

## **Technical Report BTUS13.**

### *Technical report of the 2013 Belgian Time-Use Survey (BTUS13)*

Ignace Glorieux<sup>1</sup>, Joeri Minnen<sup>1</sup>, Theun Pieter van Tienoven<sup>1</sup>, Camille Vanderhoeft<sup>2</sup>, Sarah Daniels<sup>1</sup>, Kyra de Korte<sup>1</sup>, Jef Deyaert<sup>1</sup>, Jeroen Magits<sup>1</sup>, Eszter Mészáros<sup>1</sup>, Julie Verbeylen<sup>1</sup>, Djiwo Weenas<sup>1</sup>

<sup>1</sup> Research Group TOR, Department of Sociology, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium // [www.vub.ac.be/tor](http://www.vub.ac.be/tor)

<sup>2</sup> Statistics Belgium, North Gate, Koning Albert II-laan 16, 1000 Brussel // [www.statbel.fgov.be](http://www.statbel.fgov.be)

---

## Index

1	Introduction, objectives and history of Belgian Time Use Research.....	3
2	Fieldwork and sample design.....	4
2.1	Interviewers.....	4
2.2	Reference period.....	4
2.3	Population and sampling frame.....	4
2.4	Diary days.....	4
2.5	The BTUS13 sampling design.....	5
3	Measurement instruments.....	7
3.1	LFS.....	7
3.2	Diaries.....	7
3.3	Work Grid.....	8
3.4	Drop-off questionnaires.....	10
4	Coding of the diaries and cleaning of the data.....	11
4.1	Coding of the diaries.....	11
4.2	Data cleaning: checks and quality control.....	12
4.3	Cleaning of the Work grid, Labour Force Survey and drop-off questionnaires.....	15
5	Response.....	16
6	Weighting and representativity.....	17
6.1	Integrated weighting procedure.....	17
6.2	Post-stratification weighting procedure.....	20
7	Overview of final datasets.....	25
8	List of abbreviations.....	26
9	References.....	27

# 1 Introduction, objectives and history of Belgian Time Use Research

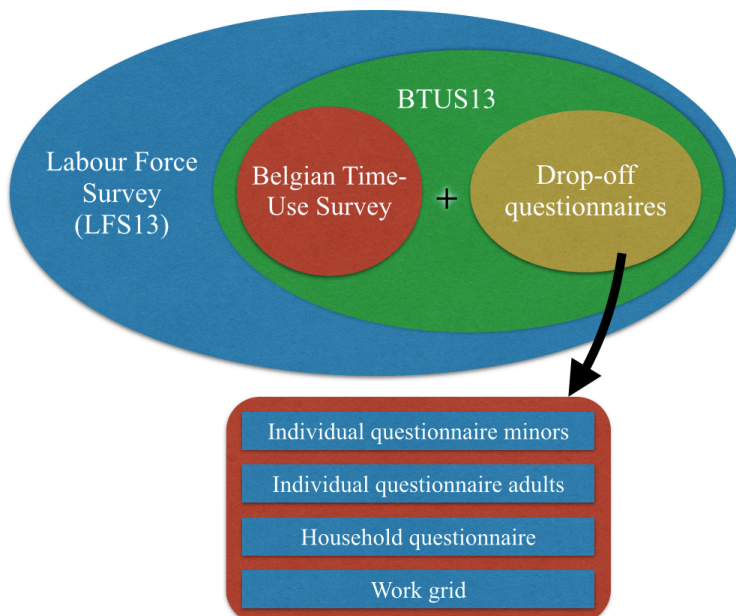
Since 1999, Statistics Belgium has been conducting Time-Use Surveys together with the research group TOR as its advisory partner. The data collection was administered and financed by Statistics Belgium, the data cleaning was administered by the research group TOR (financed by the Belgian Federal Science Policy with an AGORA- and BRAIN.be-grant).

In Time Use Research, respondents are asked to keep record of their daily activities for a number of days by means of a diary. Besides filling in a diary, they also complete additional questionnaires (see Figure 1), which enable researchers to link daily behaviour with socio demographic characteristics, labour market conditions, health conditions, well-being, attitudes, ... For the Belgian Time Use Surveys of 1999, 2005 and 2013 (hereafter BTUS99, BTUS05 and BTUS13) respondents were given a paper and pencil diary with the instruction to describe in their own words what they were doing, when they were doing this activity, where they were during this activity and who was present. These diary entries were post coded towards a fixed list of activities by trained coders of Statistics Belgium.

The BTUS99 data include 8,382 respondents from 4,275 households, the BTUS05 data 6,400 respondents from 3,474 households, and the BTUS13 5,559 respondents from 2,744 households. Further information about the 1999 and 2005 Time-Use Surveys can be found in the technical report of BTUS99 and BTUS05 (Glorieux, Minnen, & Mestdag, 2007).

In contrast to BTUS99 and BTUS05 and the coupling with the Household Budget Study (HBS), the BTUS13 was coupled with the Labour Force Survey. The sample of BTUS13 was drawn from the Belgian 2013 Labour Force Survey (hereafter LFS13) participants. This means that we have information from the LFS13, the BTUS13 and drop-off questionnaires for all participants.

**Figure 1. Structure of BTUS13 data**



## 2 Fieldwork and sample design

The BTUS13 sample was drawn from LFS13. About 1/4<sup>th</sup> to 1/3<sup>rd</sup> (depending on the quarter, as explained hereafter) of the household groups in LFS13 were invited face to face by the LFS interviewers to participate in BTUS13. When households agreed to participate, everyone living in the household of 10 years and older was asked to register all his or her activities during two specified days (one weekday and one weekend day) and the interviewer gave detailed instructions about how to do that. The interviewer visited the household for a second time to check and collect the completed diaries.

### 2.1 Interviewers

In total, about 200 interviewers were thoroughly trained to execute the fieldwork of BTUS13. All interviewers who interviewed for BTUS13 had to take an obligatory training which included the following topics: objectives and procedure of BTUS13, content of the questionnaires and practical questions.

### 2.2 Reference period

The fieldwork for BTUS13 ran from mid January 2013 until February 2014. This, of course, corresponds – with a short delay – to the reference period for LFS13, on which BTUS13 is based to sample and recruit households.

### 2.3 Population and sampling frame

The study or target population for BTUS13, as for LFS13, is the residing Belgian population living in private households. This population is restricted to households where at least one member is between 15 and 76 years old, to form the sampling frame for LFS13, and therefore also for BTUS13.

### 2.4 Diary days

After completing the LFS13 questionnaires, households were asked if they were willing to participate in BTUS13. If a household agreed, all household members of at least 10 years old were given two diaries, one for a weekday and one for a weekend day, as well as a drop-off questionnaire and a work grid for the working household members. In order to ensure a balanced spread of all possible ordered combinations of one weekday and one weekend day, Statistics Belgium developed an algorithm to automatically determine such a combination randomly as soon as the household agreed to collaborate in BTUS13. Finally, all household members were asked to complete the two diaries at the same two diary days. After the combination of weekday and weekend day was communicated to the household, household members could delay this specific combination with a maximum of three weeks<sup>1</sup>.

---

<sup>1</sup> E.g. imagine the interviewer communicates to the household they have to fill in their diaries on Thursday February 7<sup>th</sup> and on Saturday February 9<sup>th</sup>, they could delay this ordered Thursday-Saturday combination with a maximum of three weeks.

To complete the two diaries, respondents were asked to describe in their own words what they did on a specific day; based on the time slots of 10 minutes in the diaries, the duration for each recorded activity can be calculated. For analysis of duration of activities, more specifically for comparison of durations across weekdays and/or weekend days, it is important that, in the final BTUS13 respondents sample, all ordered weekday-weekend day combinations are sufficiently represented. In section 6.2 we describe how the respondents' weights are adjusted for observed imbalances in diary day combinations.

## **2.5 The BTUS13 sampling design**

### **2.5.1 The LFS13 sampling design**

Since the BTUS13 sample of households is a subsample of the LFS13 sample of households, we first describe briefly the LFS13 sampling design. For LFS13, a sample of private households was drawn from a sampling frame as described in section 2.3. This sample is drawn in two stages.

In the first stage, geographical units (generally being *sections* or *neighbourhoods* within the sub municipalities of the current 589 Belgian municipalities) are selected randomly with probability proportional to their size (i.e. the number of private households). Each quarter, 624 draws of these geographical units are made. The 624 draws are spread over the regions as follows:

- in the Brussels Capital Region: 91 draws;
- in the Flemish Region: 292 draws
- in the Walloon Region: 241 draws – among which 21 are from the German speaking Community.

It is worth mentioning that the 624 draws are uniformly spread over the 52 reference weeks in the reference year. Hence, 48 draws are assigned to each reference week.

In the second stage, for each draw of a geographical unit, a group of households from this geographical unit is selected randomly. The a priori defined group sizes are 23 for the Flemish and Walloon Regions, and 26 for the Brussels Capital Region. This amounts to 14625 households selected each quarter into the LFS13 initial sample; 483 of these households belong to the German speaking Community.

In each selected household, all members aged 15 years or more have to complete LFS13 questionnaires. The initial sample of 14625 households includes about 28000 individuals aged 15 years or more.

### **2.5.2 Sampling for BTUS13 from the initial LFS13 sample**

After 2-stage sampling of households for LFS13, a third stage is added to complete the BTUS13 sampling design. From the initial LFS13 sample of households, households are selected for BTUS13 as follows.

First, the reference weeks (being important time periods for LFS) within each quarter are transformed into months (being more practical time periods for BTUS). Each LFS13 group of households is thus assigned a month. Next, a number of households to be selected from the initial LFS13 sample for each combination of month and province (with the Brussels Capital Region as a separate “province”, and with exclusion of households in the German

---

speaking community) is calculated such that approximately the same number of households will be selected for each month, and such that the sampling fractions in the different provinces are roughly equal.

Furthermore, for quarters 1 and 2 (January – June 2013) each:

- a sub-sample was drawn from the LFS13 sample of groups of households; the overall third stage sampling fraction is about 27%;
- this results in a total of 163 groups of LFS13 households selected into the initial BTUS13 sample; households participating into LFS13 are actually invited to participate in BTUS13.

Similarly, for quarters 3 and 4 (July 2013 – February 2014) each:

- a sub-sample was drawn from the LFS13 sample of groups of households; the overall third stage sampling fraction is about 33%;
- this results in a total of 205 groups of LFS13 households selected into the initial BTUS13 sample; households participating into LFS13 are actually invited to participate in BTUS13.

All household members of 10 years and older were asked to fill in a diary for two days (one weekday and one weekend day); see section 2.4.

### 3 Measurement instruments

#### 3.1 LFS

The LFS is a socio-economic household survey, with as its main goal to classify the population of working age<sup>2</sup> into three groups: employed, unemployed and inactive persons, and to provide descriptive and explanatory data on every category. This survey belongs to the community sample surveys on active population that are coordinated by the Statistical Office of the European Communities, EUROSTAT. The purpose is obtaining information that is comparable at European level on the employment and unemployment rates according to the definitions of the International Labour Organization (ILO) but also collecting and disseminating results that are not available elsewhere (mobility of the workers, reasons for working part-time, education level of the population, profession, ...). More information about the LFS can be found on the website of Statistics Belgium: <http://statbel.fgov.be>.

#### 3.2 Diaries

The central instrument of Belgian time use research consists of paper-and-pencil diaries, consciously filled in by the respondents. The EUROSTAT driven *Harmonized European Time Use Survey* (hereafter: HETUS) model was used to design the diaries and code the open descriptions recorded by the respondents, see HETUS 2008 guidelines<sup>3</sup> (Eurostat, 2008). In BTUS13 a more extensive activity list was used than prescribed by HETUS, though the same structure and topics were covered, which makes BTUS13 perfectly suited for comparative international research.

All members of the household of 10 years and older were invited to register their daily activities in a diary during two days (one week and one weekend day) from 04:00 am until 04:00 am the day after. Respondents were asked to indicate the begin and end time of each activity in 10-minute blocks. They could extend this block, by simply drawing an arrow or using quotation marks (“”) from the start until the end of the activity (example in figure 2). In these blocks they had to describe in their own words what they were doing (main activity), what they were doing besides this main activity (secondary activity), if they used the Internet during main or secondary activity, their location or transportation mode. Through a checkbox they had to indicate the presence of other persons. All members of the household filled in their diaries on the same days.

---

<sup>2</sup> 15 years and older

<sup>3</sup> These guidelines can be downloaded via: <http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-08-014-EN.pdf>

**Figure 2. Example of a diary template (extract from the 2008 HETUS guidelines (Eurostat, 2008, pp. 200 - 202)):**

Time	<b>What were you doing?</b> <i>Record your main activity for each 10-minute period from 04.00 to 07.00!</i>  Only one main activity on each line! Distinguish between travel and the activity that is the reason for travelling.	<b>What else were you doing?</b> <i>Record the most important parallel activity.</i>  Indicate if you used, in the main or parallel activity, a computer or internet. You do not need to record the use of a computer or internet during working time.	<b>Where were you?</b> <i>Record the location or the mode of transport</i>  e.g. at home, at friends' home, at school, at workplace, in restaurant, in shop, on foot, on bicycle, in car, on motorbike, on bus, ...	<b>Were you alone or together with somebody you know?</b>					
				Mark "yes" by crossing					
				Alone	With other household members				Other persons that you know
					Partner	Parent	Household member up to 9 years	Other household member	
07.00-07.10	Woke up the children		At home	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
07.10-07.20	Had breakfast	Talked with my family	↓	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
07.20-07.30	--"	--"	↓	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
07.30-07.40	Cleared the table	Listened to the radio	↓	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
07.40-07.50	Helped the children dress	Talked with my children	↓	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
07.50-08.00	Went to the day care centre	--"	On foot	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
08.00-08.10	Went to work	Read the newspaper	Bus	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
08.10-08.20	Went to work	--"	--"	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
08.20-08.30	Work		Workplace	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
08.30-08.40			↓	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
08.40-08.50			↓	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
08.50-09.00			↓	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
09.00-09.10			↓	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
09.10-09.20			↓	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
09.20-09.30			↓	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
09.30-09.40			↓	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
09.40-09.50			↓	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
09.50-10.00			↓	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Note: a more extensive example can be found in the codebook of BTUS13 (see 1.6 Original questionnaire (diary)).

### 3.2.1 Contextual information diary days

After completing the diaries, respondents had to answer several contextual questions regarding the type of day (if it was a normal day, if they were hurried that day, ...). More information about these questions can be found in the codebook of BTUS13 under paragraph 1.7.1 Codebook background information filled in diary days.

### 3.3 Work Grid

Each household member of 18 years and older with a paid job (employee, self-employed, assisting household member) was asked to fill in a seven-day work grid. The first registration day in the work grid is the first day respondents have to fill in a diary. Respondents had to indicate their working times by drawing a line on a 24h horizontal grid, which consists of 15-minutes slots, and starts at 4am. They were also asked to indicate whether it was a normal week or an unusual week (due to temporarily absence at work or other reasons).





### **3.4 Drop-off questionnaires**

Every household completed a household questionnaire, the individual members of the household filled in an individual questionnaire, with a specific version for minors (10 – 17 years old).

#### ***3.4.1 Individual questionnaire adults***

Following topics were covered:

- Information about eating and cooking habits
- Leisure time
- Organisation and regularity of weekdays (Monday to Friday)
- Satisfaction with life and life environment
- Personal information

#### ***3.4.2 Individual questionnaire minors (10 – 17 year)***

Following topics were covered:

- Information about eating and cooking habits
- Personal planning
- Satisfaction
- Personal information

#### ***3.4.3 Household questionnaire***

Following topics were covered:

- Possession of durable goods
- Revenues of the household
- Use of childcare services

## 4 Coding of the diaries and cleaning of the data

### 4.1 Coding of the diaries

All participants completed a two-day diary and described their daily activities in their own words. After the interviewer controlled the diaries at the respondents home, a group of trained coders of Statistics Belgium coded these open descriptions towards a closed activity list based on the 2008 HETUS guidelines (Eurostat, 2008).

#### 4.1.1 Activity list

The full list of activities of BTUS13, as well as their concordance with the HETUS typology, can be found in the codebook of BTUS13 under paragraph *1.3 Activity list diary*.

#### *Hierarchical classification of activities*

In order to be able to report on time use research in a comprehensive way, the activity codes attributed by the coders to the open descriptions of the respondent – most often – provide too much detail. Also, when time use research is compared over years or between countries, the most detailed level of activity codes is often (a bit) different. That's why we structure our activity codes hierarchically. Codes directly attributed by the coders are called *level-3 codes*. When we group these codes in subgroups, we get activity codes on a more general level, which we call *level-2 codes*. The most general level of activity codes are *level-1 codes*.

Since 1999 the research group TOR uses a hierarchical classification of activities that is called the *TOR level-2 & 1 typology*. The structuring of these activities into this typology is thoroughly described in the BTUS13 codebook under paragraph *1.4 assignment level-3 activity codes BTUS99, BTUS05 BTUS13*.

#### *Translation of activities towards other typologies (TOR13 / HETUS / Eurostat codes)*

The HETUS guideline of 2008 includes a 3-level pre-defined hierarchical structure for activity codes. For reasons of comparability, we translated all BTUS13 level-3 activity codes towards:

- The TOR level-1 & 2 typology.
- The HETUS codes. Starting from the HETUS level-3 codes, it's very easy to deduce the HETUS level-2 & 1 hierarchical codes: the first digit of a level-3 code indicates the first level and the second digit the second level (e.g. level-3 *311: food preparation, baking and preserving* is categorized under level-2 *food management*, which resides under level-1 *household and family care*).
- The BTUS99 & BTUS05 EUROSTAT codes. These activity codes were used in the diary files of 1999 and 2005. In order to be able to compare on the most detailed level, we translated the HETUS based BTUS13 level-3 activity codes towards the EUROSTAT codes.

An extensive overview of these translations can be found in the BTUS13 codebook under paragraph *1.4 assignment level-3 activity codes BTUS99, BTUS05 BTUS13*.

## 4.2 Data cleaning: checks and quality control

During the coding and construction of a dataset by Statistics Belgium, several quality controls were integrated to guarantee qualitative diary data. After this process, one researcher of the Vrije Universiteit Brussel was entitled by Statistics Belgium to perform a thorough check of all information included in the diaries.

Our initial diary database consisted of 5,678 respondents. We started with general quality checks, followed by more in-depth quality controls.

During the cleaning, we kept a *list of problematic respondents*. Whenever there was a problem with one of the quality checks below or we had doubts about the credibility of a specific diary, the respondent appeared on this list. After the computer cleaning was fully performed, the diaries of the *problematic* respondents<sup>4</sup> were physically examined.

### 4.2.1 Quality checks on respondent level:

1. *Does every diary day start and end at 04:00h?*
  - Yes, all diaries start and end at 04:00h
2. *Do the data & time variables 'begdatetime' and 'enddatetime' in the activity file (TUS13-dagb.sav) display the correct dates?* E.g. if a respondent goes to the toilet at 01:20h, the date variable has to indicate the previous date + 1. In the initially delivered dataset by Statistics Belgium, all date variables indicated the date that the respondent had to register, but included no date-correction for activities starting or ending after midnight.
  - This was changed in the final diary dataset.
3. *Does every diary day contain 24 hours?* In case there are open time slots, we inserted unspecified time.
  - Note: for only two respondents, there was an open time slot of 10 minutes. Based on a physical check of the diaries at Statistics Belgium, we were able to solve this error in a correct way for both respondents.
  - Note: when we aggregated our activity files, we noticed that 5 respondents had diaries of 96 hours (2 x 48h). A detailed check of the respondent numbers and household questionnaires revealed a small error in the respondent numbers (probably a typo of the coders). We adapted these respondent numbers in all databases. Of these 5 respondents, two survived our cleaning procedure and are in the final dataset.
4. *Did every respondent fill in a weekday and a weekend day?*
  - Yes, all respondents filled in one week and one weekend day.

---

<sup>4</sup> More or less 90 diaries.

5. *Are all activities registered in a sequential order?* In other words: do the begin and end times of consecutive activities connect with each other?
  - Yes, after completing the two open time slots of check 3.
6. *Did all members of a household fill in their diaries on the same days?*
  - No, there are respondents with only one day matching with the rest of the household, and respondents with two non-matching days.
    - In the original, non-cleaned file, there are 48 respondents with at least one non-matching day with the rest of the household.
    - Of these 48 respondents, 15 respondents had only one non-matching day with the rest of the household and 33 respondents with 2 non-matching days.
    - After all cleaning procedures, we performed this check again, and in the final database (5559 respondents) there are 38 respondents with at least one non-matching day with the rest of the household. Of these, 11 respondents had only one non-matching day with the rest of the household and 27 two non-matching days.
    - We marked them in the final database with the variable ‘datum\_gezinsleden’ (see also codebook of BTUS13).
7. In those situations where respondents gave a vague description of their activities and coders had to interpret the specific activity (making use of context variables, presence of meaningful others, ...), a specific code<sup>5</sup> was used. E.g.: when a respondent forgot a transportation from / to work, and the coder interpreted this as *unspecified*, an extra code was created (instead of using code 9100: *Travel from / to work* code 9101: *Travel from / to work (filled in by coder)* was used). We controlled the diaries in the database, and – if necessary – checked the physical diaries at Statistics Belgium.
8. When the coders had some doubts, they noted their thoughts in an extra column in the diary data (see variable ‘Comm\_Act’ in the diary file: TUS13-dagb.sav). We noticed during the cleaning that some activities were coded as ‘unspecified’, even though coders gave valuable information about the activity in the ‘Comm\_Act’ column. These activities were re-checked by experienced coders of Statistics Belgium, and re-coded whenever possible or suitable.
9. In specific cases where the coder indicated a certain time slot as: 9992: *forgot an arrow*<sup>6</sup> or 9991: *unclear activity description*, we checked the temporal location of this activity. When these activities were located between 21:00h and 04:00h, within two episodes of sleep, we checked all context variables of this activity. In case we were almost certain the respondent was sleeping, we recoded these time slots as 110: *sleeping*.

---

<sup>5</sup> These codes were: 0111 9101 9201 9361 9381 9391 9401 9501 9601 9801 9001, labels can be found in the Codebook of BTUS13.

<sup>6</sup> To avoid that respondents had to note an activity for each 10-minute episode in the diary activity file, they could prolong an activity by simply drawing (see codebook BTUS13 for a good example of a filled in diary)

10. Did respondents register more than two hours of unspecified time use (codes 9980, 9990, 9991 and 9992) per day? These diaries were thoroughly checked and in case the quality of the diaries was too bad, respondents were removed from the diary dataset.

#### 4.2.2 *Quality checks on an aggregated level:*

The following quality checks were performed on a more general aggregated level, making use of Z-scores. Given we are working on the activity diary file, these scores show the distance between a specific value and the mean of all respondents.

1. *Do respondents register enough activities, so that we're able to reconstruct their day in a meaningful way?*
  - Our experience with cleaning diary data is that respondents with diaries of doubtful quality often register few activities. During this check we constructed a variable per respondent which indicated the number of activities a respondent performed on a specific day. Next, we calculated Z-scores for each respondent. We thoroughly checked respondents with an absolute Z-score (for the number of daily activities) higher than 3.
2. *Do respondents register very long or very short activities compared to all other respondents?*
  - Here, we aggregated our data based on the three-level structure of the diary codes (see further). We checked per day (such that we checked week and weekend days separately<sup>7</sup>) and calculated Z-scores for all activity codes on the three levels. This enabled us to check respondents with extraordinary long or short durations on specific activities (or, in case we looked at level-2 and 1 groups of activities).
    - E.g. if a respondent performed 10 hours of *personal care* (= TOR level-1) on a weekday, he/she will be marked by our Z-score, because the duration of this activity group deviates from the mean duration of all other respondents who performed 'personal care' on a weekday.
    - E.g. if a respondent played bridge (= TOR level-3, activity code 7321: *playing bridge*) for more than 8 hours on a Saturday, he/she will be marked because the duration is exceptionally long, given the average duration of diary respondents who performed that activity on a weekend day.
    - It's important to note here that the Z-scores were only calculated for participants of an activity, so based on the duration per participant.
  - When respondents were marked (by looking at Z-scores which in absolute sense deviated 3 or more standard deviations from the mean), we checked their diaries in the dataset and controlled the quality of the diary. If there were doubts (e.g. we were unable to reconstruct the respondents day, there were inexplicable periods of unspecified time, suspicious sequences of activities) these respondents were put on

---

<sup>7</sup> We deliberately chose to analyse week- and weekend days separately because some activities are performed more during weekends (e.g. *fishing*), while other activities are typically more performed at weekdays (e.g. *going to school*).

our *list of problematic respondents* and their written diaries were physically checked at Statistics Belgium.

- Based on these Z-score checks, we gave extra attention towards biologically necessary activities (eating, sleeping): where they performed enough? And if not, do we have a logical explanation for that (e.g. respondent which is sick in bed, on vacation, exceptional working tasks, ...)?

#### ***Final diary dataset***

After all cleaning and quality checks were performed, there were 18 respondents which had at least one registered day in their diaries which was of such bad quality that we couldn't let them in our data (8 respondents had one badly registered weekday, 8 a badly registered weekend day and 2 had a badly registered weekday and weekend day).

### **4.3 Cleaning of the Work grid, Labour Force Survey and drop-off questionnaires**

Besides the quality checks on the diary data, the data from the different questionnaires were checked and edited if necessary. Variables included in the Labour Force Survey were edited by Statistics Belgium and variables from the work grid and VUB researchers further edited the drop-off questionnaires. This editing process consisted of checking the routing of the questionnaire, checking for inconsistencies among variables, evaluating outlier values, etc...

## **5 Response**

Detailed response tables will be integrated in the full version of the technical report, which will be online October 23<sup>rd</sup> 2015.



## 6 Weighting and representativity

Because we think it is important to make the BTUS13 data comparable with the Flemish time use data (TOR13<sup>8</sup>) collected in the same period we calculated the BTUS13 weights in the same way as we calculated the TOR13 weights, namely applying a post-stratification weighting procedure for responding individuals (see section 6.2). We also calculated a so-called integrated weight (see section 6.1), as used for the BTUS99 and BTUS05 data, in order to make the three BTUS data sets comparable. In the following paragraphs we report in detail how we calculated these two different types of weights.

### 6.1 Integrated weighting procedure

The weighting, or ‘calibration’, of the BTUS13 respondent sample has been done by means of the SPSS based software *g-Calib* 2.1 (Statistics Belgium, 2006; Vanderhoeft, 2002, 2003) allowing to achieve coherence with (estimated) population figures both at individual and at household level. The (estimated) population figures, serving as benchmarks for calibration of BTUS13, are calculated from LFS13. For that purpose, LFS13 itself has been recalibrated.

In the present section, we briefly describe the calibration model for BTUS13, and, for completeness, comment on the recalibration of LFS13.

#### 6.1.1 Auxiliary variables for calibration of BTUS13

The following characteristics have been used simultaneously as auxiliary variables for calibration of BTUS13:

- individual characteristics, or *individual level auxiliary variables*:
  1. *Gender*;
  2. *Age category*, with 5 classes: 10-17, 18-35, 36-50, 51-65, 66+;
  3. *Education category*, with 3 classes: at most lower secondary education, higher secondary education, higher or university education.
- household characteristics, or *household level auxiliary variables*:
  4. *Household size*, with 3 classes: 1, 2, 3+;
  5. *Socio-professional status of the household’s reference person*, with 2 classes: working, non-working.

Furthermore, the following characteristic of both individuals and households (assuming that all members within a given household have the same place of residence) has served as *calibration stratification variable*:

6. *Region* (as shortcut for the NUTS1 level in Belgium) of residence, with 3 classes: Brussels Capital Region, Flemish Region and Walloon Region;

---

<sup>8</sup> TOR13, or the Flemish time use data, collected online in 2013 with the MOTUS infrastructure.

This means that:

- the individual respondents sample, within each *region*, is calibrated marginally on (estimated) population counts
  - of males and females;
  - for each of the five above-mentioned *age categories*;
  - for each of the three above-mentioned *education categories*;
- the household respondent sample, within each *region*, is calibrated marginally on (estimated) counts of households
  - for each of the three above-mentioned *household size* classes;
  - for each of the two above-mentioned *socio-professional statuses*.

### 6.1.2 Other characteristics of the calibration model for BTUS13

Description of the BTUS13 calibration model can be completed, mentioning that:

- initial weights for BTUS13 respondents are the *sampling weights*. Notice that all members of the same household have the same sampling weight, which is the household's sampling weight. Simultaneous calibration on the aforementioned individual and household level calibration variables, implies that the *calibrated weights* (also called *integrated weights*) are constant within households as well;
- prior to proper calibration, a global correction factor is calculated and applied within each calibration stratum. This means that the sampling weights are roughly adjusted by a constant factor within each region;
- the so-called *truncated linear method* has been applied. Notice that the ordinary *linear method* implies GREG (*generalized regression*) estimation; truncation of the ratios of calibrated to adjusted sampling weights (i.e. the so-called *g-weights*) causes some deviation from GREG estimation, but ensures that *g-weights* and final calibrated weights are positive;
- the boundaries on the *g-weights* that necessarily have to be set for the truncated linear method, are 0 and 8.

### 6.1.3 Benchmarks from LFS13

The Belgian Labour Force Survey is, for the purpose of publication of labour market statistics (e.g. quarterly unemployment rates), calibrated at individual level only, using a simple post-stratification setting; we call this the *traditional model*. (Since some years, LFS is calibrated using the SAS-macro *Calmar*; (see Deville, Särndal, & Sautory, 1993)) More specifically, the following individual level auxiliary variables are used:

- *Province* of residence, where Province is shortcut for NUTS2 level in Belgium;
- *Gender*;
- *Age category*, with 16 classes: 0-4, 5-9, 10-14, ... 75+;

and benchmarks are obtained from the National Population Register. Notice that post-stratification implies full crossing of the above three auxiliary variables, i.e. calibration of the LFS13 individual respondents sample on the joint population distribution of the above-mentioned three variables.

Since BTUS13 has to be calibrated at household level as well (using aggregated individual auxiliary information), we decided to recalibrate LFS13, using simultaneously individual and household level auxiliary information. On the one hand, the traditional LFS calibration model has therefore been extended with terms implying calibration of responding households to the joint distribution of:

- *Province* of residence, as at the individual level;
- *Household size*, with 5 classes: 1, 2, 3, 4, 5+.

On the other hand, the three-way term between the aforementioned three individual variables has been dropped, compared to the traditional LFS calibration model.

Since the final LFS13 calibration model, used to calculate benchmarks for calibrating BTUS13, is no longer a simple post-stratification model, the following choices have to be mentioned to complete the description of the LFS calibration model:

- initial weights are the sampling weights. Notice that for LFS too, all members of the same household have the same sampling weight;
- no global correction factor is calculated;
- the so-called *truncated linear method* has been applied;
- with boundaries on the g-weights set equal to 0.01 and 10.

Benchmarks for recalibrating LFS13 are calculated from The National Population Register d.d. 1/07/2013. Recalibration of LFS13 has been done using *Calmar 2* (Sautory & Le Guennec, 2003) which (just like g-Calib 2.1) allows simultaneously using individual and household level auxiliary information.

Finally, after recalibrating LFS13, the final calibrated weights (which are constant within responding households) have been used to estimate population figures at individual and at household level. Individual level benchmarks are based on LFS respondents aged 10 years or more.

#### **6.1.4 Diary day correction factors**

Although the random assignment of diary days to respondents<sup>9</sup> has been considered carefully (see section 2.4), there remain some imbalances in the frequencies of each of the registered diary days in the BTUS13 respondents sample. A so-called *diary day correction factor* is therefore calculated for each respondent per registered diary day (one week- and one weekend day), and the calibration weight for each respondent is finally multiplied with this correction factor.

In Table 1, under *Ideal Situation*, we present the probabilities that should appear under the situation of equal probabilities of week- and weekend days ( $1/7 = 14,3\%$ ).

---

<sup>9</sup> Actually to responding households, since all household members are assigned the same combination.

Under *Actual Situation*, we present the corresponding observed – CALWEI weighted – frequencies (in the individual BTUS13 respondents sample) of all registered diary days (frequencies calculated using the Belgian Diary Day File (DDFILE.sav), weighted by ‘WGHT2 - Individual response weight’, frequencies of ‘DDV1 - Day of week on which diary completed’ are presented under ‘actual situation’). The diary day correction factor is then simply the ratio between the theoretical probability (*Ideal Situation*) and the corresponding observed frequency (*Actual Situation*) per registered diary day of each respondent.

Table 1. Calculating diary day correction factors

	<b>Ideal Situation</b>	<b>Actual Situation</b>	<b>Diary day correction factors</b>
Monday	0,142857143	0,079289695	1,801711336
Tuesday	0,142857143	0,116530637	1,225919175
Wednesday	0,142857143	0,097974418	1,458106571
Thursday	0,142857143	0,106687620	1,339022680
Friday	0,142857143	0,099517629	1,435495843
Saturday	0,142857143	0,239904623	0,595474740
Sunday	0,142857143	0,260095377	0,549249066

## 6.2 Post-stratification weighting procedure

As said before, to make BTUS13 results comparable with TOR13 results, we also applied a post-stratification weighting procedure to the individual BTUS13 respondents. The following individual characteristics were used:

1. *Region*: Brussels Capital Region, Flanders Region and Walloon Region;
2. *Gender*: male, female;
3. *Age category*, with 5 classes: 10-17, 18-34, 35-54, 55-74 and 75+;
4. *Highest level of education*<sup>10</sup>, with 3 classes: < 15 years / No formal / Max. Lower Sec. Edu (ISCED 1 2), (Post) Secondary Education (ISCED 3 4) and First and second stage Tertiary Education (ISCED 5 6)].

In the original BTUS13 dataset (diaries), there are 5683 respondents, of which 18 were left out because of the low quality of their diaries (see cleaning part). 101 respondents couldn't be coupled with LFS and 5 respondents had problems with the age variable; this resulted in a final diary database of 5559 respondents.

The post-stratification weighting procedure works as follows. Firstly, the BTUS13 respondents are cross classified by the above-mentioned four individual characteristics, and relative cell frequencies are calculated, as shown in Table 2. Secondly, the weighted LFS13 respondents sample is cross-classified in the same way and weighted relative cell frequencies

<sup>10</sup> The *Not stated (aged 15 years or over)* which we use for the weighting of the TOR13 data (gathered online by the MOTUS infrastructure) is not applicable here, since we know the education level of all respondents aged 15 and above.

are calculated. These cell frequencies, shown in Table 3, serve as estimates for the corresponding population cell frequencies. Finally, the relative cell frequencies in Table 3 are divided by the corresponding relative cell frequencies in Table 2, to obtain a post-stratification weight for each BTUS13 respondent in each cell. These weights are shown in Table 5.

Some neighbouring cells in Table 2 are regrouped into larger cells (marked green), because they contained either a small number of BTUS13 respondents or no BTUS13 respondents at all. The corresponding cells in Tables 3 and 4 are regrouped in the same way. This avoids extremely large (possibly infinite) weighting factors.

The post-stratification weighting factors vary between 0,34 and 6,62.

Table 2. *BTUS13 respondents sample (N=5559); absolute and relative frequencies*

Region	Gender	Highest level of education	Age category				
			10-17	18-34	35-54	55-74	75+
Brussels Capital Region	Male	< 15 y; No formal; Max. Lower Sec. Edu (ISCED 1 2)	13 0,234%	4 0,072%	12 0,216%	9 0,162%	
		(Post) Secondary Education (ISCED 3 4)	0 0,000%	11 0,198%	10 0,180%		
		First and second stage Tertiary Education (ISCED 5 6)	0 0,000%	20 0,360%	25 0,450%	17 0,306%	2 0,036%
		< 15 y; No formal; Max. Lower Sec. Edu (ISCED 1 2)	20 0,360%	7 0,126%	7 0,126%	5 0,090%	0 0,000%
	Female	(Post) Secondary Education (ISCED 3 4)	0 0,000%	11 0,198%	15 0,270%	9 0,162%	0 0,000%
		First and second stage Tertiary Education (ISCED 5 6)	0 0,000%	32 0,576%	36 0,648%	16 0,288%	1 0,018%
		< 15 y; No formal; Max. Lower Sec. Edu (ISCED 1 2)	160 2,878%	53 0,953%	97 1,745%	164 2,950%	28 0,504%
		(Post) Secondary Education (ISCED 3 4)	1 0,018%	177 3,184%	228 4,101%	159 2,860%	18 0,324%
Flanders	Male	First and second stage Tertiary Education (ISCED 5 6)	0 0,000%	107 1,925%	238 4,281%	129 2,321%	16 0,288%
		< 15 y; No formal; Max. Lower Sec. Edu (ISCED 1 2)	186 3,346%	39 0,702%	71 1,277%	185 3,328%	40 0,720%
		(Post) Secondary Education (ISCED 3 4)	3 0,054%	178 3,202%	247 4,443%	146 2,626%	10 0,180%
		First and second stage Tertiary Education (ISCED 5 6)	0 0,000%	197 3,544%	303 5,451%	110 1,979%	11 0,198%
	Female	< 15 y; No formal; Max. Lower Sec. Edu (ISCED 1 2)	118 2,123%	45 0,809%	89 1,601%	99 1,781%	10 0,180%
		(Post) Secondary Education (ISCED 3 4)	2 0,036%	97 1,745%	103 1,853%	58 1,043%	7 0,126%
		First and second stage Tertiary Education (ISCED 5 6)	0 0,000%	69 1,241%	118 2,123%	83 1,493%	16 0,288%
		< 15 y; No formal; Max. Lower Sec. Edu (ISCED 1 2)	118 2,123%	54 0,971%	92 1,655%	126 2,267%	5 0,090%
Walloon Region	Male	(Post) Secondary Education (ISCED 3 4)	1 0,018%	111 1,997%	132 2,375%	81 1,457%	5 0,090%
		First and second stage Tertiary Education (ISCED 5 6)	0 0,000%	111 1,997%	156 2,806%	69 1,241%	6 0,108%
		< 15 y; No formal; Max. Lower Sec. Edu (ISCED 1 2)	118 2,123%	45 0,809%	89 1,601%	99 1,781%	10 0,180%
		(Post) Secondary Education (ISCED 3 4)	2 0,036%	97 1,745%	103 1,853%	58 1,043%	7 0,126%
	Female	First and second stage Tertiary Education (ISCED 5 6)	0 0,000%	69 1,241%	118 2,123%	83 1,493%	16 0,288%
		< 15 y; No formal; Max. Lower Sec. Edu (ISCED 1 2)	118 2,123%	54 0,971%	92 1,655%	126 2,267%	5 0,090%
		(Post) Secondary Education (ISCED 3 4)	1 0,018%	111 1,997%	132 2,375%	81 1,457%	5 0,090%
		First and second stage Tertiary Education (ISCED 5 6)	0 0,000%	111 1,997%	156 2,806%	69 1,241%	6 0,108%

Table 3. LFS13 weighted 10+ years aged respondents sample (N=9 846 234);  
Absolute and relative frequencies

Region	Gender	Highest level of education	Age category				
			10-17	18-34	35-54	55-74	75+
Brussels Capital Region	Male	< 15 y; No formal; Max.	51212	46957	56221	77194 0,784%	
		Lower Sec. Edu (ISCED 1 2)	0,520%	0,477%	0,571%		
		(Post) Secondary Education (ISCED 3 4)	375	53670	45767		
			0,004%	0,545%	0,465%		
		First and second stage Tertiary Education (ISCED 5 6)	0	45730	66148	29883	6990
			0,000%	0,464%	0,672%	0,303%	0,071%
	Female	< 15 y; No formal; Max.	49115	42159	49839	48162	34348
		Lower Sec. Edu (ISCED 1 2)	0,499%	0,428%	0,506%	0,489%	0,349%
(Post) Secondary Education (ISCED 3 4)		409	54847	38050	22146	10227	
		0,004%	0,557%	0,386%	0,225%	0,104%	
	First and second stage Tertiary Education (ISCED 5 6)	0	59351	69983	29451	7118	
		0,000%	0,603%	0,711%	0,299%	0,072%	
Flanders	Male	< 15 y; No formal; Max.	271970	124785	212795	310970	147102
		Lower Sec. Edu (ISCED 1 2)	2,762%	1,267%	2,161%	3,158%	1,494%
		(Post) Secondary Education (ISCED 3 4)	3558	351299	384184	222216	56318
			0,036%	3,568%	3,902%	2,257%	0,572%
		First and second stage Tertiary Education (ISCED 5 6)	0	192921	314002	169811	38954
			0,000%	1,959%	3,189%	1,725%	0,396%
	Female	< 15 y; No formal; Max.	262327	89304	171923	363409	260755
		Lower Sec. Edu (ISCED 1 2)	2,664%	0,907%	1,746%	3,691%	2,648%
(Post) Secondary Education (ISCED 3 4)		3178	294329	364148	219731	74770	
		0,032%	2,989%	3,698%	2,232%	0,759%	
	First and second stage Tertiary Education (ISCED 5 6)	0	271497	353498	139041	28316	
		0,000%	2,757%	3,590%	1,412%	0,288%	
Walloon Region	Male	< 15 y; No formal; Max.	169823	109582	148222	168762	64324
		Lower Sec. Edu (ISCED 1 2)	1,725%	1,113%	1,505%	1,714%	0,653%
		(Post) Secondary Education (ISCED 3 4)	1232	184223	194293	104275	24152
			0,013%	1,871%	1,973%	1,059%	0,245%
		First and second stage Tertiary Education (ISCED 5 6)	0	91373	150763	92813	21846
			0,000%	0,928%	1,531%	0,943%	0,222%
	Female	< 15 y; No formal; Max.	167522	77124	130272	343797 3,492%	
		Lower Sec. Edu (ISCED 1 2)	1,701%	0,783%	1,323%		
(Post) Secondary Education (ISCED 3 4)		1443	166134	181406	111092	34130	
		0,015%	1,687%	1,842%	1,128%	0,347%	
	First and second stage Tertiary Education (ISCED 5 6)	0	130056	182196	85288	17628	
		0,000%	1,321%	1,850%	0,866%	0,179%	

Table 4. *BTUS13 post-stratification weights*

Region	Gender	Highest level of education	Age category				
			10-17	18-34	35-54	55-74	75+
Brussels Capital Region	Male	< 15 y; No formal; Max. Lower Sec. Edu (ISCED 1 2)	2,224103	6,627762	2,645111	4,842477	
		(Post) Secondary Education (ISCED 3 4)	1,000000	2,754644	2,583919		
		First and second stage Tertiary Education (ISCED 5 6)	1,000000	1,290915	1,493837	0,992434	1,973212
	Female	< 15 y; No formal; Max. Lower Sec. Edu (ISCED 1 2)	1,386471	3,400312	4,019738	5,438273	1,000000
		(Post) Secondary Education (ISCED 3 4)	1,000000	2,815054	1,432155	1,389246	1,000000
		First and second stage Tertiary Education (ISCED 5 6)	1,000000	1,047140	1,097530	1,039218	4,018690
Flanders	Male	< 15 y; No formal; Max. Lower Sec. Edu (ISCED 1 2)	0,959682	1,329269	1,238558	1,070536	2,966109
		(Post) Secondary Education (ISCED 3 4)	2,008780	1,120547	0,951329	0,789050	1,766450
		First and second stage Tertiary Education (ISCED 5 6)	1,000000	1,017940	0,744873	0,743195	1,374544
	Female	< 15 y; No formal; Max. Lower Sec. Edu (ISCED 1 2)	0,796263	1,292804	1,367106	1,109048	3,680435
		(Post) Secondary Education (ISCED 3 4)	0,598080	0,933554	0,832353	0,849699	4,221375
		First and second stage Tertiary Education (ISCED 5 6)	1,000000	0,778082	0,658675	0,713636	1,453335
Walloon Region	Male	< 15 y; No formal; Max. Lower Sec. Edu (ISCED 1 2)	0,812533	1,374843	0,940263	0,962423	3,631613
		(Post) Secondary Education (ISCED 3 4)	0,347782	1,072256	1,064992	1,015030	1,947967
		First and second stage Tertiary Education (ISCED 5 6)	1,000000	0,747645	0,721339	0,631331	0,770865
	Female	< 15 y; No formal; Max. Lower Sec. Edu (ISCED 1 2)	0,801524	0,806348	0,799447	1,540487	
		(Post) Secondary Education (ISCED 3 4)	0,814691	0,845010	0,775897	0,774327	3,853832
		First and second stage Tertiary Education (ISCED 5 6)	1,000000	0,661506	0,659388	0,697855	1,658740



## 7 Overview of final datasets

For the construction of the data files & variable information: see codebook BTUS13.

- Diary data:
  - TUS13-fictieveweek.sav
  - TUS13-gewichten.sav
  - TUS13-weekdag.sav
  - TUS13-werkweek.sav
  - TUS13-zaterdag.sav
  - TUS13-zondag.sav
  - TUS990513-fictieveweek.sav
- Questionnaires:
  - TUS13-IV-volw.sav
  - TUS13-IV-mind.sav
  - TUS13-dagb-vragen.sav
  - TUS13-WG.sav
  - TUS13-hh.sav
  - LFS13.sav

## 8 List of abbreviations

BTUS99	Belgian Time Use Survey 1999
BTUS05	Belgian Time Use Survey 2005
BTUS13	Belgian Time Use Survey 2013
LFS13	Labour Force Survey 2013
TOR13	Flemish Time Use Survey 2013, collected online with MOTUS software infrastructure (more information can be found online at <a href="http://www.tijdsonderzoek.be">www.tijdsonderzoek.be</a> )
TOR research group	Tempus Omnia Revelat ( <i>time reveals everything</i> ) research group
HETUS	Harmonized European Time Use Surveys

## 9 References

- Belgium, S. (2006). *Methodological Developments towards the Implementation of the Wave Approach for the Labour Force Survey and Improving Calibration Estimation. Part A – Main Text*.
- Deville, J.-C., Särndal, C.-E., & Sautory, O. (1993). Generalized Raking Procedures in Survey Sampling. *JASA*, 88(423), 1013-1020.
- Eurostat. (2008). *Harmonised European time use surveys - 2008 guidelines*. Retrieved from Luxembourg: <http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-08-014-EN.pdf>
- Glorieux, I., Minnen, J., & Mestdag, I. (2007). *Technisch Verslag "Time&Budget"*. Retrieved from Brussel: <http://www.vub.ac.be/TOR/publication/technisch-verslag-time-budget/>
- Sautory, O., & Le Guennec, J. (2003). *La macro Calmar 2: Redressement d'un échantillon par calibration sur marges. Document de travail de la direction des statistiques démographiques and social*. Retrieved from INSEE:
- Vanderhoeft, C. (2002). *g-CALIB Release 1.0 – Generalised Calibration under SPSS®*.
- Vanderhoeft, C. (2003). *g-CALIB Release 2.0 – Supplement to the Manual for Release 1.0*. Statistics Belgium.