL ID
PB040: PERSONAL CROSS-SECTIONAL WEIGHT
PB150: SEX
PB180: SPOUSE/PARTNER ID
PB210: COUNTRY OF BIRTH
PB220A: CITIZENSHIP 1
PE040: HIGHEST ISCED LEVEL ATTAINED
PH010: GENERAL HEALTH
PL020: ACTIVELY LOOKING FOR A JOB
PL025: AVAILABLE FOR WORK
PL031: SELF-DEFINED CURRENT ECONOMIC STATUS
PL040: STATUS IN EMPLOYMENT
PL051: OCCUPATION (ISCO-08 (COM))
PL060: NUMBER OF HOURS USUALLY WORKED PER WEEK IN MAIN JOB
PL073: NUMBER OF MONTHS SPENT AT FULL-TIME WORK AS EMPLOYEE
PL074: NUMBER OF MONTHS SPENT AT PART-TIME WORK AS EMPLOYEE
PL075: NUMBER OF MONTHS SPENT AT FULL-TIME WORK AS SELF-EMPLOYED (INCLUDING FAMILY WORKER)

PL076: NUMBER OF MONTHS SPENT AT PART-TIME WORK AS SELF-EMPLOYED (INCLUDING FAMILY WORKER)
PL080: NUMBER OF MONTHS SPENT IN UNEMPLOYMENT
PL085: NUMBER OF MONTHS SPENT IN RETIREMENT OR EARLY RETIREMENT
PL086: NUMBER OF MONTHS SPENT AS DISABLED OR/AND UNFIT TO WORK
PL087: NUMBER OF MONTHS SPENT STUDYING
PL089: NUMBER OF MONTHS SPENT FULFILLING DOMESTIC TASKS AND CARE RESPONSIBILITIES
PL090: NUMBER OF MONTHS SPENT IN OTHER INACTIVITY
PL100: TOTAL NUMBER OF HOURS USUALLY WORKED IN SECOND, THIRD… JOBS
PL111: NACE REV.2
PL140: TYPE OF CONTRACT
PY010G/PY010N: EMPLOYEE CASH OR NEAR CASH INCOME
PY020G/PY020N: NON-CASH EMPLOYEE INCOME
PY050G/PY050N: CASH BENEFITS OR LOSSES FROM SELF-EMPLOYMENT
PY080G/PY080N: PENSION FROM INDIVIDUAL PRIVATE PLANS
PY100G/PY100N: OLD-AGE BENEFITS
PY110G/PY110N: SURVIVOR' BENEFITS
PERSONAL REGISTER (R-FILE)
RB010: YEAR OF THE SURVEY
RB020: COUNTRY
RB030: PERSONAL ID
RB031: YEAR OF IMMIGRATION
RB050: PERSONAL CROSS-SECTIONAL WEIGHT
RB060: PERSONAL BASE WEIGHT
RB062: LONGITUDINAL WEIGHT (TWO-YEAR DURATION)
RB063: LONGITUDINAL WEIGHT (THREE-YEAR DURATION)
RB064: LONGITUDINAL WEIGHT (FOUR-YEAR DURATION)
RB070: MONTH OF BIRTH
RB080: YEAR OF BIRTH
RB090: SEX
RB220: FATHER ID
RB230: MOTHER ID
RB240: SPOUSE/PARTNER ID
RB245: RESPONDENT STATUS
RB250: DATA STATUS
RL010: EDUCATION AT PRE-SCHOOL
RL020: EDUCATION AT COMPULSORY SCHOOL
RL030: CHILD CARE AT CENTRE-BASED SERVICES
RL040: CHILD CARE AT DAY-CARE CENTRE
RL050: CHILD CARE BY A PROFESSIONAL CHILD-MINDER AT CHILD'S HOME OR AT CHILD-MINDER'S HOME
RL060: CHILD CARE BY GRAND-PARENTS, OTHERS HOUSEHOLD MEMBERS (OUTSIDE PARENTS), OTHER RELATIVES, FRIENDS OR NEIGHBOURS
RL070: CHILDREN CROSS-SECTIONAL WEIGHT FOR CHILD CARE

The derived variables are additional computed variables concerning the statistical unit, i.e. the household and person, and are calculated in order to support further computations. These variables are calculated based on the micro-data and further used for the computation of the indicators and dimensions.

Data weighting

The database of each country contains different types of weights:

- Household cross-sectional weight (target variable DB090) to obtain the actual number of private households in the country;
- Personal cross-sectional weight (target variable RB050) to obtain the actual number of persons in the country;
- Personal cross-sectional weight for each household member aged 16 and more (target variable PB040) to obtain the number of persons aged 16 and more in the country;

Most of the indicators use the personal cross-sectional weight (RB050), because poverty status is calculated at individual level and the target group refers to the total population living in private households. For some indicators which focus on persons aged 16 and more (e.g. "share of working poor") the personal cross-sectional weight for each household member aged 16 and more (PB040) is used.

When calculating the indicators the weights are adjusted with a weighting factor to reflect the missing cases (RB050a).

Derived variables

Age

In EU-SILC, age is defined as the age calculated at the end of the income reference period. However, data collection often occurs a few months after the end of the income reference period, so household composition is captured at the time of interview. Consequently, household members who have died between the end of the income reference period and the time of the survey data collection are not registered and babies born in this interval will be recorded with negative age at the end of the income reference period is reconstructed. The algorithm calculating age uses the following relevant basic SILC variables: RB010 (year of the survey), RB070 (month of birth), RB080 (year of birth) – in R file.

- AGE= (((RB010 - 1) * 100 + 12) - (RB080 * 100 + RB070)) / 100

Note: if AGE=−1 age is set to AGE=0

Age is broken down in the following groups:

- Less than 6 years (Y_LT6)
- From 6 to 11 years (Y6-11)
- From 12 to 17 years (Y12-17)
- Less than 18 years (Y_LT18)
- From 18 to 59 years (Y18-59)
- From 18 to 64 years (Y18-64)
- Less than 60 years (Y_LT60)
- From 60 years or over (Y_GE60)
- Less than 65 years (Y_LT65)
- From 65 years or over (Y_GE65)
- Less than 75 years (Y_LT75)
- From 75 years and over (Y_GE75)

Age at the date of interview (AGE_IW)

The algorithm calculating age in work (AGE_IW) uses the following relevant basic SILC variables: RB070 (month of birth), RB080 (year of birth), HB050 (month of household interview), HB060 (year of household interview).

$$AGE\_IW = floor\ (\frac{(HB060\ X\ 100 + HB050) - (RB080\ X\ 100 + RB070)}{100})$$

Highest level of education of children's parents (HHISCED)

Highest educational level of children's parents refers to children living in a household with one or both parents and to the highest level of education attained by (at least one of) the parents. Data are classified according to the International Standard Classification of Education (ISCED): low education corresponds to ISCED levels 0-2 (pre-primary, primary and lower secondary education); medium education corresponds to ISCED levels 3 and 4 (upper secondary and post-secondary non-tertiary education) and high education corresponds to ISCED levels 5 and 6 (tertiary education).

The algorithm for highest educational level of children's parents uses the following basic SILC variables: FPE040 (highest ISCED level attained by the father) and MPE040 (highest ISCED level attained by the mother).

The calculation of the variable highest level of education of children's parents (HHISCED) is described below.

if FPE040>=MPE040 then HHISCED = FPE040
if MPE040>=FPE040 then HHISCED = MPE040

Equivalised household size (EQ_SS)

The algorithm for equivalised household size uses the following auxiliary variables:

- hm13 - Number of household members aged 13 or less
- hm14 - Number of household members aged 14 and over
- SUM_hm13 - The total number of household members (at household level) with age 13 or less
- SUM_hm14 - The total number of household members (at household level) with age 14 and over

The calculation of variables hm13 and hm14 are described schematically below:
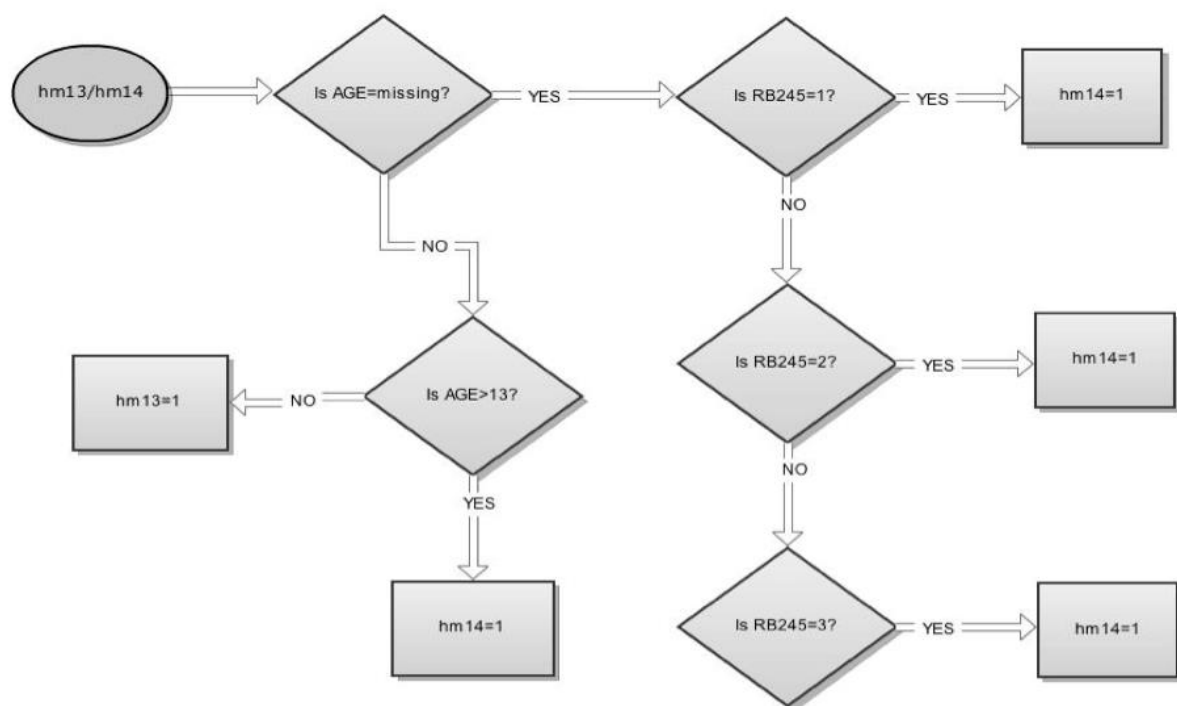
The calculation of the Equivalised household size (EQ_SS) is described below.

EQ_SS = 1 + 0.5 x (sum_hm14-1) + 0.3 x sum_ hm13.

Equivalised Disposable Income (EQ_INC)

Equivalised disposable income (EQ_INC) is the total income of a household, divided by the number of household members converted into equivalised adults; household members are equivalised or made equivalent by the following so-called modified OECD (Organisation for Economic Co-operation and Development) equivalence scale:

•   the first household member aged 14 years or more counts as 1 person
•   each other household member aged 14 years or more counts as 0.5 person
•   each household member aged 13 years or less counts as 0.3 person

Total disposable income for each household is divided by its equivalised size and the total equivalised disposable income is formed.

The following types of disposable income are calculated:

•   Equivalised disposable income after social transfers (EQ_INC20)
•   Equivalised disposable income before social transfers (excluding old-age and survivor's benefits/pensions) (EQ_INC22)
•   Equivalised disposable income before social transfers (including old-age and survivor's benefits/pensions) (EQ_INC23)

Risk of poverty threshold (ARPTXX)

The at-risk-of-poverty threshold is calculated as the XX percentage of the median value of the Equivalised disposable Income (EQ_INC) after social transfers (EQ_INC20).

$$ARPTXX = XX\% \times EQ\_IC20MEDIAN$$

The usual definition defines at-risk-of-poverty threshold as 60% of the equivalised median income after social transfers so the value ARPT60 threshold is the most commonly used. Different thresholds (ARPT40, ARPT50, ARPT70, ARPTM40, ARPT50, ARPT60) are also calculated to derive different poverty rates.

Poverty status (ARPTXXi)

The risk of poverty indicator identifies people below the Risk of poverty threshold (ARPTXX) as a poor (ARPTXXi=1) and people with Equivalised disposable Income (EQ_INC) after social transfers (EQ_INC20) above the risk of poverty threshold as a nonpoor (ARPTXXi=0).

if EQ_INC20 < ARPTXX then ARPTXXi=1

if EQ_INC20 >= ARPTXX then ARPTXXi=0

The usual definition defines at-risk-of-poverty threshold as 60% of the equivalised disposable income so the value of variable XX in the usual definition is 60 (ARPT60i).

Activity status

The most frequent activity status is the status that individuals declare themselves to have occupied for more than half the total number of months for which information on any status is available. The following classification is used for activity status in EU-SILC datasets:

- Employed persons (EMP)
- Employees (SAL)
- Employed persons except employees (NSAL)
- Not employed persons (NEMP)
- Unemployed persons (UNE)
- Retired persons (RET)
- Other inactive persons (NAC_OTH)

With regard to its calculation, the following methodological issues should be taken into consideration:

1. The most frequent activity status is defined as the status that individuals declare themselves to have occupied for more than half the total number of months for which information on any status is available. Consequently, where an individual provides information on his activity status over 12 months, his most frequent activity status will be the status he declares to have occupied for at least 7 months. Individuals who have spent only half or less than the total number of declared months in any activity status are excluded from the computation. People with less than 7 months declared in the calendar of activities are excluded.

2. The most frequent activity status for each month is based on a self-assessment by the interviewees. Therefore, it may not be entirely consistent with the ILO coding that is applied in the European Union Labour Force Survey.

3. Activity status is measured at the individual level.

For each household member aged 16 and over, the number of months in each status during the income reference period is counted. The following activity statuses will be considered:

TOT - Total number of months spent in any status during the reference period
EMP - Number of months spent in work for employed persons
SAL - Number of months spent in work for employees
NSAL - Number of months spent in work for employed persons except employees
UNEMP - Number of months spent in unemployment
RET - Number of months spent in retirement
INAC_OTH - Number of months spent as 'other inactive' (in education or training, doing housework, looking after children or other persons; in community or military service; other economically inactive)

The calculation of the current activity status of the respondent depends on the year of survey and more specifically if it is before or after 2008.

- For surveys after 2008 (DB010>2008)

For each household member the following variables will be selected: PL073, PL074, PL075, PL076, PL080, PL085, PL086, PL087, PL088, PL089, PL090. The following derived variables will be constructed:

TOT= PL073+PL074+PL075+PL076+PL080+PL085+PL086+PL087+PL088+PL089+PL090
SAL= PL073+PL074
NSAL= PL075+PL076
UNEMP= PL080
RET= PL085
INAC_OTH= PL086+PL087+PL088+PL089+PL090

The respondent is being excluded if the total number of months spent in any activity is less than seven (TOT<7). For the rest of the respondents that have reported for more than six months the activity status is calculated as follows:

If EMPTOT > 0.5 then Activity status=1
If UNEMPTOT > 0.5 then Activity status=5
If RETTOT > 0.5 then Activity status=6
If INAC_OTHTOT > 0.5 then Activity status=7
Otherwise the Activity status is missing.

- For surveys before 2009 (DB010<2009)

For each household member the following variables will be selected: PL070, PL072, PL080, PL085, PL087, PL090. The following derived variables will be constructed:

TOT= PL070+PL072+PL080+PL085+PL087+PL090 EMP= PL070+PL072
UNEMP= PL080
RET= PL085
INAC_OTH= PL087+PL090

The respondent is being excluded if the total number of months spent in any activity is less than seven (TOT<7). For the rest of the respondents that have reported for more than six months the activity status is calculated as follows:

If EMPTOT > 0.5 then Activity status=1
If UNEMPTOT > 0.5 then Activity status=5
If RETTOT > 0.5 then Activity status=6
If INAC_OTHTOT > 0.5 then Activity status=7

Household type (HHTYP)

The following household types will be considered:

- A1 Single Person (HHTYP=1-5)
- A1_LT65 One adult younger than 65 (HHTYP=1,2)
- A1_GE65 One adult older than 65 (HHTYP=3,4)
- A1_DCH Single person with dependent children (HHTYP=9)
- A1M Single male (HHTYPE=1,3)
- A1F Single female (HHTYP=2,4)
- A2 Two adults (HHTYP=6,7)
- A2_2LT65 Two adults, no dependent children, younger than 65 years (HHTYP=6)
- A2_GE1_GE65 Two adults, no dependent children, at least one adult 65 years or more (HHTYP=7)
- A2_1DCH Two adults with one dependent child (HHTYP=10)
- A2_2DCH Two adults with two dependent children (HHTYP=11)
- A2_GE3DCH Two adults with three or more dependent children (HHTYP=12)
- A_GE2_NDCH Two or more adults without dependent children (HHTYP=6-8)
- A_GE2_DCH Two or more adults with dependent children (HHTYP=10-13)
- A_GE3 Three or more adults, no dependent children (HHTYP=8)
- A_GE3_DCH Three or more adults with dependent children (HHTYP=13)
- HH_NDCH Households without dependent children (HHTYP=1-8)
- HH_DCH Households with dependent children (HHTYP=9-13)
- UNK Others (not possible to determine type) (HHTYP=16)

With regard to its calculation, the following methodological issues should be taken into consideration:

1. The classification of households is not mutually exclusive. A single man aged 66, for example, is included in both the category "one adult, older than 65 years" and in the category "single person".

2. The aim of the core variable on household composition is to collect information about the size and composition of the private household to which the respondent belongs and on the relationships between household members. The social situation of an individual is at least in part a reflection of their household arrangements.

3. The place of usual residence is used as the basis of the household membership. The existence of shared expenses in the household is used to determine who is regarded as household member.

4. All persons aged less then 18 are considered as dependent children, plus those economically inactive aged 18-24 living with at least one of their parents.

The calculation of the household type variable for the respondent uses the following auxiliary variables.

- HT_D - Number of dependent children in the household
- HT_A - Number of adults in the household
- SHTD - Total number of dependent children in the household
- SHTA - Total number of adults in the household
- SRB30 - The number of Personal IDs (RB030)

The calculation of the household type variable (HHTYP) for the respondent depends on the concepts of the adult and dependent child. Below the algorithm dividing respondents to adults or dependent children is described graphically.
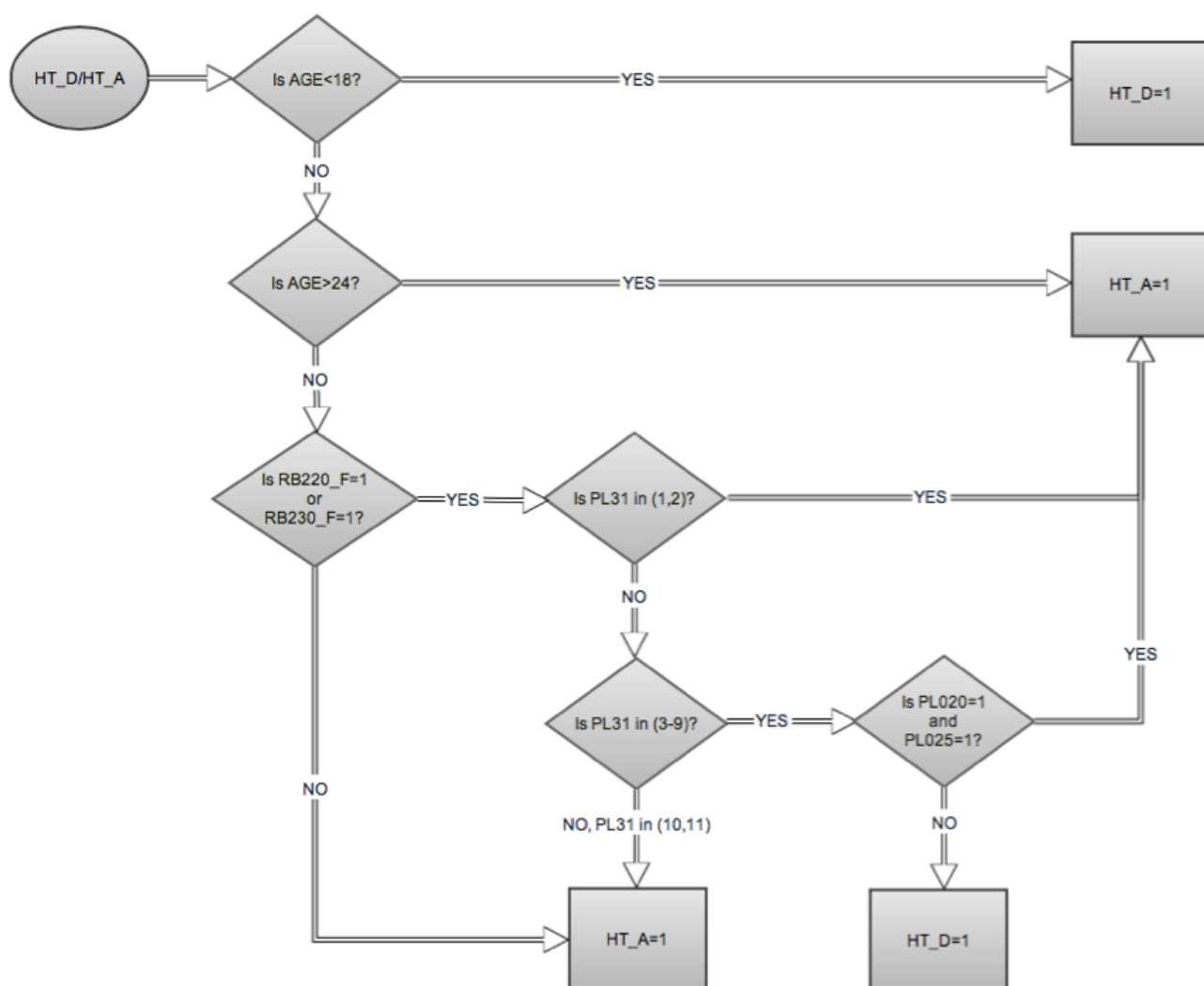
The variables HT_D and HT_A are used to derive the auxiliary variables SHTD (SHTD=sum (HT_D)) and SHTA (SHTA=sum (HT_A)). These auxiliary variables are used for the calculation of the variable household type (HHTYP):
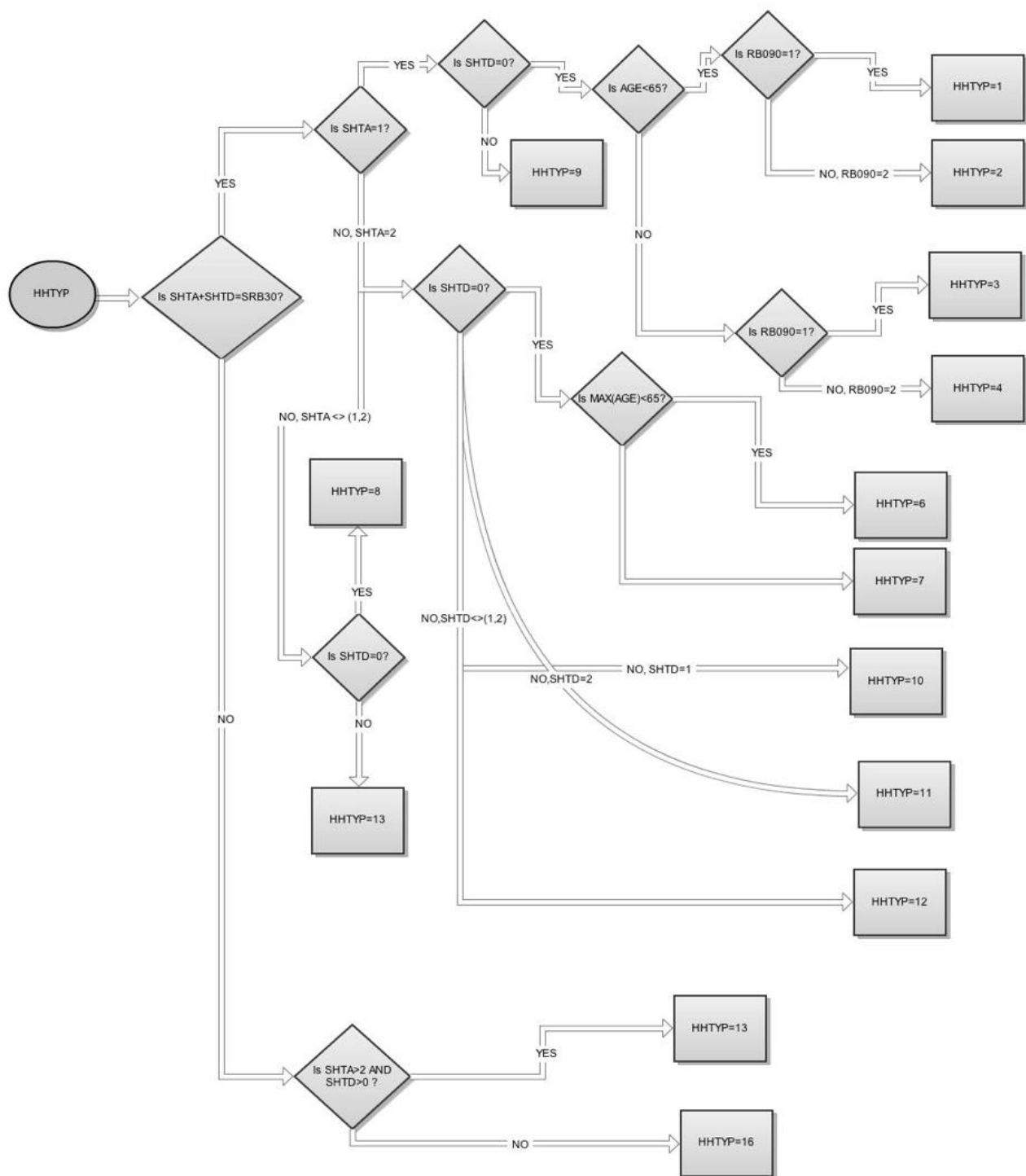
Income quantile

Income quantile groups are computed on the basis of the total equivalised disposable income attributed to each member of the household.

In EU-SILC datasets, different cut-point values (the so-called cut-off points) of income are identified, dividing the survey population into equal groups. The quantiles used in EU-SILC datasets include: the median, tertiles, quintiles , deciles, and percentiles.

Below we describe the calculation of the q – quantile interval which a person belongs to. A person belongs to the 1st q – quantile if his/her equivalised disposable income is less than or equal to the equivalised disposable income of the person with the highest equivalised disposable income within the $(1/q) \times 100\%$ of people which have the least income.

A person belongs to the kth q – quantile (0 < k<= q) if his/her equivalised disposable income is:

- less than or equal to the equivalised disposable income of the person with the highest equivalised disposable income within the $(k/q) \times 100\%$ of people which have the least income, and
- higher than the equivalised disposable income of people in $\frac{\kappa-1}{q} \, x \, 100\%$ of the population the lowest equivalised income.

The procedure for calculating the q – quantile where a person belongs is broadly similar to the procedure applied for the calculation of the median (i.e. persons will be sorted according to their equivalised disposable income (sorting order: lowest to the highest value)), but here the cut-off points will be:

$$\text{Cut-off-point}_k \; = \frac{k}{q} \, x \, 100\% \, x \, \sum_{i=1}^{n} RB050a_i$$

Where:

n = number of persons (household members)

RB050ai = is the Adjusted cross sectional weight (RB050a) for person i and k an integer satisfying the condition (0 < k<= q)
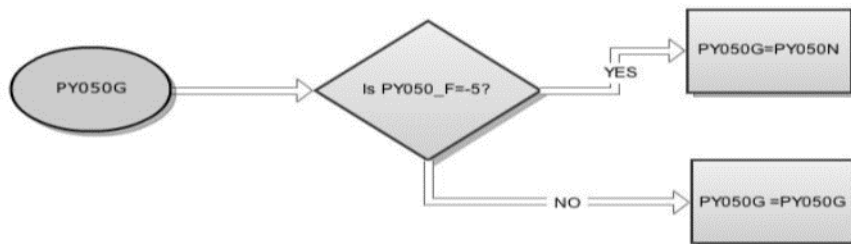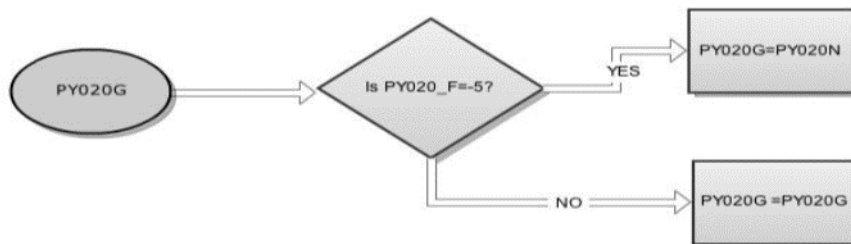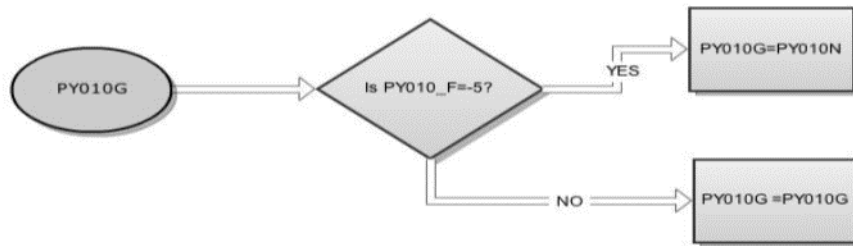
With regard to quintiles, the following methodological issues should be taken into consideration:

1. Persons have to be sorted according to their Equivalised disposable Income (EQ_INC) (after social transfers) (sorting order: lowest to highest value, household identification number and personal identification number).

2. The first quintile group represents 20 % of the population with the lowest income (an income smaller or equal to the first cut-off value), and the fifth quintile group represents the 20 % of population with the highest income (an income greater than the fourth cut-off value).

Working Income (INCWRK)

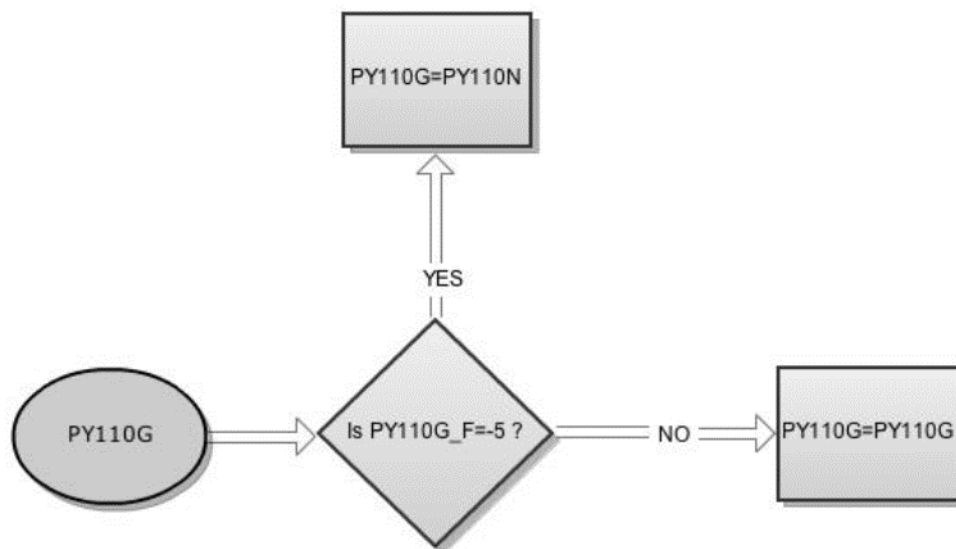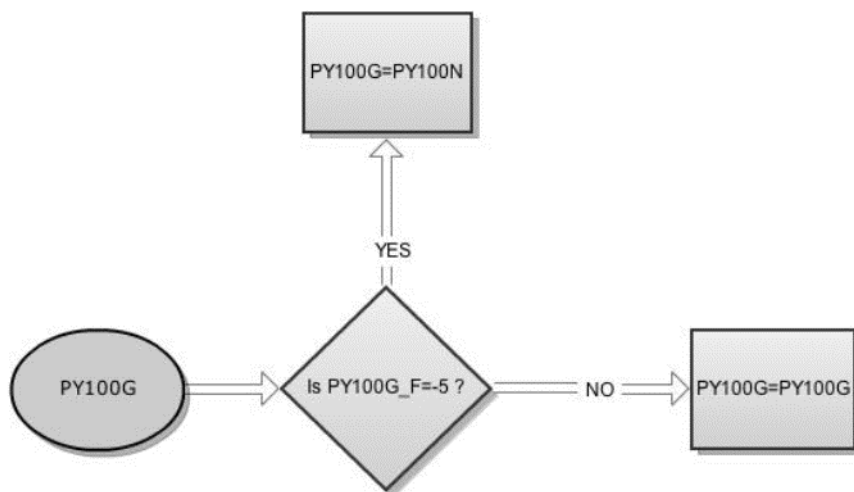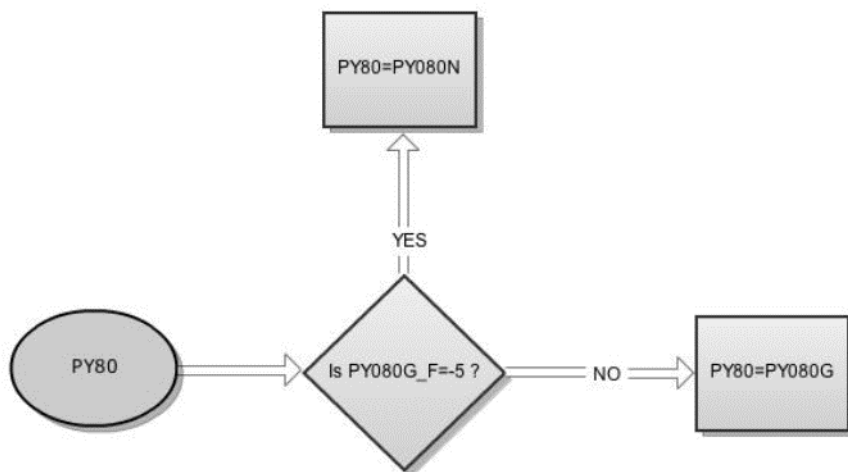The income from work variable (INCWRK) is defined as:

INCWRK = PY010G + PY020G + PY050G

PY010G → Is PY010_F=-5? → YES → PY010G=PY010N
Is PY010_F=-5? → NO → PY010G =PY010G

PY020G → Is PY020_F=-5? → YES → PY020G=PY020N
Is PY020_F=-5? → NO → PY020G =PY020G

PY050G → Is PY050_F=-5? → YES → PY050G=PY050N
Is PY050_F=-5? → NO → PY050G =PY050G

Pension Income (INCPEN)

The income from pensions variable (INCPEN) is defined as:

(INCPEN = PY080 + PY100G + PY110G

```
            ┌──────────────┐
            │ PY80=PY080N  │
            └──────────────┘
                   ▲
                   │
                  YES
                   │
  ┌────────┐   ◆ Is PY080G_F=-5 ? ◆   ──NO──▶  ┌──────────────┐
  │  PY80  │──▶                                 │ PY80=PY080G  │
  └────────┘                                    └──────────────┘


            ┌──────────────────┐
            │ PY100G=PY100N    │
            └──────────────────┘
                   ▲
                   │
                  YES
                   │
  ┌────────┐   ◆ Is PY100G_F=-5 ? ◆  ──NO──▶  ┌──────────────────┐
  │ PY100G │──▶                                │ PY100G=PY100G    │
  └────────┘                                   └──────────────────┘


            ┌──────────────────┐
            │ PY110G=PY110N    │
            └──────────────────┘
                   ▲
                   │
                  YES
                   │
  ┌────────┐   ◆ Is PY110G_F=-5 ? ◆  ──NO──▶  ┌──────────────────┐
  │ PY110G │──▶                                │ PY110G=PY110G    │
  └────────┘                                   └──────────────────┘
```

Work intensity of the household (WI)

The work intensity of the household refers to the number of months that all working age household members have been working during the income reference year as a proportion of the total number of months that could theoretically be worked within the household.

A working age is defined as a person aged 18-59, not being a dependent child. Dependent children include all persons aged below 18 as well as persons aged 18 to 24 years, living with at least one parent and economically inactive.

The following classification is used for the work intensity levels of the household in EU-SILC datasets:

- Work intensity other than very low [0.2 - 1] (NVLOW)
- Very high work intensity [0.85 - 1] (VHIGH)
- High work intensity [0.55 - 0.85] (HIGH)
- Medium work intensity [0.45 - 0.55] (MED)
- Low work intensity [0.2 -0.45] (LOW)
- Very low work intensity [0 - 0.2] (VLOW)

With regard to its calculation, the following methodological issues should be taken into consideration:

1. For each working age person (aged 18 to 64) in the household that is not classified as a dependent child, two figures are computed, using the calendar of activities of the previous year:

- the number of months in the previous year which the person has given information about his/her activity status (the 'workable' months)
- the number of months in the previous year for which the person has been classified as 'at work'

2. 'At work' comprises:

- In paid employment, whether full-time or part-time
- Including paid apprenticeship or training under special schemes related to employment
- In self-employment (with or without employees)
- Including unpaid work in family enterprise

3. Work intensity is measured at the household level.

4. When work intensity of the household can not be calculated then it is not included in the calculation of low work intensity

The calculation algorithm for the working intensity uses the following auxiliary variables:

- Hourx/Hourx2 - Total hours worked per week (constructed)
- C_mean - The mean of working hours of those who work part-time at the time of interview (constructed)
- Houratio - An estimation of part-time ratio (constructed)
- NW - Total number of workable months (constructed)
- Ne1/Ne2 - Total number of months actually worked (constructed)
- Monthratio - An estimation of the part of the year actually worked by the respondent (constructed)
- wi - The sum of month ratios of all working age members of a household (constructed)
- Size - The number of working age members of a household (constructed)
- WORK_INT - Household work intensity expressed as the average month ratio for a household (only working age household members are included) (constructed)

- LWI - Low work intensity flag

The starting point of work intensity algorithm is the calculation of the total number of hours worked per week (hourx/hourx2) for each respondent. The calculation of auxiliary variables hourx and hourx2 is presented schematically below:
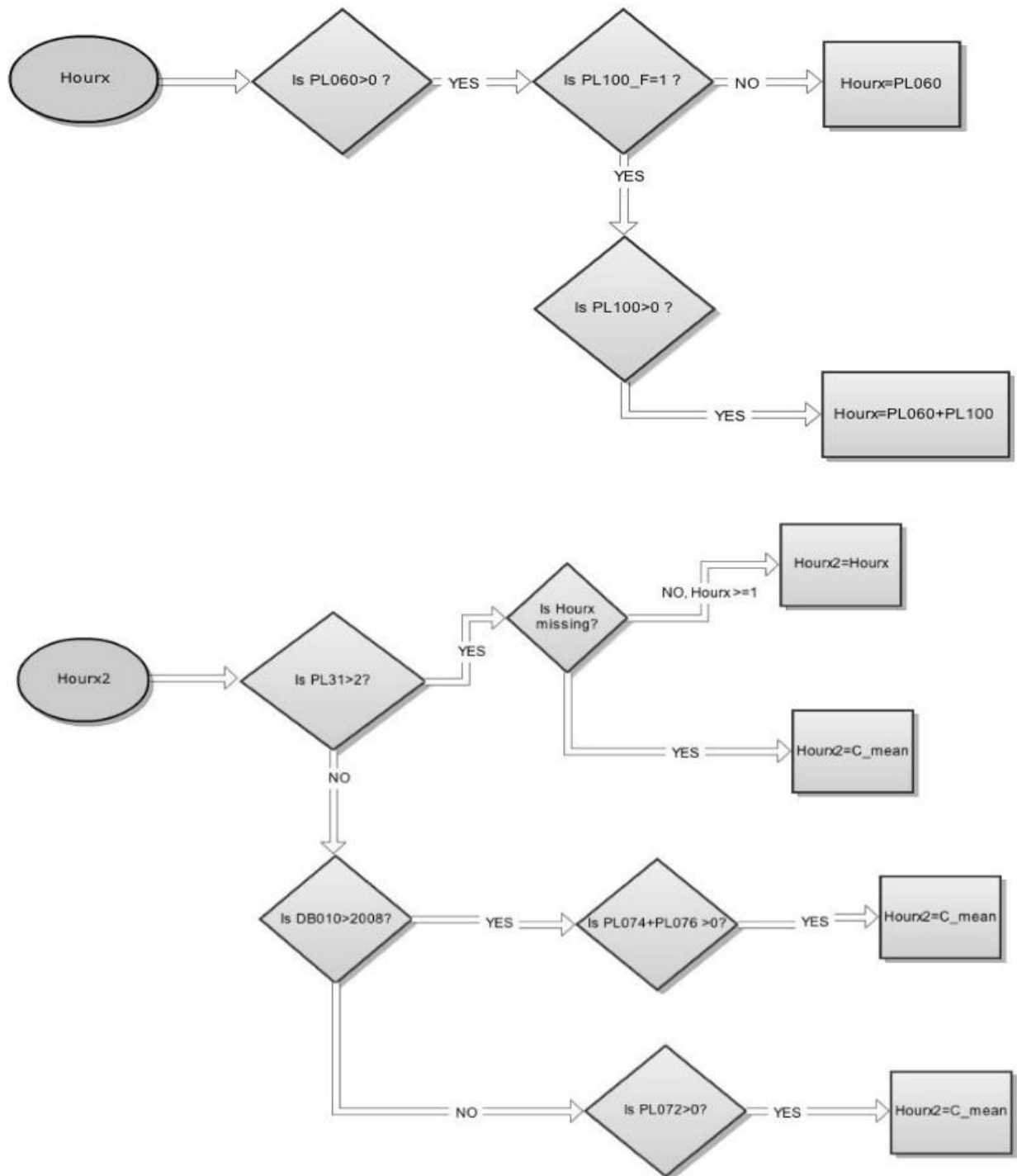
An estimation of the part-time hours ratio is needed in order to equivalise full time and part-time hours worked by the working age members of the household in order.
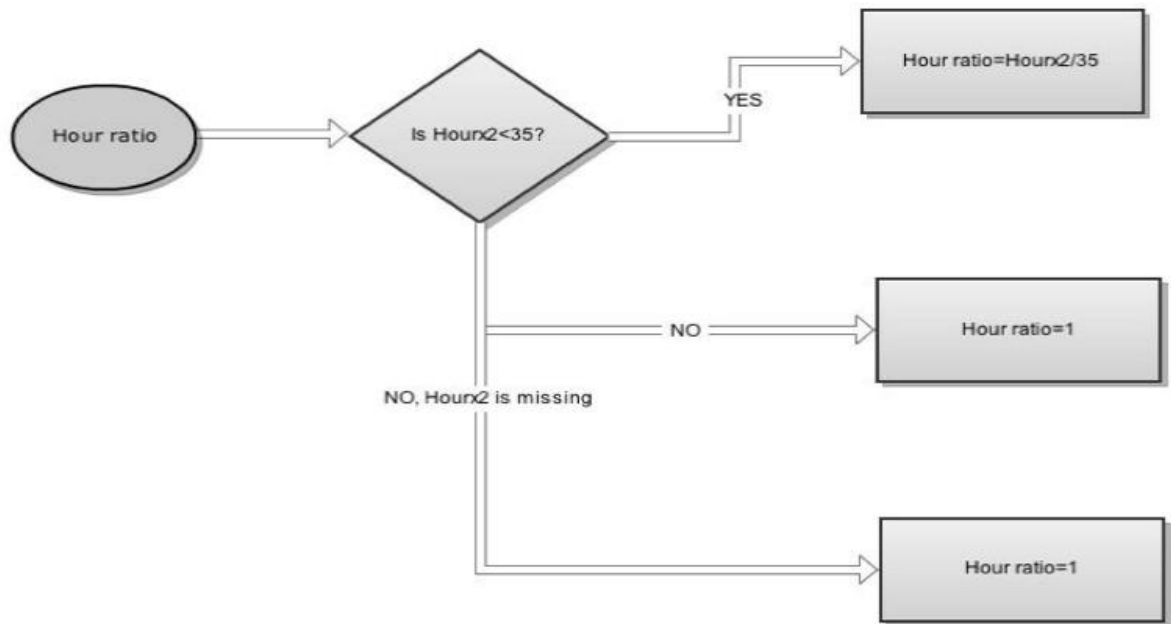
The calculation of the total equivalised months actually worked (Ne1) as well as the total number of workable months (Nw) for the working age members of the household is presented schematically below:
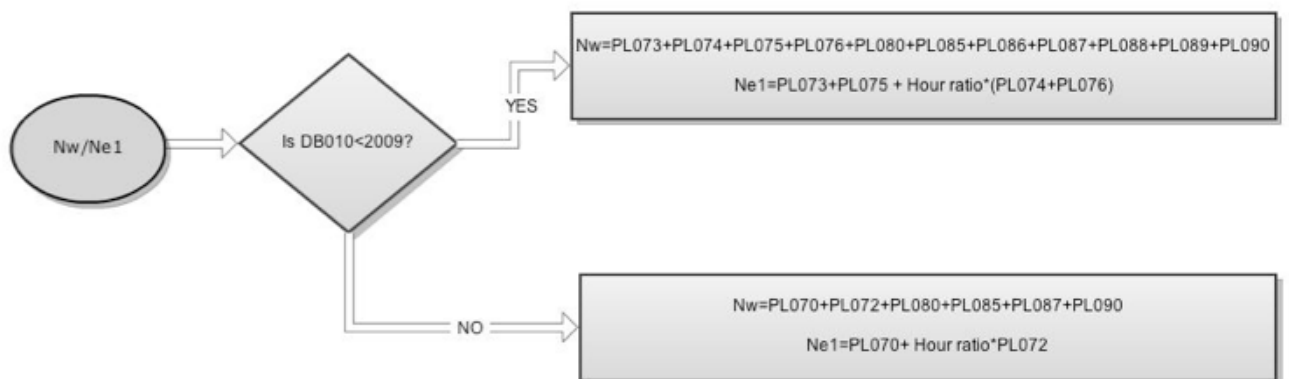
An estimation of the part of the year actually worked by each member of the household at working age can be calculated as described below:

$$month\_ratio = Ne2/12$$

The sum of month ratios for all household members at working age define auxiliary variable wi:
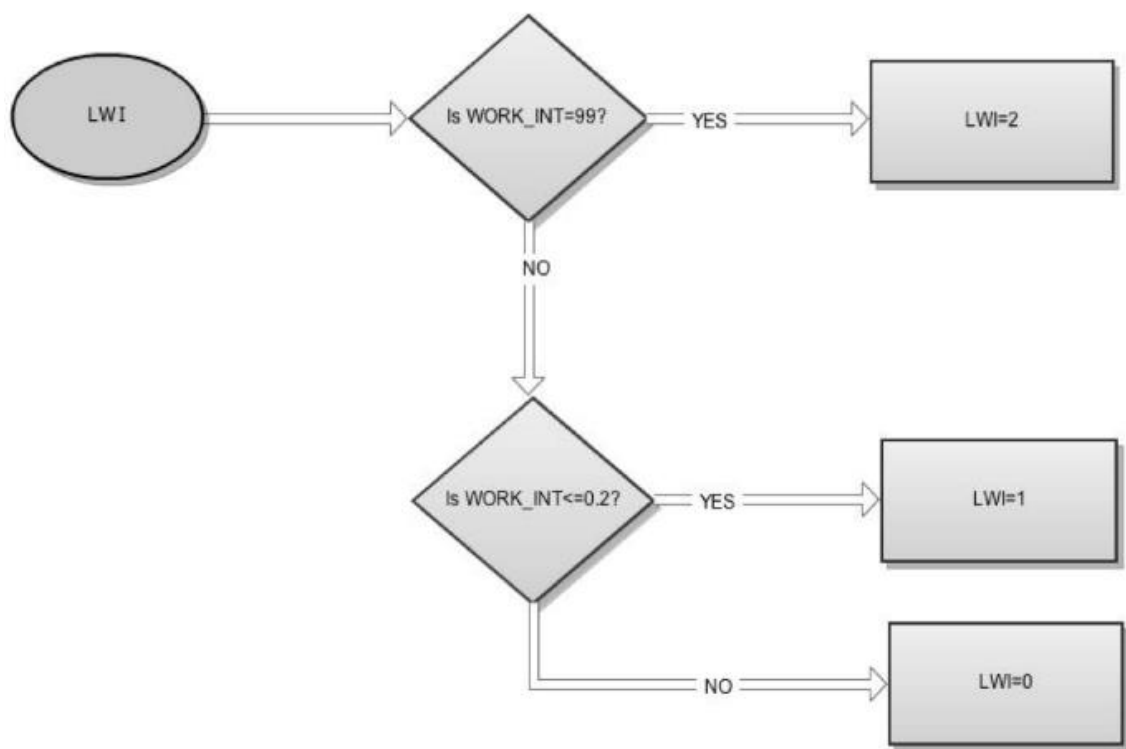
$$wi = \sum_i month\_ration\_i, i \in [1, size]$$

Finally the work intensity variable defined as WORK_INT=wi/size

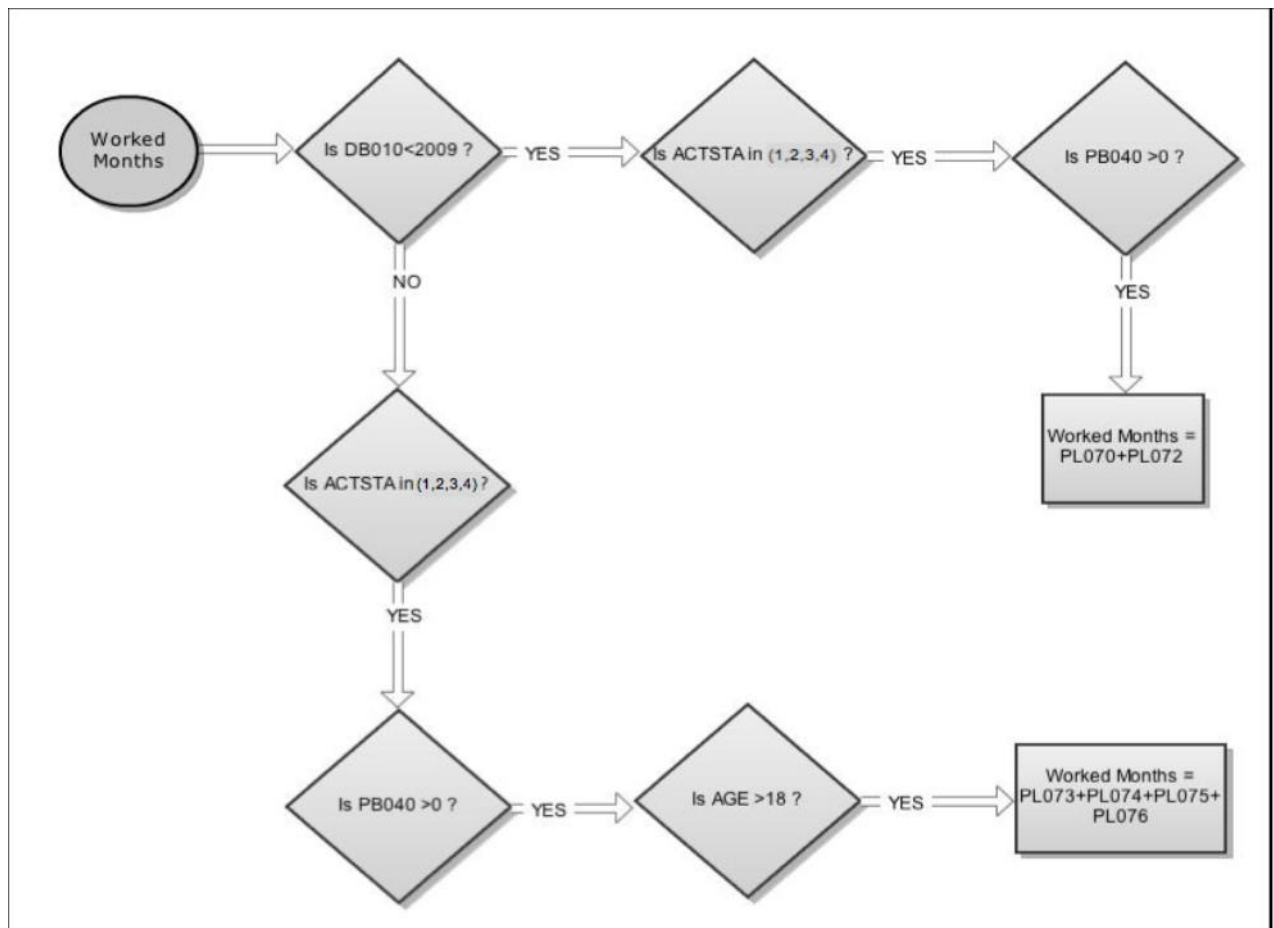If WORK_IN T> 1 then WORK_INT=1

If Age > 59 then WORK_INT=99

If WORK_INT = missing then WORK_INT=99

The work intensity variable is also used to calculate the low work intensity variable (LWI) as:

Worked Months

The calculation of the total months worked for the working age members of the household is presented schematically below:

The above calculation makes use of the derived variables Activity Status (ACTSTA) and Age.

NUTS region

The respondent's region of residence is recorded in the basic SILC variable DB040; this variable helps for the calculation of the NUTS region variable. There are two levels of aggregation for the variable NUTS:

- the NUTS1 level
    - BG3 Severna I Yugoiztochna Bulgaria
    - BG4 Yugozapadna i Yuzhna tsentralna Bulgaria
- the NUTS2 level
    - BG31   Severozapaden
    - BG32   Severen tsentralen
    - BG33   Severoiztochen
    - BG34   Yugoiztochen
    - BG41   Yugozapaden
    - BG42   Yuzhen tsentralen.
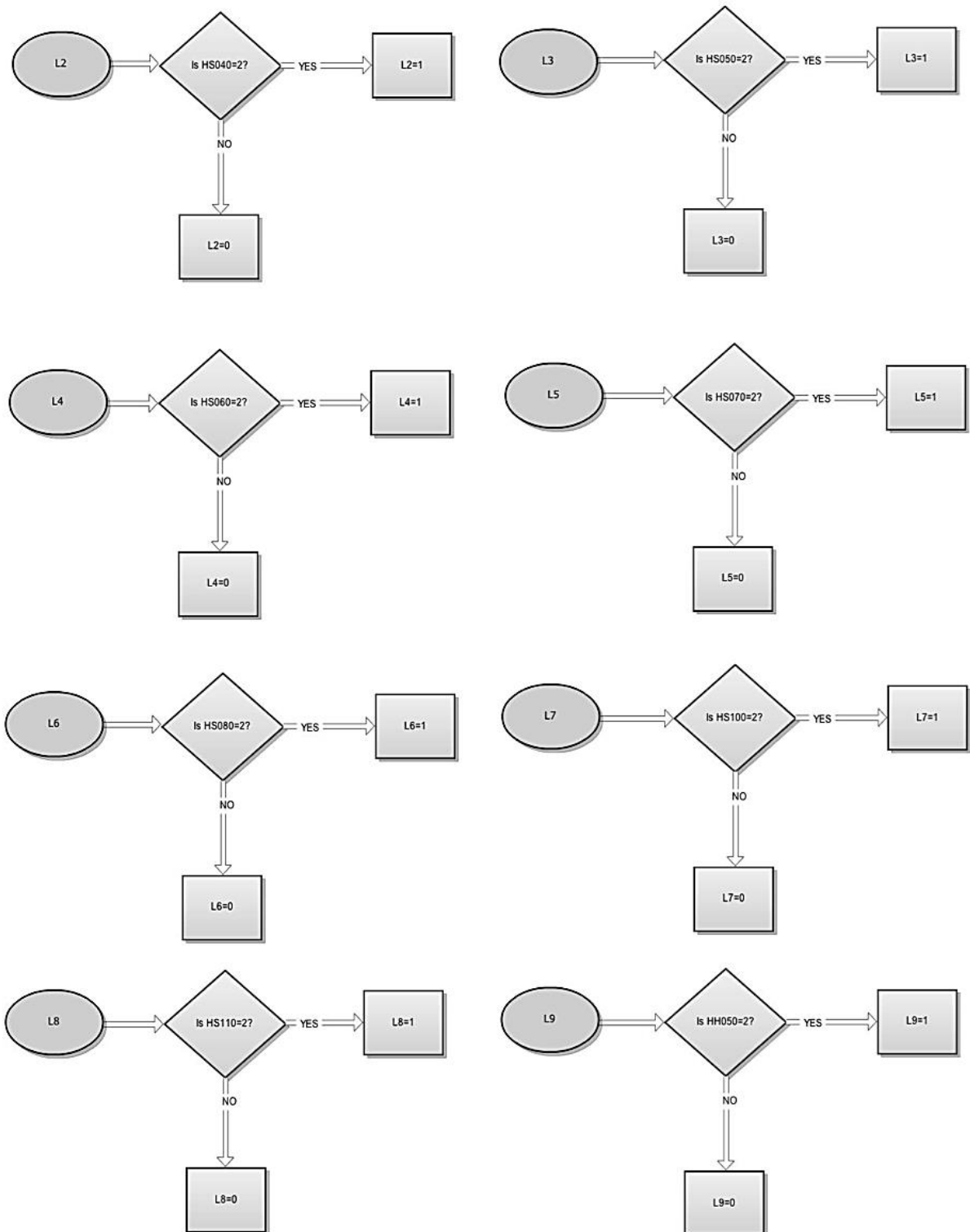
Material deprivation (MD)

The material deprivation rate refers to the situation of people who cannot afford a number of necessities considered essential to live a decent life. The nine material deprivation items considered are:

- L1-Arrears on mortgage or rent payments (basic variable HS010/HS011), utility bills (basic variable HS020/HS021), hire purchase instalments or other loan payments (basic variable HS030/HS031)
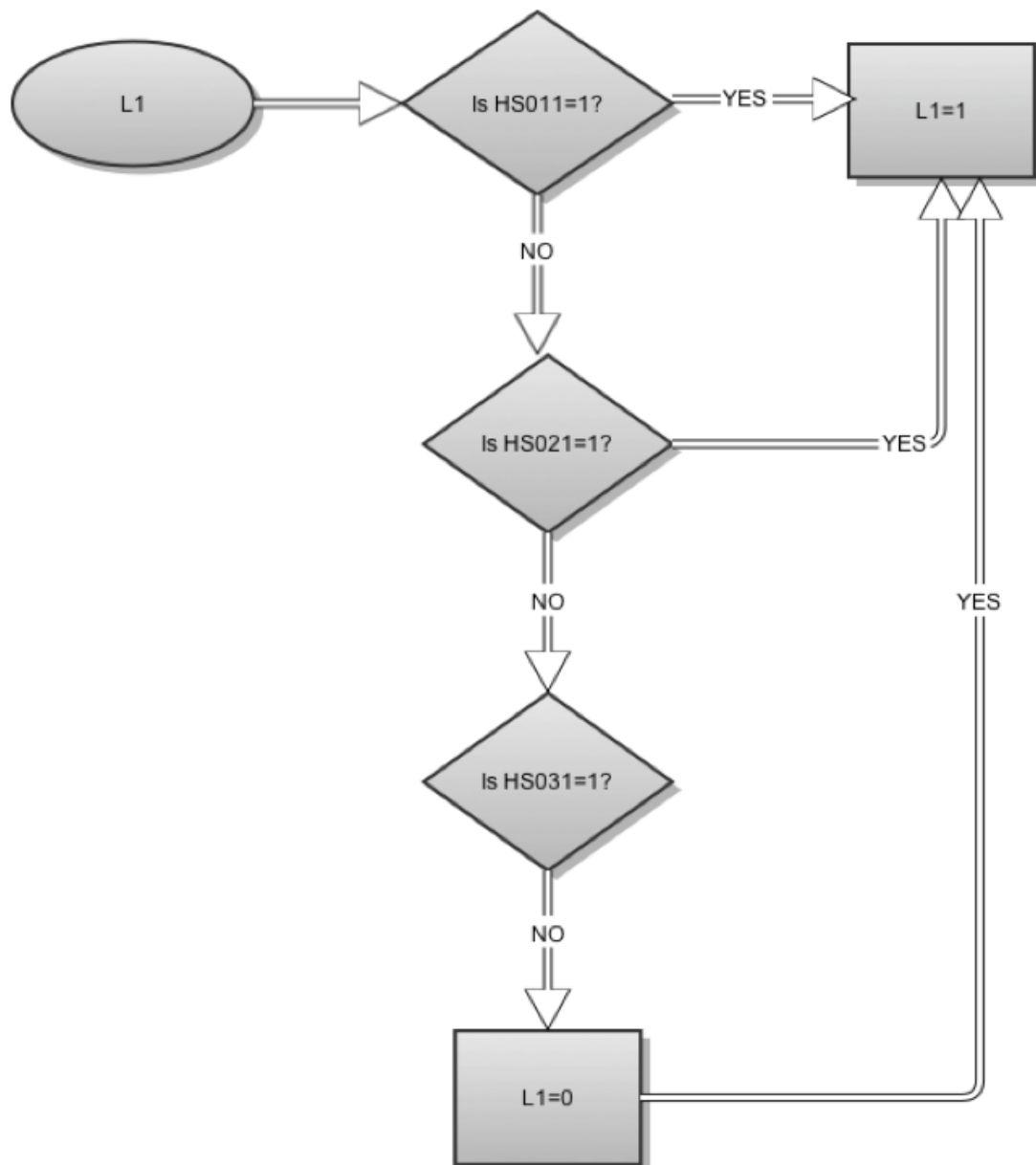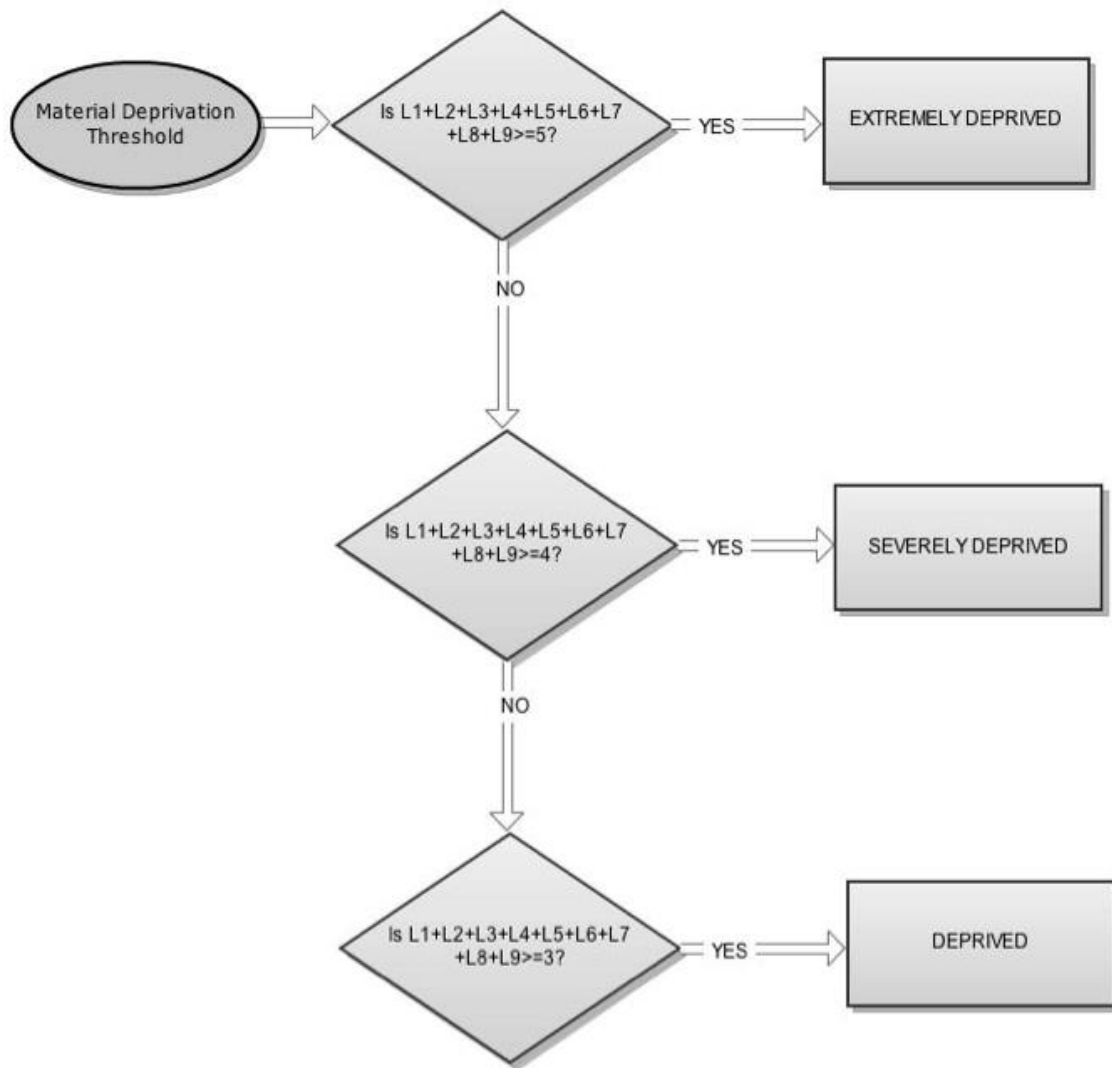

- L2-Capacity to afford paying for one week's annual holiday away from home (basic variable HS040)
- L3-Capacity to afford a meal with meat, chicken, fish (or vegetarian equivalent) every second day (basic variable HS050)
- L4-Capacity to face unexpected financial expenses (basic variable HS060)
- L5-Household cannot afford a telephone (including mobile phone) (basic variable HS070)
- L6-Household cannot afford a colour TV (basic variable HS080)
- L7-Household cannot afford a washing machine (basic variable HS100)

- L8-Household cannot afford a car (basic variable HS110)
- L9-Ability of the household to pay for keeping its home adequately warm (basic variable HH050)

Individuals are considered deprived if they have an enforced lack of at least three out of nine material deprivation items. The calculation of materially deprivation rate using the nine families of material deprivation items is presented below:

Material Deprivation Threshold → Is L1+L2+L3+L4+L5+L6+L7+L8+L9>=5? — YES → EXTREMELY DEPRIVED

NO

Is L1+L2+L3+L4+L5+L6+L7+L8+L9>=4? — YES → SEVERELY DEPRIVED

NO

Is L1+L2+L3+L4+L5+L6+L7+L8+L9>=3? — YES → DEPRIVED

As a result, 3 levels of material deprivation are obtained:

- Extremely deprived if they have an enforced lack of at least five out of nine material deprivation items;
- Severely deprived if they have an enforced lack of at least four out of nine material deprivation items;
- Deprived if they have an enforced lack of at least three out of nine material deprivation items.

Gini coefficient

The Gini coefficient measures the extent to which the distribution of income in an economy deviates from perfect (equal). It can be calculated using the Lorentz curve according to the following formula:

$$G = \frac{A}{A+B} = 1 - \frac{B}{A+B}$$

where A is the area between the line of equal distribution and the Lorentz curve and B is the area under the Lorentz curve.

The line of equal distribution of income reflects the ideal situation of equal distribution of income for all persons in the population. The closer it is to the Lorentz curve, the more evenly distributed the income is.
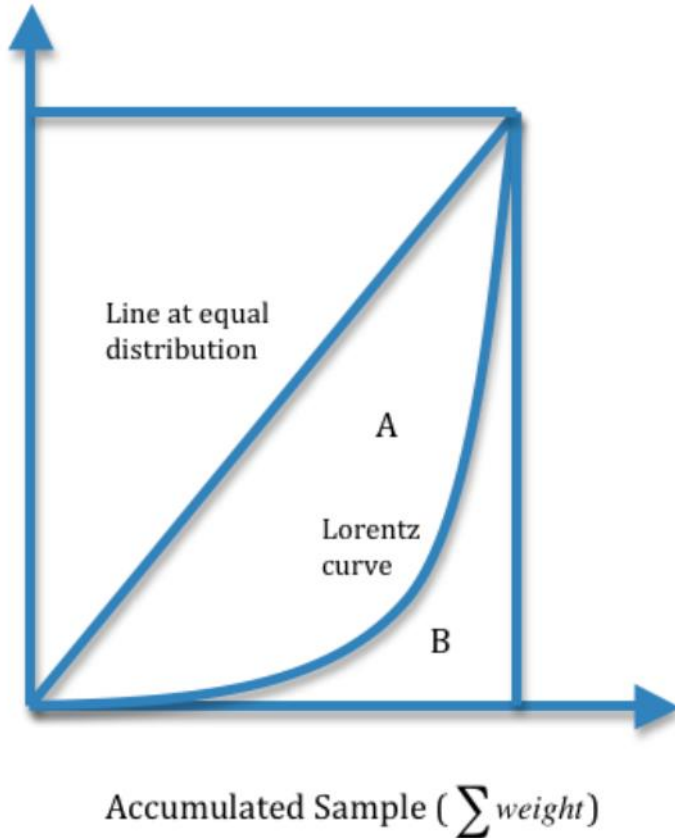
The area A+B (triangle below the line of equal distribution) will be given by:

$$A + B = \frac{\sum_{i=1}^{n}(EQ_{INCi} \times weight'i) \; x \; \sum_{i=1}^{n}(weight'i)}{2}$$

where ordinate (y) represents the equivalent disposable income of the population and abscissa (x) represents the population itself.

The area under the Lorenz curve B will be given by:

$$B= \sum_{i=1}^{n} \; (\sum_{j=1}^{i} \; EQ\_INC_i \times weight'_{i-} - 0.5 \times EQ\_INC_i \times weight'_i) \times weight'_i$$

where EQ_INCi is the equivalised disposable income of person i and weight'I is the weight for person i.

S80/S20 income quintile share ratio

The ratio of total income received by the 20% of the country's population with the highest income (top quintile) to that received by the 20% of the country's population with the lowest income (lowest quintile). Income must be understood as equivalised disposable income.

Combined indicator – At-risk-of-poverty or social inclusion (AROPE)

The indicator includes persons (as percentage of persons in the total population or as thousands of persons) who are at-risk-of-poverty (EQ_INC20<ARPT60) or severely materially deprived (SEV_DEP) or living in a household with low work intensity (WI<0.2) of the total population of the country .

With regard to the calculation of the indicators, the following methodological issues should taken into consideration:

- At-risk-of-poverty rates are assumed to be 'after social transfers' (i.e. they include social benefits such as pensions and unemployment benefits).
- At-risk-of-poverty rates relate to the at-risk-of-poverty threshold that is calculated for the total population of a member state at 60 % median equivalised disposable income level. The severe deprivation threshold is set to four and the very low work intensity threshold is 20 %.
- Income poverty risk at a given point in time may not necessarily imply low living standards in the short term, for example if the persons at risk have access to savings, to credit, to private insurance, tax credits, to financial assistance from friends and relatives etc. In particular, the cumulative impact of extended periods at risk is to be further assessed.
- Measuring incomes at the level of private households may have certain implications. The exclusion of collective households might lead to an underrepresentation of certain groups (e.g. the elderly, persons with disabilities).