Methodology report

Weightings of the Swiss Household Panel

SHP_I wave 9, SHP_II wave 4, SHP_I and SHP_II combined



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Federal Department of Home Affairs FDHA
Federal Statistical Office FSO

Neuchâtel, 2009

Statistik der Schweiz

Die vom Bundesamt für Statistik (BFS) herausgegebene Reihe «Statistik der Schweiz» gliedert sich in folgende Fachbereiche:

0	Statistische Grundlagen und Übersichten
1	Bevölkerung
2	Raum und Umwelt
3	Arbeit und Erwerb
4	Volkswirtschaft
5	Preise
6	Industrie und Dienstleistungen
7	Land- und Forstwirtschaft
8	Energie
9	Bau- und Wohnungswesen
10	Tourismus
11	Mobilität und Verkehr
12	Geld, Banken, Versicherungen
13	Soziale Sicherheit
14	Gesundheit
15	Bildung und Wissenschaft
16	Kultur, Medien, Informationsgesellschaft, Sport
17	Politik
18	Öffentliche Verwaltung und Finanzen
19	Kriminalität und Strafrecht
20	Wirtschaftliche und soziale Situation der Bevölkerung
21	Nachhaltige Entwicklung und Disparitäten auf regionaler und internationaler Ebene

Swiss Statistics

The „Swiss Statistics“ series published by the Federal Statistical Office (FSO) covers the following fields:

0	Statistical Base and Overviews
1	Population
2	Territory and Environment
3	Employment and Income from Employment
4	National Economy
5	Prices
6	Industry and Services
7	Agriculture and Forestry
8	Energy
9	Construction and Housing
10	Tourism
11	Mobility and Transport
12	Money, Banks, Insurance Companies
13	Social Security
14	Health
15	Education and Science
16	Culture, Media, Information Society, Sport
17	Politics
18	Public Administration and Finance
19	Crime and Criminal Justice
20	Economic and Social Situation of the Population
21	Sustainable Development, Regional and International Disparities

Statistik der Schweiz

Methodenbericht

Swiss Statistics

Methodology report

Weightings of the Swiss Household Panel

SHP_I wave 9, SHP_II wave 4, SHP_I and SHP_II combined

Author **Eric Graf**
Swiss Federal Statistical Office

Publisher **Swiss Federal Statistical Office**

Neuchâtel, 2009

Preface

The SHP survey is a survey of households and persons with permanent residence in Switzerland. Its main purpose is to observe social change, particularly the process of change in the living conditions of the Swiss population. The Swiss Household Panel was initially (from 1998 to 2007) a joint project of the Swiss National Science Foundation, the Federal Statistical Office (FSO) and the University of Neuchâtel.

Since early 2008, the SHP project, which is still funded by the Swiss National Science Foundation, has been conducted by FORS, the Swiss Foundation for Research in Social Sciences, located at the University of Lausanne. The FSO was given the mandate to carry out the weightings for the SHP surveys conducted in 2007–2008 (wave 9 of SHP_I and wave 4 of SHP_II).

This report provides detailed information as to how these data have been weighted. The concepts used with regard to the weighting of panel surveys and the combining of panel surveys are rather arduous. The uninformed reader may encounter some difficulty in understanding certain passages. If necessary, a list of documents which deal with the subjects in more detail can be found in the bibliography.

The work presented here is basically an up-date of the same report published last year on the weightings of the previous waves of the SHP.

Our special thanks go to Monique Graf (METH) for her in-depth reading of this report and her acute remarks, as well as to Robin Tillmann (SHP) for his comments and support.

Summary

The document has the following structure: the way data were collected is briefly described for each of the panels, then, after having presented the response rates achieved, a description is given of the different concepts related to the weighting. Next, the steps of the weighting are presented and then described one by one for each of the panels. A distinction is made between steps common to all weightings and steps peculiar to each one. Weightings at household level and individual level are also dealt with separately. The specific features of cross-sectional weightings as compared to longitudinal weightings are also noted. The final sections deal with the combined weightings of SHP_I wave 9 and SHP_II wave 4. At the end of the report, the interested reader will find a collection of appendices that cover in detail some of the steps only briefly mentioned in the main text.

Keywords

Methodology report; response rate; incentives; segmentation method: non-response; calibration; sampling plan; weight sharing; longitudinal weighting; cross-sectional weighting.

Information: Eric Graf, tel. 032 713 64 01
Eric.Graf@bfs.admin.ch

Conduct: Statistical Methods Unit, FSO

Obtainable from: Federal Statistical Office (FSO), CH-2010 Neuchâtel
Phone: +41 (0)32 713 60 60 / fax +41 (0)32 713 60 61 / email: order@bfs.admin.ch

Internet: http://www.statistik.admin.ch

Order number: 338-0058

Price: free of charge

Series: Swiss Statistics

Domaine: Statistical bases and general products

Original text: French

Translation: FSO

Graphics/Layout: FSO

Copyright: FSO, Neuchâtel 2009
Reproduction with mention of source authorised
(except for commercial purposes)

ISBN: 978-3-303-00425-8

List of contents

1 INTRODUCTION: SWISS HOUSEHOLD PANEL (SHP) 5

1.1 INTRODUCTION 5
1.3 STRUCTURE OF THE SHP SURVEY 8

2 SAMPLING PLAN AND SURVEY CONDUCT 10

2.1 SAMPLING PLAN AND GROSS SIZES OF SHP_I AND SHP_II 10
2.2 SURVEY UNIVERSE 11

3 RESPONSE RATE ACHIEVED IN 2007 13

3.1. FOLLOW-UP RULES 14
3.1.1 Rules governing contact with households *14*
3.1.2. Rules governing the follow-up of individuals *15*
3.1.3. Particular features of 2007 (waves 9 and 4) *15*

4 WEIGHTING CONCEPTS 18

4.1 CONCEPTS ON THE PRODUCTION OF LONGITUDINAL WEIGHTS 18
Target and observed longitudinal population *18*
4.1.2 Longitudinal unit *18*
4.1.3 Cohabitants *19*
4.1.4 Longitudinal eligibility *19*
4.2 CONCEPTS ON THE PRODUCTION OF CROSS-SECTIONAL WEIGHTS 19
Target and observed cross-sectional population *19*
4.2.2 Cross-sectional unit *20*
4.2.3 Unit within the scope of the survey from a cross-sectional point of view *20*
4.2.4 Cross-sectional eligibility *20*
4.2.5 Responding units *20*
4.3 ILLUSTRATION OF CONCEPTS BY EXAMPLE 20

5 STEPS IN THE WEIGHTING PROCESS 23

5.1 INTRODUCTION 23
5.2 TYPES OF WEIGHTINGS PRODUCED 24

6 STEPS COMMON TO ALL TYPES OF WEIGHTS 27

6.1 MODELLING OF NON-RESPONSE BY THE SEGMENTATION METHOD 27
6.1.2 Segmentation method *27*
6.1.2 Segmentation and logistic regression *28*
6.1.3 Application for the SHP *29*
6.2 INITIAL WEIGHTS OF SHP_I AND INITIAL WEIGHTS OF SHP_II 30
6.3 CLASSIFICATION OF INDIVIDUALS: STEP 2 31
6.4 ADJUSTMENT FOR NON-RESPONSE – GRID: STEP 3 34

7 INDIVIDUAL CROSS-SECTIONAL WEIGHTING 36

7.1 WEIGHT SHARING: STEP 4 36
7.2 ADJUSTMENT FOR NON-RESPONSE – GRID: INDIVIDUAL QUESTIONNAIRE: STEP 5 37
7.3 CALIBRATION: STEP 6 38

8 HOUSEHOLD CROSS-SECTIONAL WEIGHTING 40

8.1 WEIGHT SHARING: STEP 7 40
8.2 ADJUSTMENT FOR NON-RESPONSE – GRID: INDIVIDUAL QUESTIONNAIRE: STEP 8 41
8.3 CALIBRATION: STEP 9 42

8.4 WINSORISATION OF WEIGHTS AFTER CALIBRATION: STEP 9A 43
9 INDIVIDUAL LONGITUDINAL WEIGHTING 45
9.1 ADJUSTMENT FOR NON-RESPONSE – GRID: INDIVIDUAL QUESTIONNAIRE: STEP 10 45
9.2 CALIBRATION: STEP 11 46
10 INDIVIDUAL CROSS-SECTIONAL WEIGHTING SHP_I AND SHP_II COMBINED 48
10.1 CROSS-SECTIONAL COMBINATION OF PANELS FROM CROSS-SECTIONAL SAMPLES FROM EACH PANEL 48
10.2 COMBINATION OF THE TWO PANELS: STEP 13 50
10.3 CALIBRATION: STEP 14 51
10.4 WINSORISATION OF WEIGHTS AFTER CALIBRATION: STEP 14A 51
11 COMBINED CROSS-SECTIONAL HOUSEHOLD SHP_I AND SHP_II 53
11.1 COMBINATION OF THE TWO PANELS: STEP 15 53
11.2 CALIBRATION: STEP 16 54
11.3 WINSORISATION OF WEIGHTS AFTER CALIBRATION: STEP 16A 54
12 INDIVIDUAL LONGITUDINAL WEIGHTING OF SHP_I AND SHP_II COMBINED 56
12.1 ADJUSTMENT FOR NON-RESPONSE TO THE INDIVIDUAL QUESTIONNAIRE FOR SHP_I, REFERENCE YEAR 2004 56
12.2 COMBINATION OF THE TWO PANELS: STEP 17 58
12.3 CALIBRATION: STEP 18 59
13 CONCLUSION 61
13.1 GUIDELINES FOR THE APPLICATION OF THE WEIGHTINGS PRODUCED 61
13.2 RECOMMENDATIONS ON ESTIMATIONS OF PRECISION 62
14 BIBLIOGRAPHY 63
APPENDIX A – DATA FILES FOR SHP_I AND SHP_II 65
APPENDIX B – DERIVATION OF VARIABLES FOR THE CLASSIFICATION OF INDIVIDUALS 71
APPENDIX C – MODELLING OF NON-RESPONSE TO THE GRID QUESTIONNAIRE 74
APPENDIX D – MODELLING OF NON-RESPONSE TO THE INDIVIDUAL QUESTIONNAIRE, INDIVIDUAL LONGITUDINAL WEIGHTING 79
APPENDIX E – MODELLING OF NON-RESPONSE TO THE INDIVIDUAL QUESTIONNAIRE, INDIVIDUAL CROSS-SECTIONAL WEIGHTING 84
APPENDIX F – MODELLING OF NON-RESPONSE TO THE HOUSEHOLD QUESTIONNAIRE, HOUSEHOLD CROSS-SECTIONAL WEIGHTING 89
APPENDIX G – MODELLING OF NON-RESPONSE TO THE INDIVIDUAL QUESTIONNAIRE, INDIVIDUAL LONGITUDINAL WEIGHTING OF SHP_I AND SHP_II COMBINED. 94
APPENDIX H – CHARACTERISTICS OF DISTRIBUTIONS OF PRODUCED WEIGHTINGS 97
APPENDIX I – RESPONSE RATES 104
APPENDIX J – ALGORITHM FOR BELONGING TO TARGET POPULATION 106

1 Introduction: Swiss Household Panel (SHP)

1.1 Introduction

The SHP survey is a survey of households and persons with permanent residence in Switzerland. Its main purpose is to observe social change, particularly the process of change in the living conditions of the Swiss population. The Swiss Household Panel was initially (from 1998 to 2007) a joint project of the Swiss National Science Foundation, the Federal Statistical Office (FSO) and the University of Neuchâtel.
Since early 2008, the SHP project, which is still funded by the Swiss National Science Foundation, has been conducted by FORS, the Swiss Foundation for Research in Social Sciences, located at the University of Lausanne.

Since its creation in 1999, the survey "Living in Switzerland" has covered a wide range of subjects and methods in the area of social sciences. The survey is conducted annually from September to February by the M.I.S. Trend Institute in Lausanne and Bern using the computer assisted telephone interview technique (CATI). It constitutes a longitudinal database unique in Switzerland.

As far as international cooperation is concerned (see [1], [4], [5], [6]), the SHP has taken part in the Consortium of Household Panels for European Socio-Economic Research (CHER). The SILC (Statistics on Income and Living Conditions from Eurostat) pilot surveys (2004 and 2005) were conducted by the FSO in close cooperation with the SHP. The SHP is currently part of the Cross-National Equivalent File (CNEF).

As indicated by its title, the SHP is a panel, i.e. the same persons and households are interviewed again year after year and answer, with a few exceptions, the same questions. In contrast to a rotating panel it is a simple panel. There are, therefore, no regular activations of new units to be surveyed neither are certain other units withdrawn by choice. At present, the SHP can see the parallel progress over time of two samples which will be referred to as:

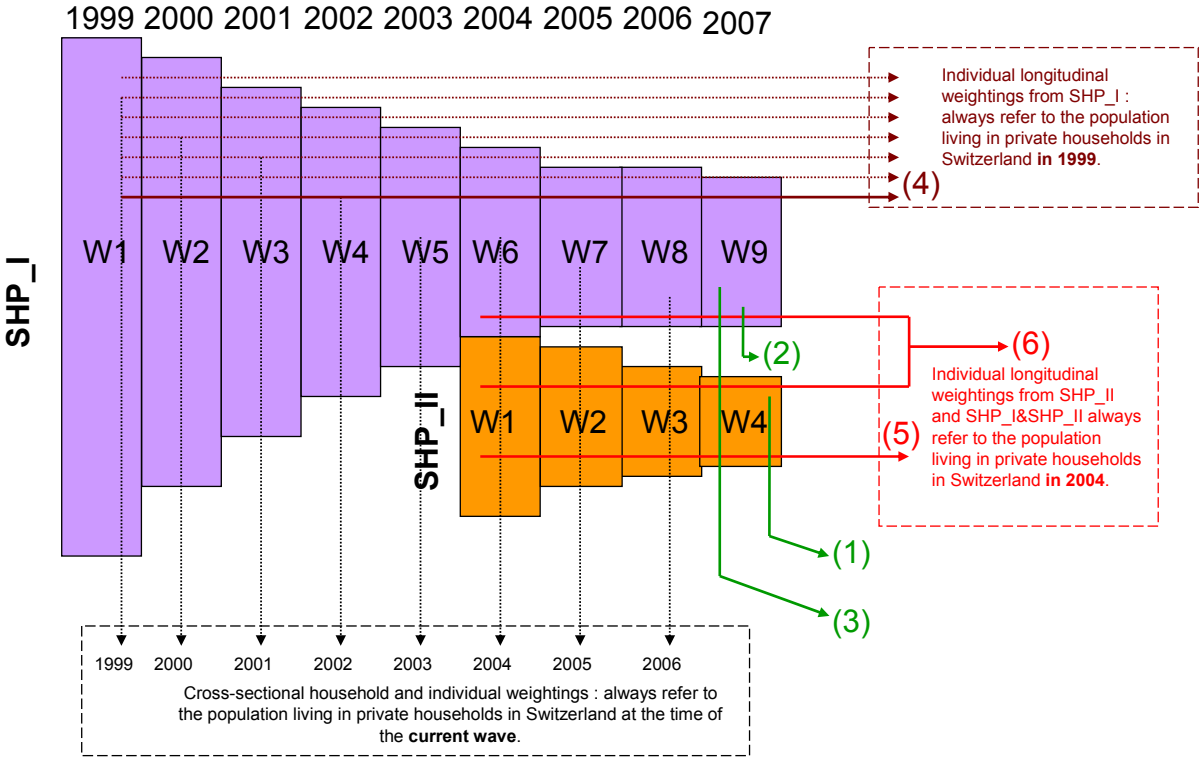
- **SHP_I wave 9:** this is the sample of households and individuals selected in 1999 and interviewed for the first time that year. This document deals with the weighting of the 9th wave of interviews of these units.
- **SHP_II wave 4:** this is the sample of households and individuals selected and interviewed for the first time in 2004. This document deals with the weighting of the 4th wave of interviews of these units.

In 2000, the 1st wave of SHP_I (data collected in 1999) was weighted by FSO methodologists (cf. [2]). After this, and until 2004 (waves 6 of the SHP_I and 1 of SHP_II), calculation of the weightings was carried out entirely by Statistics Canada. In 2005 (waves 7 and 2) the calculation was made in cooperation with the author of this document and Statistics Canada. (cf. [3]). For the years 2006 (cf. [24]) and 2007, the author carried out this task. The object of this report is an up-to-date version[1] of the same report that appeared last year regarding 2006.

[1] To facilitate comparison between weightings carried out for 2006 and 2007, we kept the same numbering of Tables in this report as that used in previous years. This is why the numbering of Tables in this document jump from 8 to 16.

The present document describes in detail the production process of weights for the SHP survey of 2007[2], i.e. those of the SHP_I wave 9, of the SHP_II wave 4 as well as the combined weighting of the samples SHP_I and SHP_II for 2007. Diagram 1 provides an outline of the nine different sets of weights produced.

Diagram 1 Outline of nine different sets of weights produced in 2007: three individual transversal weightings, three household cross-sectional weightings and three individual longitudinal weightings



Legend

(1) Individual and household cross-sectional weightings of SHP_II
(2) Individual and household cross-sectional weightings of SHP_I
(3) Combined individual and household cross-sectional weightings of SHP_I and SHP-II together
(4) Individual longitudinal weighting of SHP_I
(5) Individual longitudinal weighting of SHP_II
(6) Combined individual longitudinal weighting of SHP_I and SHP_II together

The weightings described in this document are, apart from a few minor exceptions, very similar to those from the year 2006 and those which were used by Statistics Canada for the years before and are perfectly compatible with the latter.

[2] To simplify matters, we will say "carried out in 2007" to denote the survey which took place from September 2007 to February 2008.

1.2 The problem of attrition

Diagram 2 and Table 3 show the loss of individuals and households that the SHP has experienced since its inception.
In order to compensate for this loss from the original sample, in 2004 a second panel (the SHP_II) was started in parallel to the first one. From 2004 onwards, various measures have been taken to fight this attrition (incentives, supplementary training of interviewers, reactivation of households that hadn't taken part in the survey for more than two years, etc.). For more details regarding the progress of the survey in the field, please refer to notes [5] and [6].
For the rules concerning the follow-up of individuals, please see paragraph 3.1; for the response rate achieved at the end of fieldwork, please see chapter 3 and Appendix I.

Diagram 2 Attrition of SHP_1 and SHP_II surveys. Number of individual and household questionnaires completed each year.

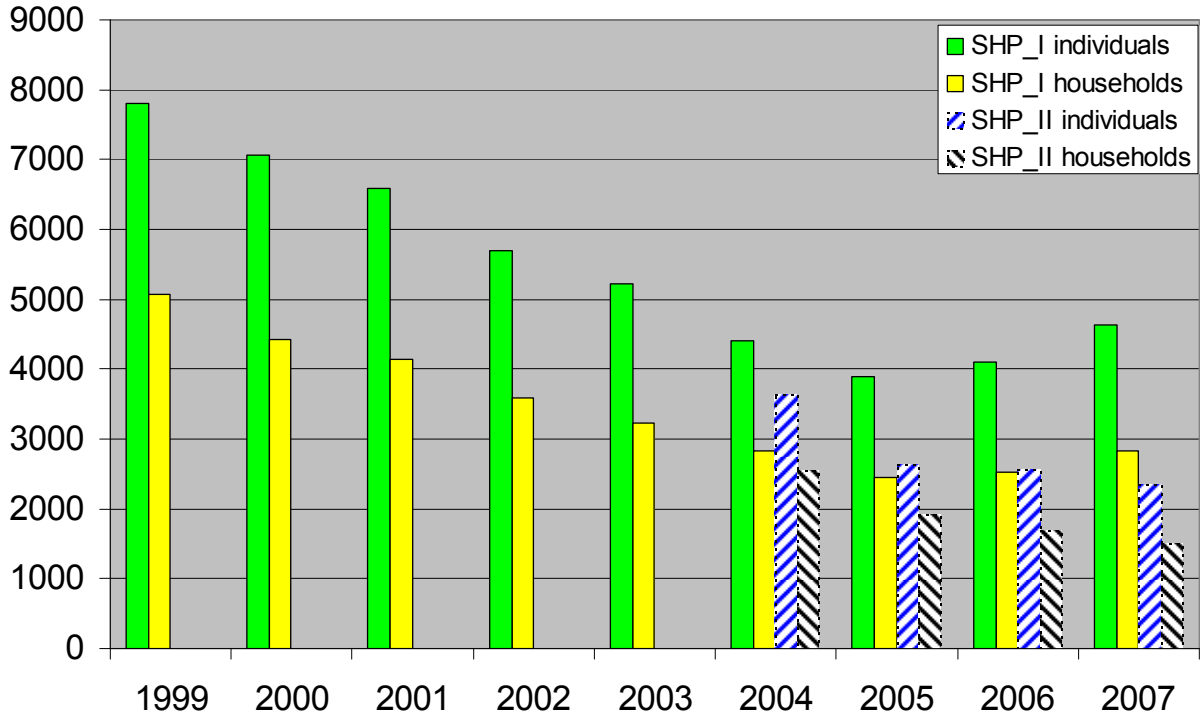


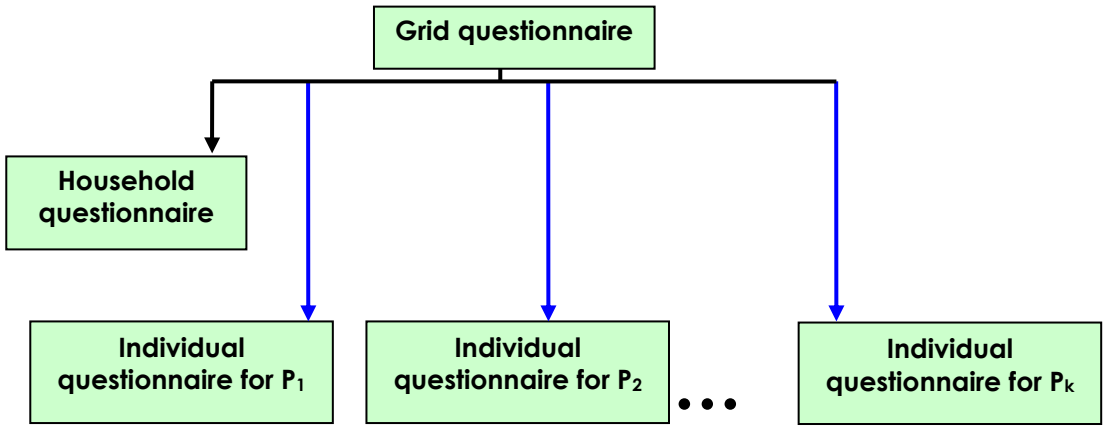
Table 3 Number of individual/household questionnaires completed in the SHP survey. Relative attrition is the proportion of the number of cases interviewed during the year n compared with the total number of cases of the year (n-1).
Please note that these are the numbers used for the calculation of weightings. These totals may differ slightly from those published by the SHP. In fact, the weightings are based on the one hand on the statistics and rules of monitoring and on the other hand their calculation often begins before the data from the survey can be considered definitive for all variables.

	Individual level			*Household level*		
Year	SHP_I	Relative attrition	Absolute attrition	SHP_I	Relative attrition	Absolute attrition
1999	7799	1.00	1.00	5074	1.00	1.00
2000	7071	0.91	0.91	4425	0.87	0.87
2001	6593	0.93	0.85	4139	0.94	0.82
2002	5691	0.86	0.73	3581	0.87	0.71
2003	5220	0.92	0.67	3227	0.90	0.64
2004	4406	0.84	0.56	2831	0.88	0.56
2005	3885	0.88	0.50	2457	0.87	0.48
2006	4093	1.05	0.52	2534	1.03	0.50
2007	4634	1.13	0.59	2822	1.11	0.56
	SHP_II			SHP_II		
2004	3648	1.00	1.00	2538	1.00	1.00
2005	2648	0.73	0.73	1909	0.75	0.75
2006	2570	0.97	0.70	1682	0.88	0.66
2007	2351	0.91	0.64	1495	0.89	0.59

1.3 Structure of the SHP survey

The SHP is a survey comprising 3 questionnaires: the grid, the household questionnaire and the individual questionnaire.

Diagram 4 Structure of the SHP survey



The questionnaire grid is a type of household screening (or inventory), lasting less than 10 minutes. It enables, amongst other things, the collection of basic socio-demographic information on all members of the household, as well as establishing whether they are eligible for interviews. Neither the household questionnaire nor the

individual questionnaire can be completed before the questionnaire grid has been carried out. Usually the *reference person*[3] answers the questionnaire grid but if this person is not available at the initial contact, another adult (aged 14 and above) can answer.
Once the grid has been completed, the interviewer can move on to the household and individual questionnaires. The household questionnaire should preferably be completed by the reference person or by an adult who "knows the household well". All individuals aged 14 and above and who live in the household are eligible to answer the individual questionnaire.

A description of the data sheets used for the weighting of wave 9 of the SHP_I and wave 4 of the SHP_II as well as their combination can be found in Appendix A.

[3] The reference person must be eligible and know his or her household "well ". This will generally be the person who responds to the household questionnaire year after year.

2 Sampling plan and survey conduct

2.1 Sampling plan and gross sizes of SHP_I and SHP_II

The samples of SHP_I and SHP_II are stratified by major geographic region (NUTS II), in proportion to the number of households per stratum. The addresses of the gross sample are distributed in the following proportions (1990 and 2000 censuses):

Table 5 Stratification of gross sample. The proportions are given by $P_h^{99} = \frac{N_h^{99}}{N^{99}}$ and $P_h^{04} = \frac{N_h^{04}}{N^{04}}$ (see also Table 6).

Strata *h*	Cantons	Proportion of addresses P_h^{99} (as %) SHP_II	Proportion of addresses P_h^{04} (as %) SHP_II
1: Lake Geneva region	VD, VS, GE	18.45	18.22
2: Espace Mitteland	BE, FR, SO, NE, JU	23.25	22.92
3: North-west Switzerland	BS, BL, AG	13.44	13.86
4: Zurich	ZH	17.51	18.22
5: Eastern Switzerland	GL, SH, AR, AI, SG, GR, TG	15.68	13.70
6: Central Switzerland	LU, UR, SZ, OW, NW, ZG	7.20	8.75
7: Ticino	TI	4.47	4.33
Total		100%	100%

The selection is thus **strictly** proportional to the number of households per major region and does **not take into account** the average number of persons in the households per region. The size of Ticino is not increased.
The survey was based on the SRH[5] which is updated quarterly in March, June, September and December. The SHP_I was taken from the SRH of 02/99, the SHP-II from the SRH 02/04. The size of strata at the moment of the selection for SHP_I and SHP_II were as follows:

Table 6 Sizes (SRH) N_h of strata h (h = 1,..,7) at the moment of both selections

Strata *h*	SHP_I, 1999 N_h^{99}	SHP_II, 2004 N_h^{04}
1: Lake Geneva region	714'725	648'590
2: Espace Mitteland	837'452	784'266
3: North-west Switzerland	484'667	455'833
4: Zurich	646'469	587'850
5: Eastern Switzerland	531'731	493'606
6: Central Switzerland	313'548	306'605
7: Ticino	180'623	160'123
TOTAL (N=)	**3'709'215**	**3'436'873**

[5] FSO register for household surveys.

Table 7 indicates the distribution of the gross sample, i.e. the number of addresses activated for the survey by stratum, marked $S_{h,}$ in wave 1.

Table 7 Sizes S_h from the gross sample by strata. The S_h represent the number of addresses activated for the survey. These addresses do not necessarily coincide with the private households, but do so in the majority of cases. Please refer to Appendix H for a detailed description of the final status attributed to each address.

Strata *h*	SHP_I, v1 1999 S_h^I	SHP_II, v1 2004 S_h^{II}
1: Lake Geneva region	2615	912
2: Espace Mitteland	3296	1146
3: North-west Switzerland	1904	692
4: Zurich	2482	912
5: Eastern Switzerland	2223	686
6: Central Switzerland	1021	436
7: Ticino	633	216
TOTAL	**14174**	**5000**

2.2 Survey universe

The SHP covers all individuals living in private households in Switzerland who had a telephone connection (landline or mobile) registered in the telephone directory. Individuals living in old peoples' homes, institutions, collective households or prison, are not covered by the survey. Persons who could not be contacted by telephone or whose telephone connection is not listed, are not covered by the survey (under coverage).
An estimated 98.5% of private households had a telephone connection at the time of the selection of the sample for the SHP_II in 2004. Furthermore, the SRH (survey frame for household surveys) for its part, covered about 93% of these households. In 1999, at the time of the selection of the sample for the SHP_I, the SRH's coverage was about 95%.

The survey frame SRH is subject to the following errors:

- under-coverage: some households were not listed in the directory at the time of selection. This includes households whose numbers are not listed or households that could not be contacted by telephone. This problem may produce a bias, namely differences between the estimates based on the observed population (SHP survey) and those that would have been produced based on the target population (all individuals living in private households in Switzerland at a given time).
- *duplicates:* despite meticulous checking of the SRH to ensure that only one number is kept per household, some households appear more than once in the survey frame. This problem results in a faulty determination of some of the initial selection probabilities. In spite of this, a correction factor is not calculated for households with several telephone lines. The information is available but the effect is negligible.
- *over-coverage*: selection of units outside the target population (businesses, homes, prisons, collective households, second homes ….). It should be noted

that for a panel, this problem is only encountered at wave 1 and that these addresses are usually considered as outside the sample (non-sample cases).

3 Response rate achieved in 2007

Table 8 summarises the response rates obtained in relation to the gross sample, i.e. in relation to the number of addresses activated for the enquiries regardless of their final status (validity). It should be pointed out that a wave 1 behaves very differently from subsequent waves. For more details regarding the response rate, please refer to Appendix I.

Table 8 Response rates at different levels of data collection for the SHP_I in 1999 (wave 1) and 2007 (wave 9), as well as for the SHP_II in 2004 (wave 1) and 2007 (wave 4) (source: Final statistics M.I.S. Trend SA).

		SHP_I wave 1, 1999				SHP_I wave 9, 2007			
Household level	Gross sample (number of addresses activated)	14174	100%			3976	100%		
	Number of grids completed	7063	50%	100%		2898	73%	100%	
	Grids + quest. Households completed	6001	42%	85%	100%	2822	71%	97%	100%
	Number of complete households	5074	36%	72%	85%	2770	70%	96%	98%
Individual level	Number of individuals eligible for indiv.quest.	10293	100%			5974	100%		
	Number of individual quest. completed	7799	76%			4634	78%		

		SHP_II wave 1, 2004				SHP_II wave 4, 2007			
Household level	Gross sample (number of addresses activated)	5000	100%			2037	100%		
	Number of grids completed	2704	54%	100%		1549	76%	100%	
	Grids + quest. Households completed	2535	51%	94%	100%	1495	73%	97%	100%
	Number of complete households	2335	47%	86%	92%	1457	71%	94%	97%
Individual level	Number of individuals eligible for indiv.quest.	5384	100%			3108	100%		
	Number of individual quest. completed	3652	68%			2351	76%		

As in 2006, we used incentives and took the following steps to motivate respondents:

- a **thank you** for respondents: all those responding to the individual questionnaire could choose between a CFF voucher for CHF 15 (18%), a voucher for the cinema (33%) or a donation of CHF 15 to a charity (49%).

[6] As mentioned in the previous chapter, the totals presented here may differ slightly from the numbers published by the SHP.

[7] A household is said to be "complete" if the questionnaire grid for the household and at least one individual questionnaire have been carried out.

- before calling those who had refused to take part, a letter was sent to the household in the case of household refusal, to the individual in the case of individual refusal.
- possibility for interviewers to interview all household members at individual level, regardless of their decision of previous years to participate or not in the survey.
- two premiums for the interviewers dependent on objectives being achieved or not, one based on the general participation rate and the other on the rate of convincing those who had refused to take part.
- information on progress of work regularly transmitted to those conducting survey.

In addition, we reactivated 1,354 former addresses of respondents for the last time during waves 2, 3 and 4, as we did in 2006 by reactivating former addresses of respondents for the last time during wave 5 as this procedure had proved to be successful regarding the final number of persons interviewed as well as for curbing attrition.

Furthermore, all households whose address had not been activated at the start of data collection but at a later stage and directly during the call to those who had refused, received a CFF voucher for CHF 50 with the preliminary letter. Whereas the success rate last year for the reactivated W5 households was 35% of complete households, it rose to 52% for the households lost in W2, W3 or W4. The difference was probably due not only to the fact that we had left these households in peace for longer but also thanks to the CFF voucher for CHF 50 which was sent with the preliminary letter.

3.1. Follow-up rules

Although the panel monitors individuals and not households, the first contact is always with (an individual in) a household! From wave 2 on, when making contact again based on information collected the previous year, we do not yet know whether the household we are contacting still includes longitudinal or not. This is why there are follow-up rules for households to determine which households are to be recontacted in relation to the previous year and follow up rules for individuals to determine who should actually be interviewed.

3.1.1 Rules governing contact with households

The general rule is to interview all households that completed at least one grid during the previous wave. We proceed with interviews as long as a longitudinal individual agrees to fill in the household questionnaire and his or her individual questionnaire (it is always possible to catch up with the other individuals in a future wave). However:

1. We permanently reject for following waves households that were not contacted at all during the 1st wave or those that did not supply any information at the time of the 1st wave (not even a grid) or those who only completed a non-response questionnaire for wave 1.
2. **For SHP_I** we permanently rejected all households that only replied to the grid for wave 1. For **SHP_II**, we changed this rule and only permanently rejected

households that had only completed a grid for wave 1, if at wave 2 we did not collect more than its grid.

3. We permanently reject households which do not reply for two successive waves, those who give a "final refusal" (households where no one is willing to respond to a household interview after all attempts to change their mind), those who move away from Switzerland and those who are fully and permanently institutionalised.

3.1.2. Rules governing the follow-up of individuals

4. Longitudinal respondents (OSM) and their children are continuously monitored, whereas co-habitants[10] are only (re-)interviewed as long as they live with an OSM.
5. The minimum age of eligibility is 14 years.
6. As a general rule, longitudinal respondents (OSM[11]) are always monitored until they die or are permanently institutionalised (or leave the target population for another reason).
7. Individuals giving a "final refusal" for two consecutive waves, as well as those who send us a written refusal are permanently rejected.

3.1.3. Particular features of 2007 (waves 9 and 4)

For wave 9, it was decided to release the addresses of all complete households interviewed in wave 2 (2001), 3 (2002) and 4 (2003), including those who had since been rejected due to a double refusal (i.e. refusal for two consecutive waves). However we did not try to get in touch again with households that had been rejected due to a written refusal. In this way more than 1500 households were "reintegrated" into the gross sample. 40% of them agreed to participate in the survey again. This operation was, therefore, a success.

Furthermore, as in 2006 the rule number 7 mentioned above was abolished. We thus gave the interviewers the possibility to interview all the members of a household regardless of their previous refusals.

In addition to this, as far as the monitoring of invalidated addresses and number is concerned, M.I.S. Trend SA used the services of Dcl.data care, specialised in tracing household moves and other changes of address.

[10] For more details regarding the concepts of longitudinals and cohabitants, please see chapter 4.
[11] OSM = Original Sample Member.

For the management of addresses, in 2007, the follow-up rules were as follows:

Survey	Description	Decision
SHP_I	1. Complete households in wave 8 or only household questionnaire or only individual questionnaire	Activate
	2. Grids only in wave 8 not coming from *reactivated ex-respondents* from wave 5 and with at least the household questionnaire or individual questionnaire in wave 8	Activate
	3. Grids only in wave 8 and coming from reactivated ex-respondents from wave 5	Permanently reject
	4. Total non-respondents in wave 8 and coming from reactivated ex-respondents from wave 5	Permanently reject
	5. Total non-respondents in wave 8 but at least the household questionnaire or individual questionnaire in wave 7	Activate in call to refusers with specialised interviewers
	6. Total non-respondents in wave 8 and only grid questionnaire in wave 7	Activate in call to refusers with specialised interviewers
	7. Total non-respondents in wave 8 and total non-respondents in wave 7	Permanently reject
	8. Complete households for the last time in wave 2, 3 or 4 AND: not already included in categories 3 or 4 AND: not rejected due to a written refusal	Activate in call to refusers with specialised interviewers
SHP_II	9. Complete households in wave 3 or only household questionnaire or only individual questionnaire	Activate
	10. Grid questionnaire only in wave 3 and with at least the household questionnaire or individual questionnaire in wave 2	Activate
	11. Total non-respondents in wave 3 but at least the household questionnaire or individual questionnaire in wave 2	Activate in call to refusers with specialised interviewers
	12. Total non-respondents in wave 3 and only grid questionnaire in wave 2	Activate in call to refusers with specialised interviewers
	13. Total non-respondents in wave 3 and total non-respondents in wave 2	Permanently reject

However, all written refusals received each year, deaths as well as addresses that could not be found following the return of the first letter have never been reactivated.

The follow-up rule n°8, as for the previous year, had a positive effect on the absolute number of interviews conducted at household and individual levels (see Diagram 2 and Table 3) This explains the higher number of persons and households interviewed in 2007 than in 2006 (in the same manner that those of 2006 were higher than those of 2005 for the same reasons)!

3.2 Conduct of surveys SHP_I wave 9 and SHP_II wave 4

The period of fieldwork lasted six months, from 3 September 2007 to 1 March 2008. As during the previous wave, the mailing list to be activated in the survey was divided into three folders containing approximately the same number of addresses in order to contact the maximum number of households during the week following the initial letter. All households had received a letter from the SHP beforehand, notifying them of the coming telephone call from an interviewer from M.I.S. Trend (see Table 16).
Last year, a brochure explaining to households the importance of their repeated participation each year was sent with the initial letter. This was not the case this year, except for letters sent to those who had refused in previous years as well as households lost a few years ago and reactivated for this wave.
Furthermore, all households whose address had not been activated at the start of data collection but at a later stage and directly during a call to those who had refused (please see paragraph), received a CFF voucher for CHF 50 with the preliminary letter.

For a more precise and detailed description on the conduct of the surveys and of the pilot work with the survey institute, please refer to references [5], [6] and [25].

Table 16 Dates of mailing and address activation: the mailing of the initial letter took place each time about a week before the release of a new block of addresses. These letters were sent before each activation of addresses (also for refusal calls).

Mailing of initial letter	Activation of sample addresses	Mailing of letter for refusal calls	Activation of sample addresses of refusal calls
31 August 2007	5 September 2007	26 September 2007	4 October 2007
7 September 2007	13 September 2007	10 October 2007	16 October 2007
14 September 2007	19 September 2007	16 October 2007	23 October 2007
		23 October 2007	1 November 2007



4 Weighting concepts

Here is a summary of the terminology and philosophy concerning how to construct current weightings of the SHP. Readers unfamiliar with these concepts may refer to [9], [15], [16], [17] and [18] for a more in-depth study of the subject.

4.1 Concepts on the production of longitudinal weights

Cross-sectional and longitudinal weightings were carried out basically in the same way as by Statistics Canada in previous years. The concepts on weighting are presented here in the context of the SHP. An example demonstrating the theoretical aspects touched upon below and which are difficult for an unfamiliar reader to fully understand, is found at the end of this chapter.

Target and observed longitudinal population

The *target longitudinal population* comprises all individuals living in private households in Switzerland in September 1999 for the SHP_I, in September 2004 for the SHP_II, on which date data collection began. The survey frame used is SRH (see chapter 2).
The observed longitudinal population comprises all individuals living in private households in Switzerland who had a valid telephone number, listed in the directory in the SRH 02/99 for the SHP_I, respectively in the SRH 02/04 for the SHP_II, on which date the sample was selected.

4.1.2 Longitudinal unit

When the sample is selected, all individuals belonging to households selected and aged 14 years and over at the time of selection, become members of the longitudinal sample, regardless of whether they have been respondents or non-respondents to the individual questionnaire, and are called "longitudinals". These individuals are considered members of the longitudinal sample for the life of the panel.
In practice, the follow-up rule of the longitudinals was that their household had at least answered the grid questionnaire in wave 1. In order to simplify the weighting procedure, we had called longitudinal all individuals in the longitudinal sample who lived in households that had at least responded to the grid in wave 1.

If a longitudinal individual moves house, dies or leaves the target population, he or she will always remain part of the target longitudinal population and will have a longitudinal weight. In fact, the control totals used for calibration take into account all units of the population of wave 1 regardless of what may happen to them in the future. We must therefore be consistent and assign a weight to all units even if they have no analytical usefulness. Furthermore, it is desirable that all estimates produced should be coherent with those from other sources. In this way, it is possible to evaluate the representativeness of the sample over time.

In principle, no other person can become member of the longitudinal sample of the panel[12]. The longitudinal sample is thus created at the moment of the panel's introduction, at the time of wave 1, and remains as it is for the duration of the panel. The population targeted by the longitudinal sample will stay the same for the life of the panel. From a longitudinal point of view, the unit of interest is the individual; it would be difficult to use the household as longitudinal analytical tool due to its dynamic nature over time. Moreover, the calculation of a longitudinal weight at household level would require additional theoretical development. In fact, a large number of definitions of longitudinal household have to be considered before making a choice.

4.1.3 Cohabitants

Cohabitants are individuals who have joined the longitudinals' households between the first wave and the current wave. The cohabitants are classified according to two categories:

- *initially present* cohabitants, who belonged to the cross-sectional target population at the moment the panel was selected;
- *initially absent* cohabitants, who joined the cross-sectional target population between the first wave and the current wave. The latter are immigrants.

4.1.4 Longitudinal eligibility

An individual from the original sample is eligible for the *longitudinal weighting* if he or she belonged to the target population at the time of the selection of the original sample (in 1999 for the SHP_I, in 2004 for the SHP_II) and has filled in the individual questionnaire for the current year.
It should be noted that for data users, this rule presents the disadvantage of a non-respondent longitudinal individual in wave 1 being given a weight of zero in wave 1, but should he or she respond to wave 2, he or she will be given a longitudinal weight in wave 2.

4.2 Concepts on the production of cross-sectional weights

Target and observed cross-sectional population

The *cross-sectional target population* for wave 9 of the SHP_I and wave 4 of the SHP_II comprises private households in Switzerland in existence in September 2007.
The *observed cross-sectional population* for wave 9 of the SHP_I comprises private households existing in Switzerland in September 2007 and having at least one member who belongs to the observed longitudinal population. Specifically, this population therefore excludes households composed exclusively of immigrants who arrived in Switzerland after September 1999 for the SHP_I, after September 2004 for the SHP_II, as well as certain individuals who were institutionalised or non-selectable for another reason in September 2007.

[12] However, as was done last year, we broke this rule and considered children of longitudinals who had become adult as longitudinals and not as initially absent cohabitants. We did this so as not to lose households where children who were under 14 at wave 1 and are now over 14 are the only longitudinal representatives or respondents of a household.

4.2.2 Cross-sectional unit

From a cross-sectional point of view, we are more interested in individuals aged 14 and over than in households. The SHP interviews all persons living with at least one longitudinal individual. The cross-sectional sample for a given year is thus made up of all longitudinal individuals in the scope? of the survey during the relevant reference year, and all persons who live with them at the time of the individual and who are in the scope of the survey.

4.2.3 Unit within the scope of the survey from a cross-sectional point of view

A household from the sample is considered to be within the scope of the survey for the current wave if it is situated in one of the seven major regions sampled and belongs to the target population.
Every longitudinal individual who is within the scope of the survey from a cross-sectional point of view will also belong to the cross-sectional sample.

4.2.4 Cross-sectional eligibility

An individual in the sample is eligible for *cross-sectional weighting* if he or she belongs to the target population for the reference year being dealt with and has responded to the individual questionnaire for the current year. Equally, a household in the sample is eligible for *cross-sectional weighting* if it belongs to the target population for the reference year being dealt with and has responded to the household questionnaire as well as (compulsorily) to the grid for the current year.

4.2.5 Responding units

An individual is considered a *respondent* if he or she has completed the individual questionnaire.
A respondent household is considered as such if it has completed the household grid and questionnaire.

It may happen that a household only completes the grid and its household questionnaire, and also, for a limited number of cases, that a household completes the grid as well as one or several individual questionnaires without having completed the household questionnaire.

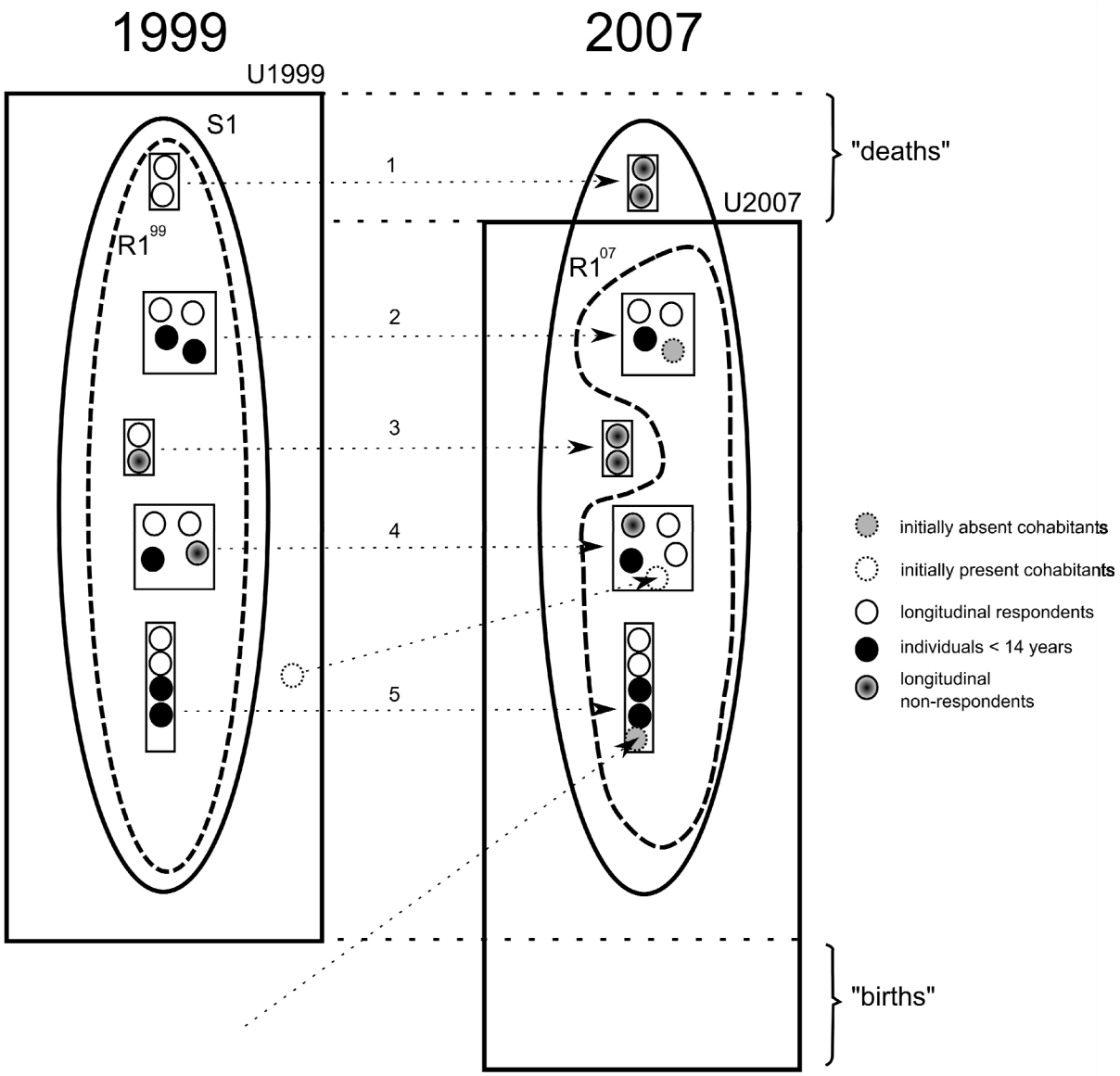
4.3 Illustration of concepts by example

Diagram 17 shows the definitions and concepts listed in this chapter. The circles represent individuals and the rectangles the households in which they live. We have taken 5 frequent situations (numbered arrows in the Diagram) of transition between two consecutive years which we have commented on below:

1. All longitudinals from a household leave the cross-sectional target population of the current year. These are "dead" in the widest sense of the word as these individuals are not necessarily deceased. These individuals maintain a longitudinal weight in 2007, but they will have no cross-sectional weight.

2. One of the children from the household becomes eligible for individual interview. Strictly speaking, this child should be treated like an initially absent individual as he or she did not belong to the target population in 1999 for the SHP_I, in 2004 for the SHP_II. In this case, from 2007 onwards, if he or she responds, this individual is eligible for cross-sectional weighting but not for longitudinal weighting. However, as we have noted previously, since wave 5, we have decided to consider this child of longitudinals as longitudinal too, as soon as he or she is old enough to be interviewed. He or she is therefore eligible for longitudinal and cross-sectional weighting. From a theoretical point of view, this strategy is not completely correct as it carries the risk of provoking a bias in the analyses, but it does allow monitoring to be continued of households where the child (in 1999 for the SHP_I, 2004 for the SHP_II) would remain the only individual from the household responding at the time of the current wave.
3. The individuals from a household responding in wave 1 do not respond for subsequent waves but they remain in the cross-sectional target population. In this case they are no longer eligible for any of the weightings.
4. A non-respondent longitudinal individual in wave 1 starts responding to a subsequent wave. This individual will have a cross-sectional and longitudinal weight.
   Another longitudinal individual, respondent in wave 1, does not respond to a subsequent wave but remains in the cross-sectional target population. This individual will have neither cross-sectional nor longitudinal weight.
   An individual who did not belong to the sample S1, but who belonged to the cross-sectional target population in wave 1, joins the household. He or she is therefore an initially present cohabitant who could have been selected to take part in the survey in wave 1. This individual will have a cross-sectional weight (if he or she responds) but no longitudinal weight).
5. An individual who did not belong to the cross-sectional target population in wave 1 joins the household. This individual will have a cross-sectional weight (only if he responds) but no longitudinal weight.

Diagram 17 Illustration of the basic concepts for longitudinal and cross-sectional weightings. This fictive example is set in the SHP_I. The situation for the SHP_II is exactly the same - simply replace 1999 by 2004.



U1999 Cross-sectional target population of SHP_I en 1999 = longitudinal target population of SHP_I in 2007

U2007 Cross-sectional transversal target population of SHP_I in 2007

S1 Gross sample of SHP_I

R1^{99} Sample of individuals in households that responded to at least the grid questionnaire in 1999, SHP_I (= net sample)

R1^{07} Sample of individuals in households that responded to at least the grid questionnaire in 2007, SHP_I (= net sample)

"**Deaths**" Individuals who have left the cross-sectional target population of the current year (deceased persons, gone abroad, institutionalised…)

"Births" Individuals who did not belong to the target population the previous year (newborns, immigrants who have settled in Switzerland…)

5 Steps in the weighting process

5.1 Introduction

This paragraph describes the general principle: the details of each step are explained in paragraph 6.
For all weightings, individuals are first classified according to their status and their answers to the different SHP questionnaires. Next an adjustment is made for non-response to the grid.

For the *individual cross-sectional weighting*, the weights are shared, and then the weight of the longitudinals and cohabitants[13] are adjusted for non-response to the individual questionnaire. Lastly, a margin calibration is carried out to adjust the weights so that the total number of estimated persons is equal to the Swiss population in 2007.

For the *household cross-sectional weighting*, the weights are shared then adjusted for the non-response to the household questionnaire. Lastly an integrated calibration adjusts the weights so that the total number of persons equals the population of Switzerland in 2007, while at the same time obliging the weights of persons in the same household to remain equal amongst themselves.

The method of analysis by segmentation is used in all non-response adjustments. It should be stated that for the individual weighting, no adjustment was carried out for non-response to the household questionnaire. Non-response adjustment for the individual questionnaire is therefore made on the basis of all individuals from households which responded to the grid. Lastly, for the individual cross-sectional weighting, the longitudinals and cohabitants had to be modelled together. The reasons for these choices are explained in detail in the relevant sections of the document. The distribution of household cross-sectional weighting in the SHP_I wave 9 and the household and individual cross-sectional weighting in the SHP_I and SHP_II combined, have been Winsorised to the 1st and 99th percentile due to the presence of some extreme and non-realistic values which were potentially too disruptive to the analyses.

The marginal calibrations were carried out using an SAS macro developed by Jean-François Naud and Caroline Cauchon of Statistics Canada for the Swiss Household Panel. More precisely, the method used is a calibration by generalised regression which is based on a distance function G_k defined by (see [10], [19]):

$$G_k(d_k, w_k) = \frac{(w_k - d_k)^2}{2\, d_k q_k}$$

where w_k is the weight after calibration for the unit k, d_k the weight before calibration, and q_k is an adjustment factor independent of the d_k. $q_k = 1$ is placed most often for all observations. Please note that $G_k(d_k, w_k)$ may produce negative weights. The same results can be achieved with the CALMAR macro developed by the INSEE (French National Institute of Statistics and Economic Studies) by using the "linear" method. Most FSO survey use (CALMAR with) the "raking ratio" method,

[13] See chapter 4 : « Concepts related to weighting » for the definition of cohabitants.

precisely in order to eliminate any chance of obtaining negative results. This last method, however, does present certain disadvantages. For example, it has been found empirically that for the SILC and SHP survey, it produced more weight in the interval (0.1) as well as very high weights.

Both methods present advantages and disadvantages. We have chosen to keep to the linear method in order to guarantee comparability and compatibility with the weightings produced for previous waves. Furthermore, at household level, an integrated calibration can easily be carried out by the macro developed by Statistics Canada.

For the cross-sectional weighting, which combines the sample from the SHP_I wave 9 with that of the SHP_II wave 4, there are two ways of proceeding:

1. The panels can be combined in one go, after the grid adjustment[14], i.e. from the longitudinal samples. This is the procedure followed by Statistics Canada methodologists to combine the surveys SHP_I wave 6 and SHP_II wave 1 carried out in 2004.
2. The alternative is to combine cross-sectionally (from cross-sectional samples), i.e. after adjustment for non-response to the individual questionnaire for the individual cross-sectional weighting and after adjustment for non-response to the household questionnaire for the household cross-sectional weighting.

As was the case last year, it is the second method which is applied and described in this document. It should be noted that it is similar to the one used in the SILC survey in other countries (see [15]).
It should also be noted that the surveys are combined at the level of the major regions, as the factors of panel allocation[15] vary from one region to another. Therefore seven different allocation factors are calculated (one per region) instead of calculating just one at Swiss level.
For a more detailed discussion of the two alternatives for combining the panels, please refer to references [12], [13] and [15].
The content of the data files used to construct the weightings are described in Appendix A.

5.2 Types of weightings produced

For wave 9 of the SHP_I and wave 4 of the SHP_II, three weightings were produced:

1 – household cross-sectional weight
- *representing the household population in 2007 (waves 9 and 4)*

2 – individual cross-sectional weight
- *representing the population of individuals aged 14+ in 2007 (waves 9 and 4)*

3 – individual longitudinal weight
- representing the population of individuals aged 14+ in 1999 for the SHP_I wave 9; in 2004 for the SHP_II wave 4.

For the combined samples SHP_I wave 9 and SHP_II wave 4, the same three types of weightings were produced. The longitudinal weighting of SHP_I is used for analyses which cover more than 5 years. For all other operations, the combined weightings

[14] Statistics Canada go by [14] to make the combination with longitudinals alone.
[15] See paragraph 10.1 for more details regarding the panels' allocation factors.

are generally used as they allow all data collected in 2007 to be used and have the advantage of a larger sample, which results in improved precision.

1 – household cross-sectional weight

- representing the population of household in 2007

2 – individual cross-sectional weight

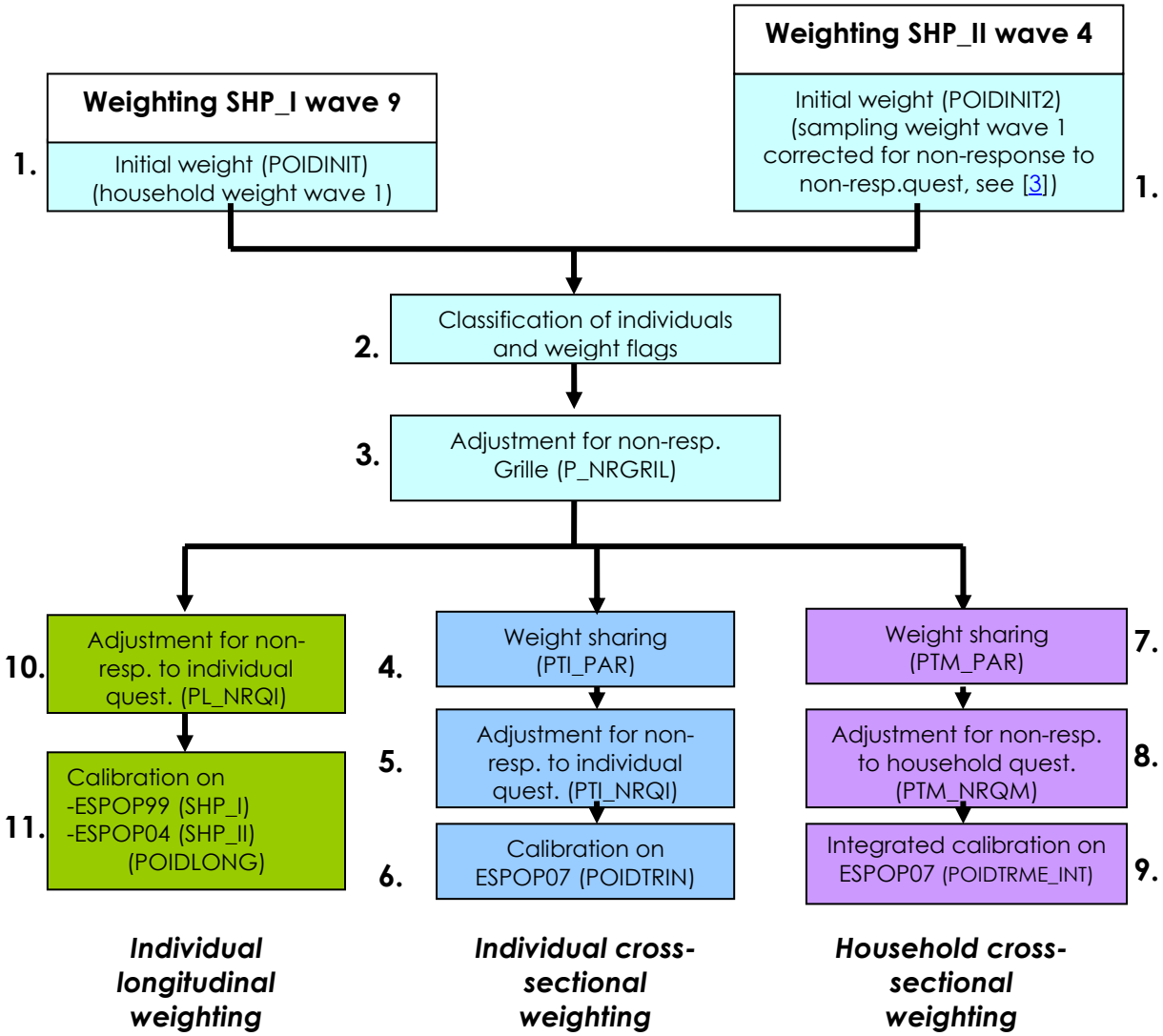
- representing the population of individuals aged 14+ in 2007

3 – individual longitudinal weight

- representing the population of individuals aged 14+ in 2004

The three first steps in the weighting procedure are the same for all three types of weights, whereas the following are specific to each one. Diagram 18 summarises the weighting process for wave 9 of SHP_I and wave 4 of SHP_II, whereas Diagram 19 deals with the weighting process for the combined samples.

Diagram 1 Steps in the weighting process of SHP_I wave 9 and SHP_II wave 4. In brackets, in the boxes, features the name of the set of weights created at the step under consideration. Step 9a was only carried out for the household cross-sectional weighting of the SHP_I.



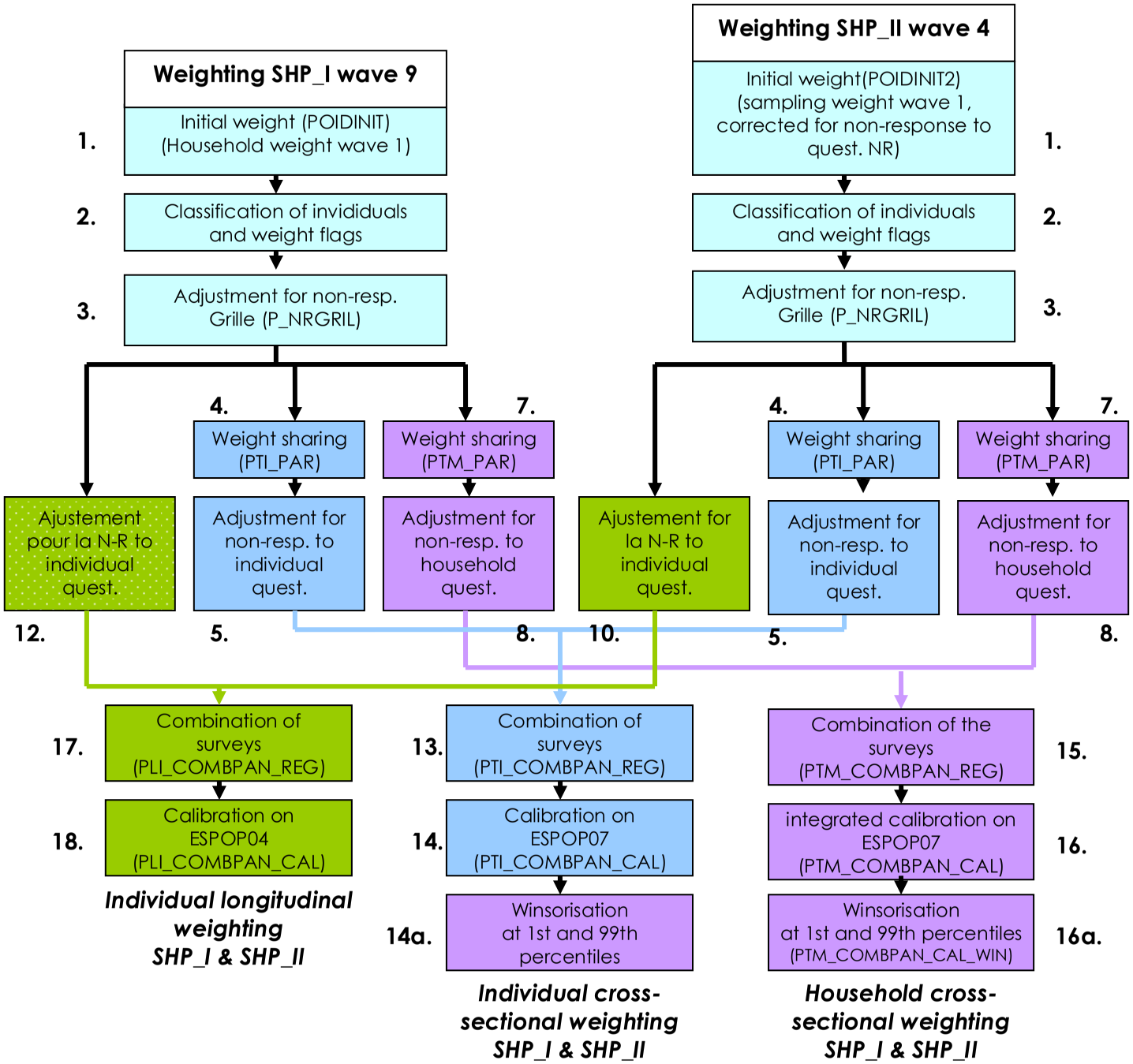


Diagram 19 Steps of the weighting enabling aggregation of the sample of SHP_I wave 9 with that of SHP_II wave 4. There is an additional step compared with the separate weighting by survey. This is the combining of the samples. The samples have to be combined for each weighting (household and individual). For the cross-sectional weightings, the steps starting with the calculation of initial weights and going to the adjustment for non-response at individual (or household) level are the same as those calculated during the separate weightings of the two surveys. For the combined longitudinal weighting, a new adjustment has to be made for the non-response to the individual questionnaire for the individuals of the SHP_I (see chapter 12 for more details). The name of the set of weights created at the step under consideration always features in brackets in the boxes.

6 Steps common to all types of weights

With reference to Diagrams 18 and 19, this section describes steps 1 to 3 which are common to the production of weights at both individual and household level, cross-sectionally as well as longitudinally. We proceeded in the same way to weight wave 9 of the SHP_I and wave 4 of SHP_II with the small difference that the initial weight of SHP_II was obtained after additional adjustment for the non-response to the (non-response) questionnaire (see [3]). After a short introduction to the segmentation method, we discuss the calculation of initial weight, then the way in which individuals are classified according to their status and their response and lastly the adjustment for non-response to the grid.

6.1 Modelling of non-response by the segmentation method

6.1.2 Segmentation method

Once the initial weights have been calculated, an adjustment factor has to be calculated for non-response (total) for each unit (household or individual according to each case) considered as respondent. The adjustment factor for non-response applied to each respondent unit is defined as being the inverse of the response rate of the response homogeneity group (RHG) in which it is situated. The RHG are made by bringing together units which share similar characteristics explaining their non-response. If the RHG are defined in such a way that the non-response is completely random within a group, the bias due to non-response becomes negligible. In order to find a suitable adjustment to boost the weight of responding units so that the latter are representative of all units (respondent and non-respondent), we carry out modelling of non-response.

The first step of modelling the non-response is to determine the explanatory variables which have to be considered in the model in order to be able to predict the dichotomous response variable. The choice of variables is often quite limited as we need information on both respondent and non-respondent units. This involves making a list of available variables which could describe the non-response process. The nature of this list varies according to the themes the survey is dealing with, the type of interview, the sampling plan etc. Every factor which might explain non-response has to be taken into account as far as possible. The variables may be the values measured in the survey itself or additional information that can be obtained by other means (for example registers) on the respondent and non-respondent units.
In a survey panel such as the one we are interested in here, from wave 2 on, we have a much wider choice of variables to model the non-response to household (see paragraph 8.2) or individual (see paragraph 7.2) questionnaires, as at that stage we can use information gathered in wave 1.
It should also be noted that category variables with a number of modalities (m) greater than 2, must be recoded into *m* dichotomous variables.

The method used for modelling the non-response is analysis by segmentation (see [8]). This method is used for all adjustments necessary due to non-response and which are presented in this document.
Modelling by segmentation uses an iterative process to partition the data file and thus make the decision tree. The first knot in the decision tree is quite simply the

complete set of data. The first iteration defines the first branch of the tree. The latter is made by determining the variable which has the most significant influence on the response variable among those variables used for the modelling. A statistical measure which reflects the importance of the variable in order to differentiate non-response is calculated for each variable; after this has been done the variable which has the greatest significance is chosen. The first branch is thus formed, creating a knot at each end. For each knot formed, a new iteration is carried out by repeating the same mechanism. In this way new branches and new knots are created. If for a given knot, no variable is significant to explain the non-response, or one of the stopping criteria for the algorithm is met, no new branch is made.

Several methods can be used to determine the variable which has the most influence on the answer (see [22]).Here we use the CHAID[17] algorithm, which selects the variable with the highest Pearson's chi-square statistical value.

6.1.2 Segmentation and logistic regression

An alternative for modelling the non-response would be to use logistic regression. As we wished to remain as compatible as possible with the weightings from previous years, we did not call into question the method which has been used for 9 years.
Furthermore it should be mentioned that the authors of [7] have shown in two empirical studies[19] similar to the SHP, that the segmentation method was better than logistic regression for creating homogeneous response groups (RHG) and consequently for reducing the bias. In fact, the segmentation method is more flexible than logistic regression. Logistic regression forces the model to be symmetric: once the significant variables have been determined, the RHG are formed by taking all the combinations possible from the intersection of these variables. By forcing the model to be symmetric, small cells or cells which are not significant can be formed, which can affect the reliability of estimates. Modelling by segmentation, for its part, determines at each branch of the decision tree which variable is the most significant, thus guaranteeing that each RHG formed is significant. This usually results in an asymmetric model. This also presents the advantage of being able to use a great number of variables to model the non-response.

A disadvantage of modelling by segmentation is directly related to the decision tree itself. When a decision tree is made, it can become quite heavy: as soon as it has several branches, it becomes more difficult to interpret. Decision trees are also quite unsTable: adding, removing or changing a variable or unit may have a major effect on the tree itself.

Table 20 summarises some comparative characteristics of the modelling methods of non-response by segmentation and by logistic regression.

---

[17] CHAID: Chi-Square Automatic Interaction Detection, algorithm invented by Kass (see [8]).

[19] SLID (*Survey of Labour and Income Dynamics*), a rotating panel, and NLSCY (*National Longitudinal Survey of Children and Youth*) a simple panel.

Table 20 Comparison between logistic regression and segmentation method for the modelling of non-response

Logistic regression	Segmentation
➢ Step by step procedure (stepwise) to obtain the list of the most significant explanatory variables of the predisposition to respond. ➢ The RHG are created on the basis of 2^k combinations of the k predictors selected. ➢ Approach which can be qualified as symmetric. ➢ We could add the constraints for the RHG (typically n>30 and the response rate >50% by RHG). ➢ The number of variables which can be used to model the non-response is limited.	➢ Uses the CHAID algorithm which divides the sample into smaller sub-groups according to their response rate until no significant predictor can be found. The final sub-groups form the RHG. ➢ Approach which can be qualified as non-symmetric. ➢ The constraints have to be satisfied for each RHG (e.g. n>30 and response rate >50%: the choice of these constraints has a direct influence on the final number of RHG). ➢ Can use a wider range of variable to model the non-response. But the tree of segmentation can become unstable if too many explanatory variables are selected or if some of them are too interrelated.

6.1.3 Application for the SHP

We use an SAS macro which was developed by Jean-François Naud of Statistics Canada when he received the mandate to carry out weightings for the Swiss Household Panel for the first time in 2001. The macro enables modelling of the non-response for several types of adjustment (please refer to figures 18 and 19 for the definition of the steps mentioned):

- adjustment for non-response to the questionnaires grid (parameter `Type_aj=1`), step 3.
- adjustment for non-response to the individual questionnaire for the longitudinal weighting (parameter `Type_aj=2`), step 10.
- adjustment for non-response to the individual questionnaire for the cross-sectional weighting (parameter `Type_aj=3`), step 5.
- adjustment for non-response to the household questionnaire for the cross-sectional weighting (parameter `Type_aj=4`), step 8.

The weight before adjustment of the respondent units of each RHG is multiplied by a factor which corresponds to the inverse of the weighted response rate of their RHG.

The minimum requirements (stopping criteria for the algorithm in determining additional branches, in addition to finding no further significant variable explaining the response variable), for each RHG have been, depending on the case, either 50%

of the minimal response rate and 30 units at least, or 30% of response rate and at least 50 units.

6.2 Initial weights of SHP_I and initial weights of SHP_II

Step 1

For the SHP_II, as was the case for the previous years, the initial weight is the household weight calculated at wave 1 (see [2]). This is the sampling weight corrected for non-response to the grid and to the household questionnaire in 1999. The correction factors were then calculated using the logistic regressions model.

$$POIDINIT_k^{SHP_I} = \frac{1}{\pi_k \cdot p_k^1 \cdot p_k^2 \cdot p_k^3 \cdot f}$$

where: π_k is the sampling weight of household k (wave 1, 1999)

p_k^1 is the probability of response to the grid in wave 1 (assuming uniform non-response) given the sampling plan

p_k^2 is the probability of response to the household questionnaire in wave 1 (obtained by a logistic non-response model) given the distribution of the grid

p_k^3 is the probability of at least one individual questionnaire being completed in wave 1 (obtained by a logistic non-response model), i.e. that the household should be complete, given the distribution (of response) to the grid and to the household questionnaire

f is a correction factor which was applied in wave 1 so that the household weights multiplied by the size of the household are equal to the size of the resident population in Switzerland according to the ESPOP 1999.

We should remember that in 2000, during wave 2 of the SHP_I, only those households complete in wave 1 were contacted again. All households for which only the grid questionnaire or the grid questionnaire and the household questionnaire had been completed, were rejected. Thus it makes sense to take the household weight *POIDINIT* described above as the initial weight.

For the SHP_II, the initial weight is the POIDINIT2 calculated as being the sampling weight in wave 1 (2004) multiplied by a corrective adjustment factor for non-response to the non-response questionnaire of wave 1. This adjustment factor was calculated on a segmentation model described in [3].

$$POIDINIT2_k^{SHP_II} = \frac{1}{\pi_k \cdot f_k^{QNR}}$$

where: π_k is the sampling weight of household k (wave 1, 2004)

f_k^{QNR} is a corrective factor obtained by the segmentation method obtained for the non-response of household k to the non-response questionnaire.

Remember that in 2005, during wave 2 of SHP_II, only households which had completed at least a grid questionnaire in wave 1 were contacted again. Households which had responded to the non-response questionnaire were permanently rejected. Thus it is normal to take the household weight *POIDINIT2* described above as initial weight.

The descriptive statistics of the initial weights of the SHP_I and SHP_II are in the Appendix H.

6.3 Classification of individuals: step 2

This step aims to create a basic weighting file which contains all individuals characterised by the classification variables used for the subsequent steps.

The file contains all individuals from the households which responded to a list the grid questionnaire in wave 1 and in the current wave[22]. For SHP_I, all individuals can be found from INDIVIDUS 1999, INDIVIDUS 2007 et LONGOUT 2007, for SHP_II, all those from INDIVIDUS 2004, INDIVIDUS 2007 et LONGOUT 2007 (these files are described in Appendix A). The variables contained in the "classification" file are as follows:

[20] Households responding to the non-response survey in 2004 thus have no weight (final weight = 0) in contrast to what was done in wave 1 by Statistics Canada (see [4]). In the later stages of the weighting, these households are not taken into consideration at all. The information that they provided is "contained" in the initial weights (POIDINIT2).

[22] At this stage, the SHP_I is dealt with separately from SHP_II. There is, therefore, a separate classification for these two surveys.

Name of variable	Included in the classification file SHP_I v9	Included in the classification file SHP_IIv4	Specifications
IDPERS	(x)	(x)	Individual identifier
idhous_o	(=IDHOUS99)	(=IDHOUS04)	household identifier at original wave (=wave 1)
idhous_c	(=IDHOUS07)	(=IDHOUS07)	household identifier at current wave (=wave 2007) (= wave 9 for SHP_I, wave 4 for SHP_II)
f_plong	(x)	(x)	Indicator of longitudinal weight
f_ptrin	(x)	(x)	Indicator of individual cross-sectional weight
f_ptrme	(x)	(x)	Indicator of household cross-sectional weight
agecorro	(x)	(x)	age of individual at original wave
agecorrc	(x)	(x)	age of individual at current wave
poidinit	(x)		Initial household weight wave 1, SHP_I (step 1)
poidinit2		(x)	initial weight corrected for the NR or QNR (step 1), SHP_II
longst	(x)	(x)	status of individual (longitudinal/cohabiting)
coderep	(x)	(x)	response code at current wave
cr_vx	cr_v2 to cr_v8	cr_v1 to cr_v3	response code at wave x
	13752 obs.	6813 obs.	

Details on how the variables were constructed can be found in Appendix B.

Tables 21 and 22 give the distributions of individuals according to their status, their response code and whether they should have a non-zero weight.

Table 21 Distribution of individuals from SHP_I and SHP_II according to their status and response code17. The number of longitudinal adults and children does not generally change from one wave to another (for SHP_I) as reference is always made to the status that individuals had in the year of selection of the panelised sampling. After a replausibilisation of age, two longitudinal children from SHP_II were changed to adults compared with the previous year. Still for SHP_II, the numbers are given without including individuals from households having responded to the non-response questionnaire in 2004. After data plausibilisations, we have here 8 (4) fewer individual questionnaires less carried out for the SHP_I (SHP_II) compared with the numbers given in the Table 8.

	Individual status - LONGST									
	Longitudinal adult		Longitudinal child		Initially present cohabitant		Initially absent cohabitant		**TOTAL**	
Response code **CODEREP**	SHP_I	SHP_II	SHP_I	SHP_II	SHP_I	SHP_II	SHP_I	SHP_II	SHP_I	SHP_II
0 – < 14 at the current wave	0	0	1051	796	30	16	488	100	1569	912
1 – Cross-sectionally outside target	189	41	5	1	0	0	0	0	194	42
2 – Total non-respondent	5382	2610	679	112	0	0	0	0	6021	2722
3 – Grid respondent only	134	117	18	5	7	8	1	0	160	130
4 – Only grid and household questionnaire respondent, in a household where all individuals are non-respondent to individual questionnaire	72	68	16	5	4	3	2	0	94	76
5 – Only grid and household questionnaire respondent, in a household where at least one individual is respondent to the individual questionnaire	844	524	201	33	152	56	19	8	1216	621
6 – Grid respondent, household and individual questionnaire	3768	2136	607	112	218	89	28	1	4621	2338
7 – Grid respondent and individual questionnaire only	2	8	3	1	0	0	0	0	5	9
Total	**10351**	**5504**	**2580**	**1065**	**411**	**172**	**538**	**109**	**13880**	**6850**

[23] Children aged under 14 are all classified as cross-sectionally out of scope.
[24] See paragraph 4.1.3 for the definitions of cohabitants.

Table 22 Distribution of individuals from SHP_I and SHP_II according to their longitudinal household cross-sectional weight[26], and individual cross-sectional weight. For SHP_II, the numbers are given without including individuals from households who responded to the non-response questionnaire in 2004

Weight indicator	SHP_I V7	SHP_I V8	SHP_I V9	SHP_II V2	SHP_II V3	SHP_II V4
F_PLONG = 0 – longitudinal weight = 0	10166	10135	9921	4137	4360	4665
F_PLONG = 1 – longitudinal weight > 0	3495	3617	3959	2589	2453	2185
Total	**13661**	**13752**	**13880**	**6726**	**6813**	**6850**
F_PTRME = 0 – cross-sectional hh weight = 0	8469	9664	7949	4078	4245	3815
F_PTRME = 1 – cross-sectional hh weight > 0	5192	4088	5931	2648	2568	3035
Total	**13661**	**13752**	**13880**	**6726**	**6813**	**6850**
F_PTRME = 0 – cross-sectional ind. weight = 0	9776	8380	9254	3026	3374	4503
F_PTRME = 1 – cross-sectional ind. weight > 0	2885	5372	4626	3700	3439	2347
Total	**13661**	**13752**	**13880**	**6726**	**6813**	**6850**

6.4 Adjustment for non-response – grid: step 3

All households chosen to be reactivated for the 2007 survey and in which longitudinals were living, were likely to respond to the grid.
However, as the longitudinal unit is the individual and not the household, the probability of response to the grid has been modelled at the individual level. We therefore model the probability that an individual belongs to a household having responded to the grid. Adult and child longitudinals are used. They are however separated according to whether they are children from the first step of modelling. The sum of the longitudinal weights before and after adjustment for non-response to the grid remains the same for children and for adults. Modelling is carried out based on the household and individual data from wave 1.
We created a file which only contained the adult longitudinals (LONGST=1) and children (LONGST=2). We defined the variable:

REPGRIL = 0 Non-respondent to the grid (CODEREP=0, 2).
1 Respondent to the grid (CODEREP=3, 4, 5, 6 or 7).

Table 23 Distribution of longitudinals according to response to the grid. We have 11686 and 5731 individuals in total as we removed cohabitants and individuals cross-sectionally out of scope, as well as those < 14 at the current wave (see Tables 21 et 22)

Response to the grid (REPGRIL)	Number of longitudinals (Proportion %)		
	Adults	Children	Total
Non-respondent to the grid SHP_I	5342	679	**6021 (51.5)**
SHP_II	2610	112	**2722 (47.5)**
Non-respondent to the grid SHP_I	4820	845	**5665 (48.5)**
SHP_II	2853	156	**3009 (52.5)**
Total SHP_I	**10162**	**1524**	**11686**
SHP_II	**5463**	**268**	**5731**

[26] Children aged under 14 have F_PTRME=0 even if their household has POIDTRME > 0.

For SHP_I, the constraints of the segmentation model had to be eased compared with previous years: a weighted response rate of at least 30% and a minimum number of 50 individuals were required for each RHG. For SHP_II the constraints were a weighted response rate of at least 50% and a minimum number of 30 individuals for each RHG.
No adjustment for non-response was made for individuals cross-sectionally out of scope. (CODEREP=1). Their adjusted weight for non-response will be the same as the initial weight.

The non-response model created with the analysis by segmentation lead to the creation of 64 RHG for the SHP_I and 6 for the SHP_II. Table 24 gives some information on the adjustment factor thus calculated and applied to the initial weights to obtain the P_NRGRIL weights.

Table 24 Descriptive statistics of the adjustment factor, f_aj_nrgril , for the 5665 and 3009 longitudinal individuals in the households responding to the grid (those for whom REPGRIL=1), see Table 23). The 194 and 45 cross-sectionally out-of-scope individuals also have a P_NRGRIL>0 weight but do not undergo any adjustment for non-response to the grid (i.e. P_NRGRIL=initial weight, f_aj_nrgril=1).
SHP_I: P_NRGRIL = f_aj_nrgril * POIDINIT
SHP_II: P_NRGRIL = f_aj_nrgril * POIDINIT2

	SHP_I	SHP_II		SHP_I	SHP_II
Average	2.06	1.92	Min	1.18	1.45
Median	1.87	1.98	Max	3.33	1.98
Standard deviation	0.61	0.15	Range	2.15	0.53

Details of this modelling of non-response can be found in Appendix C.

7 Individual cross-sectional weighting

In this chapter, we describe steps 4, 5 and 6 of figure 18. The procedure is the same for the SHP_I wave 9 and SHP_II wave 4. The starting weight of the individual cross-sectional weighting is the adjusted weight for non-response to the grid of the longitudinals aged 14 and over at the current wave. The step of sharing weights allows weights to be attributed to cohabitants who join longitudinal households between the first and the current wave. We then proceeded to an adjustment for non-response to the individual questionnaire and lastly to a marginal calibration by using the control totals for the resident population in private households in Switzerland in 2007.

7.1 Weight sharing: step 4

After the adjustment for non-response to the grid step, only longitudinals from households responding to the grid have a P_NRGRIL non-zero weight. The step of sharing weights allows a part of these weights to be attributed to cohabitants who joined longitudinal households between the first and the current wave. The sharing of weights is carried out separately for each of the two panels SHP_I and SHP_II.
For theoretical considerations and validations as to the method of weight sharing, see [9].
Please note that longitudinal children who reached the age of 14 at the current wave are treated like longitudinals and not like initially absent cohabitants. In this way we take into account the possibility that a longitudinal child of more than 14 years may be a household's only longitudinal individual.

Procedure

We take all individuals belonging to the cross-sectional target population and who responded to the grid (CODEREP = 3,4,5,6 or 7 variable). The basic weighting (before sharing) is P_NRGRIL. It is nil for all cohabitants.
The weight after sharing is PTI_PAR. **This weight may be different for individuals from the same household.**
The shared weight is calculated as follows:
i : individual
L : number of longitudinals in the household
P : number of initially present cohabitants in the household

- *if the household includes initially present cohabitants*, the shared weight is the same for everyone and is equal to the sum of the longitudinal weights divided by the number of longitudinals and initially present cohabitants in the household.

$$PTI_PAR_i = \frac{\sum_{i=1}^{L} P_NRGRIL_i}{L+P}$$

- *if there is no initially present cohabitant in the household*, then the weights of the longitudinals stays the same and those of the initially absent cohabitants are the same as the average weights of the longitudinals.

$$PTI_PAR_i = \begin{cases} P_NRGRIL_i & \text{for i longitudinals} \\ \dfrac{\sum_{i=1}^{L} P_NRGRIL_i}{L} & \text{For i initially absent cohabitants} \end{cases}$$

We use the SAS `%partage()` macro developed by Jean-François Naud of Statistics Canada to carry out weight sharing. For more details regarding the use of this macro please refer to its source code.

7.2 Adjustment for non-response – grid: individual questionnaire: step 5

Once the household has completed the grid, individuals can respond to the individual questionnaire regardless of whether the household questionnaire has been completed. This is why, to be consistent with the collection process and because little information from the household questionnaire proved useful for the modelling, we carried out the adjustment for non-response to the individual questionnaire directly from the shared weight which represents individuals in households having responded to the grid.

Procedure

We created a file containing individuals (longitudinal adults and cohabitants) whose households responded to the grid and belonged to the cross-sectional target population, i.e. all those for whom PTI_PAR > 0.

We defined: REPQI = 0 Non-respondent to the individual questionnaire (CODEREP=3, 4 or 5)
1 Respondent to the individual questionnaire (CODEREP=6 or 7).

Table 25 Distribution of adults belonging to the cross-sectional target population and who belong to households responding to the grid according to the response to the individual questionnaire.

	SHP_I Wave 9	*SHP_II Wave 4*
Response to I.Q. (REPQI)	Number of individuals (Proportion as %)	Number of individuals (Proportion as %)
Non-respondent to I.Q.	1470 (24.1)	827 (26.1)
Respondent to I.Q.	4626 (75.9)	2347 (73.9)
Total	**6096**	**3174**

Table 26 Descriptive statistics of the adjustment factor, f_aj_nrtiqi, for the 4626 and 2347 individuals responding to the individual questionnaire (those for whom REPQI=1, see Table 25). PTI_NRQI = P_NRGRIL * f_aj_nrtqi.
In both surveys there are individuals for whom f_aj_nrtqi=1 because, as we can see in the Appendix E, some homogeneous response groups formed by the segmentation algorithm reached a response rate of 100%, hence no adjustment for those RHG.

	SHP_I	SHP_II		SHP_I	SHP_II
Average	1.32	1.36	Min	1.00	1.00
Median	1.24	1.27	Max	3.29	1.96
Standard deviation	0.31	0.28	Range	2.29	1.00

The adjustment is made at the individual level. Given that the number of cohabitants is very low (see Table 21), it was necessary to carry out the modelling of longitudinals and cohabitants together. Data coming from the grid of the current wave were used for the modelling.
For SHP_I wave 9, the model obtained with analysis by segmentation contained 39 RHG and 23 RHG for the SHP_II wave 4 (see Appendix E). As was the case for the adjustment for non-response to the grid, a weighted response rate of at least 50% and a minimum number of 30 individuals were required for each RHG for the SHP_II, whereas for the SHP_I, the constraints were a response rate of 30% and 50 units for each RHG.

The adjusted weight for non-response (individual questionnaire) is PTI_NRQI.

7.3 Calibration: step 6

The cross-section adjusted weight for non-response to the individual questionnaire (PTI_NRQI) represents the population of Switzerland at the current wave. The calibration step adjusts these weights so that the totals of certain control variables are the same as those of the Swiss population in 2007 according to the ESPOP (from December 2007). This step is carried out for both surveys. The totals can be found in Table 28.

Weight before calibration: PTI_NRQI

Control totals
1. Sex x (14-24 years; 25-34 years; 35-44 years; 45-54 years; 55+ years)
2. Nationality (Swiss, foreign)
3. Region (Lake Geneva region & Central Switzerland & Ticino; Espace Mittelland & North-western Switzerland & Zurich & Eastern Switzerland)
4. Civil status (married; not married)

Final weight: POIDTRIN[27]
Table 27 gives some numbers which characterise the calibrations carried out.

[27] In the descriptives of the weightings given to users, the variable final individual cross-sectional weight, thus POIDTRIN, is renamed:
- Wp07TP1P for the SHP_I wave 9
- Wp07TP2P for the SHP_II wave 4

Table 27 Characteristics of marginal calibrations, individual cross-sectional weighting. We have 4626 and 2347 observations with a POIDTRIN>0 for SHP_I et SHP_II.

	Calibration adjustment factor = POIDTRIN/PTI_NRQI		*POIDTRIN*	
	SHP_I	*SHP_II*	*SHP_I*	*SHP_II*
Average	0.92	0.88	1358	2676
Median	0.84	0.77	1033	2232
Standard deviation	0.31	0.40	1043	1508
Min	0.61	0.49	166	257
Max	2.40	2.59	10736	12306
Range	1.79	2.11	10569	12050

Table 28 Estimated numbers of the population before and after calibration for SHP_I. Individual cross-section weight. Note that these totals reflect the population aged 14 and above. The total of the pti_nrqi du SHP_II wave 4 is higher due to numerous corrections for non response (non-response questionnaire, grid questionnaire, individual questionnaire). In fact, the corrected weights for non-response generally add up to a higher number than the sum of weights after calibration. The larger and more selective the attrition, the more pronounced the phenomenon.

	Estimate with the weight before calibration (PTI_NRQI)		*Population (and estimate with the weight after calibration POIDTRIN)*
Variable	*SHP_I wave 9*	*SHP_II wave 4*	
Male 14-24 years	631598	741323	485575
Male 25-34 years	339159	324648	485027
Male 35-44 years	569641	659394	610375
Male 45-54 years	549138	719048	550107
Male 55 years and over	1065601	1022071	944765
Female 14-24 years	559773	545506	461919
Female 25-34 years	329447	395063	490511
Female 35-44 years	646153	752947	608983
Female 45-54 years	633059	820173	543721
Female 55 years and over	1231258	1270238	1099785
Swiss	5832564	6576356	4955440
Foreign	722261	674055	1325328
Geneva, Central and Ticino	2300250	2286437	1999361
Mittelland, N-W, Zurich, Eastern	4254575	4963974	4281408
Not married	3028509	3210740	2867978
Married	3526316	4039671	3412790
Total	**6554824**	**7158158**	**6280768**

8 Household cross-sectional weighting

In this chapter, we describe steps 7, 8 and 9 of diagram 18. The procedure is the same for the SHP_I wave 9 and SHP_II wave 4. The starting weight of the household cross-sectional weighting is also the adjusted weight for non-response to the grid of the longitudinal adults. The step of sharing weights allows weights to be attributed to cohabitants who join longitudinal households between the first and the current wave and to obtain a weight at household level. We then proceeded to an adjustment for non-response to the household questionnaire and lastly to a marginal calibration by using the control totals for the resident population in private households in Switzerland in 2007.

8.1 Weight sharing: step 7

After the adjustment step for non-response to the grid, all longitudinals from households responding to the grid have a weight of P_NRGRIL > 0. The step of weight sharing consists in attributing a part of these weights to cohabitants who have joined the longitudinal households since wave 1.

As in the case of the individual cross-sectional weighting, it should be noted that longitudinal children who have reached the age of 14 are treated as longitudinals and not as initially absent cohabitants as we could also have chosen to do. This is in order to take account of the possibility that a longitudinal child aged over 14 may be a household's only longitudinal individual.

Procedure

We take all individuals belonging to the cross-section target population and who responded to the grid (CODEREP = 3,4,5,6 or 7 variable). The basic weight (before sharing) is P_NRGRIL. The weight after sharing is PTI_PAR. **This weight is the same for all of the individuals from a household.**

The shared household weight is the same as the sum of the weight of longitudinals divided by the number of longitudinals and initially present cohabitants in the household.

$$PTI_PAR_i = \frac{\sum_{i=1}^{L} P_NRGRIL_i}{L + P}$$

i : individual
L : number of longitudinals in the household
P : number of initially present cohabitants in the household

The resulting file is at household level (identifier IDHOUS07).
As for the individual cross-sectional weighting (see paragraph 7.1), we use the `%partage()` macro to carry out the weight sharing.

8.2 Adjustment for non-response – grid: individual questionnaire: step 8

After the weight sharing step, all households having responded to the grid of the current wave have a cross-sectional weight. As only households who responded to the household questionnaire have a household cross-sectional final weight, another adjustment for non-response has to be carried out, this time at household level. Only data from the grid (**current wave**) are used for the modelling. The modelling is done for both surveys (separately).

Procedure

We created a file containing the households that had responded to the grid and that belonged to the cross-sectional target population, i.e. all those for whom PTM_PAR > 0 and that had the correct response code (CODEREP=3, 4, 5, 6 or 7).

We defined: REPQM= 0 Non-respondent to the household questionnaire (CODEREP = 3 or 7 for all individuals)
1 Respondent to the household questionnaire (CODEREP = 4,5 or 6 for all individuals).

Table 29 Distribution of households belonging to the cross-section target population and who responded to the grid according to the response to the household questionnaire.

	SHP_I Wave 9	*SHP_II Wave 4*
Response to QM (REPQM)	Number of households (Proportion as %)	Number of households (Proportion as %)
Non-respondent to the household questionnaire	76 (2.6)	54 (3.5)
Respondent to the household questionnaire	2817 (97.4)	1489 (96.5)
Total	**2893**	**1543**

Modelling is based on the grid of the current wave i.e. either wave 9 for SHP_I or wave 4 for SHP_II. The model obtained with the analysis by segmentation includes 10 RHG for SHP_I and 7 for SHP_II. A weighted response rate of at least 30% and a minimum of 50 individuals were required in each RHG for the SHP_I; for SHP_II the constraints were 50% and 30 individuals per RHG.

The adjusted weight for non-response (individual questionnaire) is PTM_NRQM. Table 30 shows some numerical characteristics of this adjustment.

Table 30 Descriptive statistics of the adjustment factor, f_aj_nrtm, for the 2817 and 1489 responding households (see Table 29) for SHP_I and SHP_II. PTM_NRQM = f_aj_tm_nrqm * PTM_PAR. As can be seen in Appendix F, some RHG formed have a response rate of 100%, that is why the minimum is equal to one for the adjustment factor.

	f_aj_tm_nrqm	
	SHP_I wave 9	SHP_II wave 4
Average	1.03	1.04
Median	1.00	1.00
Standard deviation	0.09	0.11
Min	1.00	1.00
Max	1.44	1.45
Range	0.44	0.45

For more details on the modelling, please refer to Appendix F.

8.3 Calibration: step 9

As for the individual cross-sectional weighting, we wish the number of persons in households to represent the total of the resident population of Switzerland. Here an integrated marginal calibration[28] is made for the household weight by using the Swiss population 2007 (current wave), according to the l'ESPOP in December 2007.

The cross-section adjusted weight for non-response to the household questionnaire (PTM_NRQM) represents the population of Switzerland at the current wave. The calibration step adjusts these weights so that the totals of certain control variables are the same as those of the population in 2007. This step is carried out for both surveys. The totals can be found in Table 32.

Weight before calibration: PTM_NRQM

Control totals

1. Sex x ((< 14 years; 14-24 years; 25-34 years; 35-44 years; 45-54 years; 55+ years)
2. Nationality (Swiss, foreign)
3. Region (Lake Geneva region & Central Switzerland & Ticino; Espace Mittelland & North-western Switzerland & Zurich & Eastern Switzerland)
4. Marital status (married; not married)

[28] In an integrated calibration, we wish to satisfy the totals of each margin at individual level, while obliging the final weights (after calibration) of individuals from the same household to remain equal.

For the SHP_I, the considerable attrition that the sample underwent can be seen at this step of the calibration. In fact, as in the previous year only one household was attributed a negative weight after calibration, which is unacceptable. As the methodologists from Statistics Canada suggested (see [10]), one way of solving this problem is to carry out a second calibration after having first reattributed the weight before the first calibration (i.e. PTM_NRQM) to the observations which were proving to be problematic.
For SHP_II one calibration was enough.
Tables 31 et 32 give some characteristic numbers from the calibrations carried out for SHP_I and SHP_II.

Calibrated weight: POIDTRME_INT

Table 31 Characteristics of marginal calibrations, household cross-sectional weighting. We have 7034 individuals in 2817 households and 3617 individuals in 1489 households with a POIDTRME>0 for SHP_I and SHP_II. The characteristics of the adjustment factors are given at the level of individuals, whereas those of the weights after calibration are at household level.

	Calibration adjustment factor = weight before / weight after				*Calibrated weights*	
	SHP_I 1st calibration	*SHP_I 2nd calibration*	*SHP_II*	*SHP_I Weight after 1st calibration*	*SHP_I POIDTRME_INT 2nd calibration*	*SHP_II POIDTRME_INT*
	Individual level in households responding to the household questionnaire			*Household level responding to household questionnaire*		
Average	0.92	1.00	0.88	1045	1048	2037
Median	0.80	1.00	0.79	798	796	1756
Standard deviation	0.45	0.01	0.44	803	802	1101
Min	**-0.32**	0.93	0.14	**-353**	100	47
Max	3.72	1.01	2.77	7206	7287	7947
Range	4.05	0.08	2.63	7558	7187	7900

8.4 Winsorisation of weights after calibration: Step 9a

For the SHP_I, the ratio of maximum weight to minimum weight equals 73, which is less than the previous year (227.8), but in order to remain consistent from one year to the next we nevertheless decided to Winsorise the calibrated weights at the 1st and 99th percentiles. The same applied to SHP_II where the aforementioned ratio equals 169. Thus the max/min ratio changed to 20 for the SHP_I, and to 13 for the SHP_II. The integrated calibration (equal weight for all individuals of a household) represents, in addition to the margins which have to be respected, a major constraint for the problem of optimisation that it implies. This is why, with attrition such as that experienced by the SHP_I, the final weights include some extreme values which are not necessarily realistic.

Another way of avoiding extreme weights, apart from Winsorisation, would be for example to forego calibration on the civil status. However, we did not choose this option because it involves greater differences in the distribution of weights for 2007 compared with those of previous years than if we truncate the distribution as was done.

Final weight: POIDTRME_INT_WIN[29]

The descriptive statistics of the Winsorised weights can be found in Appendix H.

Table 32 Estimated numbers of population before and after calibration for SHP_I and SHP_II. Household cross-sectional weight. Please note that these totals reflect all of the population (all ages, including those under 14). The totals after Winsorisation obtained with the POIDTRME_INT_WIN are only slightly different than the ones here :

	Estimate with the weight before calibration (PTM_NRQM)		*Population (and estimate with the weight after calibration POIDTRME_INT)*	*POIDTRME_INT_WIN*	
Variable	*SHP_I Wave 9 (before 1st calibration*	*SHP_II Wave 4*		*SHP_I Wave 9*	*Wave 4 vague 4*
Male < 14 years	580389	711161	559468	562083	557453
Male 14-24 years	597207	712552	485575	485499	488611
Male 25-34 years	327991	349557	485027	478379	485480
Male 35-44 years	566482	645053	610375	607834	608230
Male 45-54 years	564549	685990	550107	550191	552705
Male 55+ years	1089160	1094235	944765	949319	946982
Female < 14 years	554507	651840	529354	530410	524566
Female 14-24 years	552785	559461	461919	462600	457426
Female 25-34 years	322006	399606	490511	486688	488868
Female 35-44 years	623909	728327	608983	610126	610781
Female 45-54 years	655691	753529	543721	542034	546768
Female 55+ years	1269935	1296985	1099785	1104428	1101721
Swiss	6908202	7681608	5781145	5824299	5806118
Foreign	796410	906687	1588445	1545291	1563472
Geneva, Central and Ticino	2719061	2731192	2364422	2364839	2373875
Mittelland, N-W, Zurich, Eastern	4985551	5857103	5005168	5004751	4995715
Not married	4120575	4586378	3956800	3960945	3955896
Married	3584037	4001917	3412790	3408645	3413694
Total	**7704612**	**8588296**	**7369590**	**7369590**	**7369590**

[29] In the descriptives of weightings given to users, the variable final household cross-sectional weight, thus POIDTRME, has been renamed:

- Wh07TP1P for the SHP_I wave 9
- Wh07TP2P for the SHP_II wave 4

9 Individual longitudinal weighting

Once the weights of the longitudinal adults have been adjusted for non-response to the grid, the longitudinal weighting continues with separate steps from those of the cross-sectional weighting. The weights of the longitudinal adults adjusted for non-response to the grill are adjusted for non-response to the individual questionnaire. Then, a marginal calibration allows the estimated number of the population to correspond to certain control totals of the Swiss population at wave 1 (we use the numbers from ESPOP 1999 for the SHP_I and those of ESPOP 2004 for the SHP_II). In this chapter, we describe steps 10, 11 and 6 of figure 19.

9.1 Adjustment for non-response – grid: individual questionnaire: step 10

Once the household has completed the grid, individuals can respond to the individual questionnaire regardless of whether the household questionnaire has been completed. This is why, to be consistent with the collection process and because little information from the household questionnaire proved useful for the modelling, we carry out the adjustment for non-response to the individual questionnaire directly from the shared weight for non-response to the grid.

Procedure

We created a file containing adult longitudinal individuals (LONGST=1) who are in households having responded to the grid or who are cross-sectionally out of scope (CODEREP=1,3,4,5,6 or 7) i.e. all those for whom P_NRGRIL > 0.

We defined: REPQI = 0 Non-respondent to the individual questionnaire (CODEREP=3, 4 or 5)
1 Respondent to the individual questionnaire (CODEREP=6 or 7).

Cross-sectionally out-of-scope individuals (CODEREP=1) do not undergo adjustment for non-response. Their adjustment factor is, therefore, equal to 1 (weight before adjustment = weight after adjustment) and the SAS macro for the segmentation `%aj_nr(…)` attributes them a RHG=. (missing).

Table 33 Distribution of adults belonging to the cross-section target population and who belong to households responding to the grid according to the response to the individual questionnaire. Please note that we have 189 and 41 cross-sectionally off-target longitudinals for SHP_I and SHP_II. The latter keep their P_NRGRIL. Note that these numbers can be found in Table 21.

	SHP_I Wave 9	*SHP_II Wave 4*
Response to I.Q. (REPQI)	Number of individuals (proportion as %)	Number of individuals (proportion as %)
0 - non-respondent to I.Q.	1050 (21.8)	709 (24.9)
1 - respondent to I.Q.	3770 (78.2)	2144 (75.1)
Total	**4820**	**2853**

Table 34 Descriptive statistics of the adjustment factor, f_aj_nrlqi, for the 3770 and 2144 individuals responding to the individual questionnaire (those for whom REPQI=1, see Table 33). PL_NRQI = P_NRGRIL * f_aj_nrlqi. The RHG that are formed where the response rate reached 100%, lead to an adjustment factor of 1.00 (i.e. no adjustment).
For cross-sectionally out-of-scope individuals (not included in the statistics of the above Table) we have f_aj_nrlqi=1.

	SHP_I	SHP_II		SHP_I	SHP_II
Average	1.28	1.34	Min	1.00	1.00
Median	1.20	1.28	Max	2.38	1.95
Std. deviation	0.26	0.28	Range	1.38	0.95

The adjustment is made at the individual level. Data coming from the grid of the **current wave** were used for the modelling.
For SHP-I wave 9, the model obtained with analysis by segmentation contained 26 RHG and 20 RHG for the SHP_II wave 4 (see Appendix D). As was the case for the adjustment for non-response to the grid, a weighted response rate of at least 50% and a minimum number of 30 individuals were required for each RHG for the SHP_II, whereas for the SHP_I, the constraints were a response rate of 30% and 50 units for each RHG.

The adjusted weight for non-response (individual questionnaire) is PL_NRQI.

The (3770+189) 3959 and (2144+41) 2185 of SHP_I and SHP_II having a non-zero weight PL_NRQI are those who will have a final non-zero longitudinal weight (F_PLONG=1).

9.2 Calibration: step 11

The adjusted longitudinal weight for non-response to the individual questionnaire (PL_NRQI) represents the population of Switzerland at the moment of selection of the sample of SHP at wave 1. The step of calibration adjusts these weights so that the totals according to certain control variables are the same as those of the Swiss population in 1999 according to the ESPOP for the SHP_I; in 2004 for the SHP_II. This step is carried out for both surveys. The totals can be found in Table 36.

Weight before calibration: PL_NRQI
Control totals[31]

1. Sex x (14-24 years; 25-34 years; 35-44 years; 45-54 years; 55+ years)
2. Nationality (Swiss, foreign)
3. Region (Lake Geneva region & Central Switzerland & Ticino; Space Mittelland & North-western Switzerland & Zurich & Eastern Switzerland)

[31] We do not calibrate according to martial status for the longitudinal weighting in order to ensure uniformity of this weighting over the years. In fact, for the cross-sectional weighting, we have calibrated according to martal status since 2004, waves 6 of SHP_I and 1 of SHP_II.

Final weight: POIDLONG[32]

Table 35 gives some numbers which are characteristic of the calibrations carried out.

Table 35 Characteristics of the marginal calibration (including cross-sectionally out-of-scope individuals), individual longitudinal weighting. We have 3959 and 2185 observations with a POIDLONG>0 for SHP_I et SHP_II.

	Calibration adjustment factor = POIDLONG/PL_NRQI		*POIDLONG*	
	SHP_I	*SHP_II*	*SHP_I*	*SHP_II*
Average	1.02	0.88	1515	2783
Median	0.93	0.85	1258	2476
Standard deviation	0.25	0.31	912	1284
Min	0.66	0.56	374	629
Max	1.96	1.90	8291	11265
Range	1.30	1.34	7917	10637

Table 36 Estimated numbers of the population before and after calibration for SHP_I. Individual cross-section weight. Please note that these totals reflect the population aged 14 and above.

	Estimate with the weight before calibration (PL_NRQI)		*Population (and estimate with the weight after calibration POIDLONG)*	
Variable	SHP_I wave 9	SHP_II wave 4	SHP_I (ESPOP1999)	SHP_II (ESPOP 2004)
Male 14-24 years	501917	685307	463320	469973
Male 25-34 years	507049	420987	530410	481790
Male 35-44 years	590288	730396	585647	611688
Male 45-54 years	524310	610018	497434	513613
Male 55 years and over	675158	916782	824282	888469
Female 14-24 years	387220	483780	445927	449624
Female 25-34 years	569875	463404	540607	491394
Female 35-44 years	641663	833054	571212	611586
Female 45-54 years	609298	748890	492648	509749
Female 55 years and over	764981	1064097	1046725	1052980
Swiss	5019412	6208605	4861829	4841446
Foreign	752348	748110	1136383	1239420
Geneva, Central and Ticino	1983768	2220163	1900541	1924056
Mittelland, N-W, Zurich, Eastern	3787991	4736551	4097671	4156810
Total	**5771759**	**6956714**	**5998212**	**6080866**

[32] In the descriptives of the weightings given to users, the variable final individual cross-sectional weight, thus POIDLONG, is renamed as:

- Wp07LP1P for the SHP_I wave 9
- Wp07LP2P for the SHP_II wave 4

10 Individual cross-sectional weighting SHP_I and SHP_II combined

10.1 Cross-sectional combination of panels from cross-sectional samples from each panel

The combined cross-sectional weighting is done by combining the two panels cross-sectionally.
The following abbreviations are agreed:

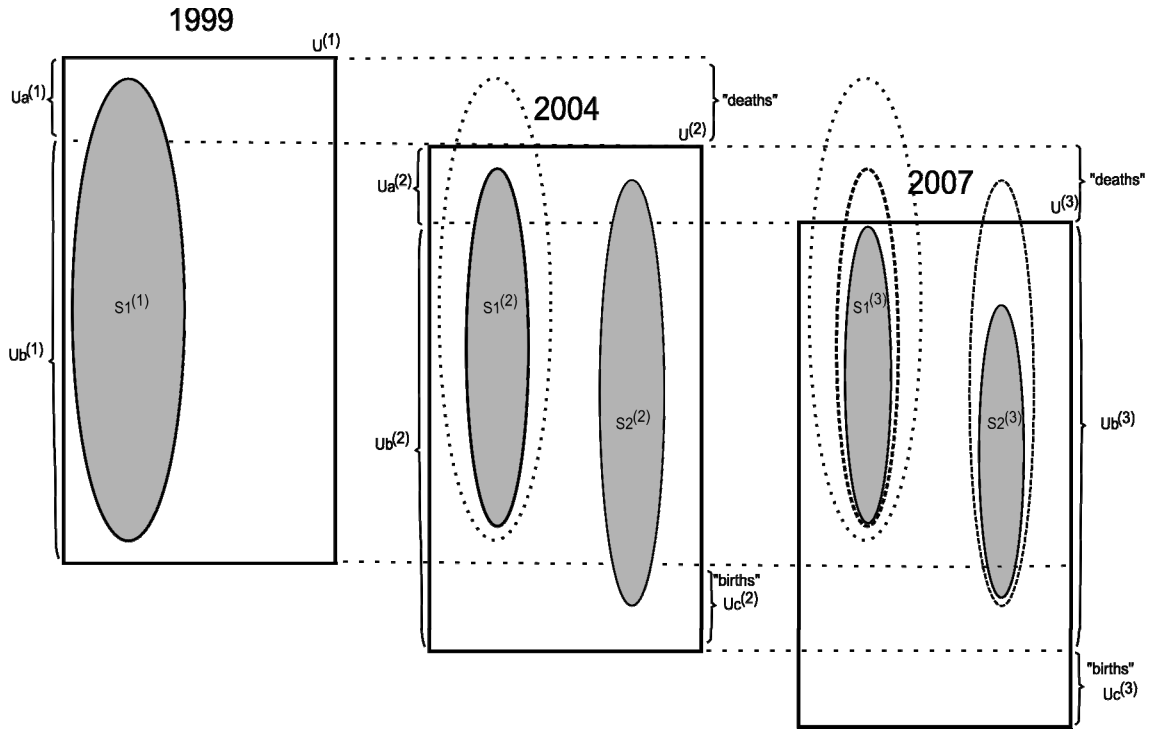
- t year of the current wave: 2007,
- t_{S1} year of selection of SHP_I: 1999,
- t_{S2} year of selection of SHP_II: 2004.

To combine SHP_I cross-sectionally to SHP_II, we first created weights so that each of the two samples was representative of the cross-sectional population at the time t. In fact, we retrieved the adjustment steps for non-response to the grid, from weight sharing for cohabitants and adjustment for non-response to the individual questionnaire which had already been done separately for both of the surveys. For each of the SHP we start with a weight PTI_NRQI (step 5, see Diagram 19). Therefore, we obtain two panels which are representative of the population. After this we combine the two cross-sectional samples.

Please note that the individuals from $U_a^{(1)}$ (see Diagram 37) having left the target population between the selection of SHP_I and the time t_{S2} of the selection of SHP_II, cannot belong to SHP_II. In addition to the individuals from $U_a^{(1)}$ and $U_a^{(2)}$ having left the target population between the selection of SHP_I and the time t do not belong to the combined sample $S_1^{(3)} \cup S_2^{(3)}$.

The differences between the populations $U^{(1)}$, $U^{(2)}$ and $U^{(3)}$ are due to individuals who left the target population between September 1999 and September 2004, as well as between September 2004 and September 2007 (persons deceased, institutionalised, gone abroad etc.)and to new arrivals in the population after September 1999 or 2004 (immigrants, births). Part of the sample from the first panel includes individuals from leaving populations $U_a^{(1)}$ and $U_a^{(2)}$, just as a part of the sample from the second panel includes individuals from the leaving population $U_a^{(2)}$.
The individuals from these populations do not, therefore, belong to the Swiss population at the moment of the current wave we are concerned with in 2007. The individuals from $U_a^{(1)}$ have zero probability of being chosen for the SHP_II in 2004.
On the other hand, a part of the second panel is composed of individuals who belonged to the Swiss population at the moment of selection of this sample but who were not present at the time of the selection of the first panel ($U_c^{(2)}$); these persons had no chance (zero probability) of being chosen at the moment t_{S1}.
In summary, the individuals who left the target populations (the "dead") between the selection of the panels and the time t are removed from the sample. After the weight sharing, the first panel will be representative of the population at time t less the households composed solely of immigrants who joined the population between t_{S1} and t. In the same way, the second panel will be representative of the population at the time t, less the households composed solely of immigrants who joined the population between t_{S2} and t.

Diagram 37 Combination of panels SHP_I and SHP_II.



Legend

$U^{(1)} = U_a^{(1)} \cup U_b^{(1)}$	Longitudinal target population for SHP_I (panel 1)
$U^{(2)} = U_b^{(2)} \cup U_c^{(2)}$	Longitudinal target population for SHP_II (panel 2). Combined target longitudinal population for SHP_I and SHP_II.
$U^{(3)} = U_b^{(3)} \cup U_c^{(3)}$	Cross-sectional target population in 2007 for SHP_I alone (wave 9), for SHP_II alone (wave 4) and for SHP_I and SHP_II together (wave 4).
$S1^{(1)}$	Sample of respondents from SHP_I in 1999, wave 1
$S1^{(2)}$	Sample of respondents from SHP_I in 2004, wave 6
$S1^{(3)}$	Sample of respondents from SHP_I in 2007, wave 9
$S2^{(2)}$	Sample of respondents from SHP_II in 2004, wave 1
$S2^{(3)}$	Sample of respondents from SHP_II in 2004, wave 4

In order to develop a methodology which allows information from cross-sectional samples from both panels to be combined, consideration of both panels at the moment of introduction of the second panel is enough. The population that has to be considered is $U^{(2)}$, the one in effect at the time of the introduction of the second panel. The longitudinal samples to be considered are the original samples chosen at wave 1 of the panels. The sample of the first panel was reduced due to individuals leaving the population and non-respondents. The method presented here was developed by Merkouris (see [14] and [16]). Let *Y* be a parameter of the population. To estimate *Y* one has to use the following combined estimator (which gives a relative importance to each of the panels according to a certain factor *p*) :

$$\hat{Y} = p \cdot \hat{Y}_1 + (1-p) \cdot \hat{Y}_2$$

where $\hat{Y}$ is the cross-sectional estimate for a variable of interest, $\hat{Y}_i$ is the estimate produced for this same variable of interest for the panel *i*, and *p* is the allocation factor of the panels. **The optimal allocation factor of the panels was determined in**

order to ensure minimal variance of $\hat{Y}$**.** The allocation factor of the panels is defined as follows:

$$p = \frac{n_1}{n_1 + n_2(\frac{d_1}{d_2})}$$

where d_i designates the *effect of the sampling plan*[33] for the panel *i* and n_i represent the number of individuals[34] (or households in the following chapter) of the panel *i* responding at time *t* (t being the year of reference in question, in this case 2007).
As the survey plans for SHP_I and SHP_II are identical, we make the hypothesis that the effects of the plan are roughly the same ($d_1/d_2=1$). To calculate the allocation factors, we take the number of individuals (or households) responding.

It should be noted that although allocating the factors in this way reduces the variance, ideally the mean square error (MSE) should be minimised. However as this latter cannot be easily measured in the context of the SHP, we chose to limit ourselves to variance reduction. It should further be mentioned that in our case, as the SHP_I is twice the size of the SHP_II in terms of number of respondents, this method gives greater importance to the older panel (SHP_I). We suspect that the SHP_I is more biased than SHP_II. We would therefore not wish the allocation factors to move too far away from ½, which is something that occurs anyway, due to the differences in number of observations by major region (Table 38).
In this chapter we deal with steps 13 and 14 in Diagram 19.

10.2 Combination of the two panels: step 13

We combined both panels in order to form a new panel which includes both samples. The combination is carried out by adjusting the weights from each unit according to the allocation factor of the panels mentioned above.

Calculation of combined weight:

- We applied *p* to individuals from SHP_I.
- We applied (*I-p*)to individuals from SHP_II.

Table 38 presents the allocation factors calculated for each major region.

[33] The effect of the sampling plan is defined as being the ratio between the overweighted estimate variance (i.e. without marginal calibration in order to respect the control totals for the population, terminology used by StatCan) and variance which would have resulted from a simple random sample of the same size.
[34] The Statistics Canada methodologists suggest combing the two panels before integrating the cohabitants by the weight sharing method, i.e. that the n_i represent the number of longitudinal respondents only. Omitting the cohabitants in the calculation of *p* factors contributes to a reduction in the importance of the older panel and thus minimises the potential impact due to a bias of attrition.
But as the number of cohabitants is quite low in our survey in proportion to that of the longitudinals, we have not made a distinction between the cohabitants or the longitudinals for the calculation of the n_i , which represent, in our case, the sum of both.

The weight after combination is obtained as follows:

$$\text{PTI_COMBPAN_REG} = \begin{cases} \text{PTI_NRQI} * p_h & \text{for individuals from SHP_I} \\ \text{PTI_NRQI} * (1 - p_h) & \text{for individuals from SHP_II} \end{cases}$$

Table 38 Allocation factors by major region. Combined individual cross-sectional weighting of SHP_I and SHP_II.

REGION07 h	*Total number of individuals* n_h	*Number of individuals SHP_I:* n_{h1}	*Number of individuals SHP_II:* n_{h2}	*Allocation factor of the panels* p_h
1	1234	821	413	0.67
2	1719	1135	584	0.66
3	987	692	295	0.70
4	1243	781	462	0.63
5	869	576	293	0.66
6	669	438	231	0.65
7	252	183	69	0.73
TOTAL	6973	4626	2347	

10.3 Calibration: step 14

The combined weight (PTI_COMPAN_REG) represents the population of Switzerland at the current wave. The calibration step adjusts these weights so that the totals of certain control variables are the same as those of the Swiss population in 2007. The totals can be found in Table 40.
Weight before calibration: PTI_COMBPAN_REG.

The control totals are the same for the individual cross-sectional weightings taken separately for each survey:

1. Sex x (14-24 years; 25-34 years; 35-44 years; 45-54 years; 55+ years)
2. Nationality (Swiss, foreign)
3. Region (Lake Geneva region & Central Switzerland & Ticino; Espace Mittelland & North-western Switzerland & Zurich & Eastern Switzerland)
4. Marital status (married; not married)

Final weight: POIDTRIN.

Table 39 gives some numbers which characterise the calibrations carried out.

10.4 Winsorisation of weights after calibration: step 14a

The ratio of maximum weight to minimum weight is 94. We decided to Winsorise the calibrated weights at the 1st and 99th percentiles. Thus the max/min ratio changes to 16. We observed (Table 40) that Winsorisation has the greatest effect on "foreign" margins and the age range 25 to 34 for men and women alike.

Final weight: POIDTRIN_WIN[35]

The descriptive statistics of the Winsorised weights can be found in Appendix H, Table 39 and Table 40

Table 39 Characteristics of marginal calibrations, individual cross-sectional weighting. We have 6973 individuals with a POIDTRIN>0 for SHP_I and SHP_II combined.

	Calibration adjustment factor = POIDTRIN/PTI_COMBPAN_REG	*POIDTRIN*	*POIDTRIN_WIN*
Average	0.85	901	901
Median	0.80	695	705
Standard deviation	0.27	714	639
Min	0.57	98	268
Max	2.14	9221	4161
Range	1.57	9124	3893

Table 40 Estimated numbers of population before and after calibration for SHP_I and SHP_II combined. Combined individual cross-sectional weight.

Variable	Estimate with the weight before calibration (PTI_COMBPAN_REG)	*Population (and estimate with the weight after calibration POIDTRIN) ESPOP 2007*	*Estimate with Winsorised weight POIDTRIN_WIN*	*Relative differences with nbr. ESPOP for POIDTRIN_WIN estimates*
Male 14-24 years	688236	485575	487444	-0.38%
Male 25-34 years	360004	485027	475576	1.95%
Male 35-44 years	638304	610375	604916	0.89%
Male 45-54 years	627357	550107	552521	-0.44%
Male 55 years and over	1096250	944765	946982	-0.23%
Female 14-24 years	563392	461919	465801	-0.84%
Female 25-34 years	384723	490511	480114	2.12%
Female 35-44 years	731286	608983	608278	0.12%
Female 45-54 years	720359	543721	549681	-1.10%
Female 55 years and over	1291893	1099785	1109456	-0.88%
Swiss	6293386	4955440	5022093	-1.35%
Foreign	808417	1325328	1258676	5.03%
Geneva, Central and Ticino	2393124	1999361	1988624	0.54%
Mittelland, N-W, Zurich, Eastern	4708679	4281408	4292145	-0.25%
Unmarried	3204303	2867978	2874550	-0.23%
Married	3897500	3412790	3406218	0.19%
Total	**7101803**	**6280768**	**6280768**	

[35] In the descriptives of weightings given to users, the variable combined final cross-sectional weight, i.e. POIDTRIN, is renamed wp07T1P.

11 Combined cross-sectional household SHP_I and SHP_II

The combined cross-sectional household weighting is done in the same way as the combined individual cross-sectional weighting except that two cross-sectional samples composed of households and not individuals are combined. The panels are, therefore, combined after adjustment for non-response to the household questionnaire. The weight before combination is PTM_NRQM (which was calculated at the time of the household cross-sectional weightings of each of the surveys, step 8, see Diagram18). Lastly, we carry out an integrated calibration by using the totals of the resident population in private households in Switzerland in 2007.
Here we will cover steps 15 and 16 from Diagram 19. For introduction to the theory, please refer to paragraph 10.1 of the previous chapter.

11.1 Combination of the two panels: step 15

We combined both panels in order to form a new panel based on both samples. The combination is carried out by adjusting the weights from each unit according to the allocation factor of the panels. As for the individual cross-sectional weighting, we applied an allocation factor (see paragraph 10.1) except that this time the factor is calculated at household level.

Calculation of combined weight:

- We apply *p* to individuals from SHP_I.
- We applied (*l-p*)to individuals from SHP_II.

Table 41 presents the allocation factors calculated for each major region in Switzerland.

Table 41 Allocation factors by region. Household cross-sectional weighting of SHP_I combined with SHP_II.

REGION07 h	*Total number of individuals* n_h	*Number of individuals SHP_I:* n_{h1}	*Number of individuals SHP_II:* n_{h2}	*Allocation factor of the panels:* p_h
1	773	505	268	0.65
2	1070	693	377	0.65
3	592	418	174	0.71
4	781	490	291	0.63
5	537	345	192	0.64
6	386	249	137	0.65
7	167	117	50	0.70
TOTAL	4306	2817	1489	

The weight after combination is obtained as follows:

$$\text{PTM_COMBPAN_REG} = \begin{cases} \text{PTM_NRQM} * p_h & \text{for households from SHP_I} \\ \text{PTM_NRQM} * (1 - p_h) & \text{for households from SHP_II} \end{cases}$$

11.2 Calibration: step 16

As for the household cross-sectional weighting, we wish the number of persons in households to represent the total of the resident population of Switzerland. We therefore carried out another integrated marginal calibration by using the accounts from the ESPOP (from December 2007).
The combined cross-sectional weight (PTI_COMBPAN_REG) represents the population of Switzerland at the current wave. The adjustment factors can be found in Table 42 and the totals in Table 43.

Weight before calibration: P_ PTM_COMBPAN_REG.

Control totals

1. Sex x (< 14 years; 14-24 years; 25-34 years; 35-44 years; 45-54 years; 55+ years)
2. Nationality (Swiss, foreign)
3. Region (Lake Geneva region & Central Switzerland & Ticino; Espace Mittelland & North-western Switzerland & Zurich & Eastern Switzerland)
4. Marital status (married; Unmarried)

Final weight: POIDTRME

As for the household cross-sectional weighting of the SHP_I (see paragraph 8.3), the calibration of the household cross-sectional weight was carried out twice because after the first calibration there was one household which had a negative weight (with a value of -104). Before doing the second calibration, we put the negative weights back to their values prior to the first calibration (i.e. PTM_COMBPAN_REG). POIDTRME is the final household cross-sectional weight of the combined weighting (thus after the second calibration). Table 42 summarises the facts.

11.3 Winsorisation of weights after calibration: step 16a

The ratio of maximum weight to minimum weight is 136. We decided to Winsorise the calibrated weights at the 1st and 99th percentiles. Thus the max/min ratio changes to 17. We observed (Table 43) that as in the case of the individual cross-sectional weighting, Winsorisation most affects "foreign" margins and the age range 25 to 34 for men and women alike.

Final weight: POIDTRME_WIN[36]

The descriptive statistics of the Winsorised weights can be found in Appendix H, Table 42 and Table 43

[36] In the descriptives of weightings given to users, the variable combined final cross-sectional weight, i.e. POIDTRME, is renamed wh07T1P.

Table 42 Characteristics of marginal calibrations, combined household cross-sectional weighting for SHP_I and SHP_II. We have 10651 individuals in 4306 households with a POIDTRME>0 for SHP-I combined with SHP_II. The characteristics of the adjustment factors are given at the level of individuals, whereas those of the weights after calibration are at household level.

	Calibration adjustment factor (=p before / p after)				
	1st calibration	*2nd calibration*	*After first calibration*	*POIDTRME*	*Final weight POIDTRME_WIN*
	Individual level where household responded to household questionnaire		*Household level respondents to the household questionnaire*		
Average	0.88	1.00	735	735	735
Median	0.80	1.00	606	605	612
Standard deviation	0.40	0.00	479	480	439
Min	**-0.15**	0.96	**-104**	40	168
Max	3.08	1.01	5494	5447	2853
Range	3.23	0.05	5598	5407	2684

Table 43 Estimated numbers of population before and after calibration for SHP_I and SHP_II combined. Combined cross-sectional household weight.

Variable	*Estimate with the weight before calibration* PTM_COMBPAN_REG	*Population ESPOP 2007 and estimation with the weight after calibration: POIDTRME)*	*Estimate with Winsorised weight* POIDTRIN_WIN	*Relative differences with nbr. ESPOP for POIDTRIN_WIN estimates* POIDTRME_WIN
Males less than 14 years	642256	559468	557388	0.37%
Male 14-24 years	647094	485575	486450	-0.18%
Male 25-34 years	349497	485027	478831	1.28%
Male 35-44 years	613532	610375	603908	1.06%
Male 45-54 years	616815	550107	550776	-0.12%
Male 55+ years	1097556	944765	954377	-1.02%
Males less than 14 years	598546	529354	526743	0.49%
Female 14-24 years	557530	461919	464171	-0.49%
Female 25-34 years	365953	490511	481791	1.78%
Female 35-44 years	676899	608983	608120	0.14%
Female 45-54 years	696063	543721	546869	-0.58%
Female 55+ years	1299649	1099785	1110167	-0.94%
Swiss	7267376	5781145	5838753	-1.00%
Foreign	894014	1588445	1530837	3.63%
Geneva, Central, Ticino	2769851	2364422	2364028	0.02%
Mittelland, N-W, Zurich, Eastern	5391539	5005168	5005562	-0.01%
Unmarried	4375570	3956800	3962520	-0.14%
Married	3785820	3412790	3407070	0.17%
Total	**8161390**	**7369590**	**7369590**	

12 Individual longitudinal weighting of SHP_I and SHP_II combined

It is possible to combine two panels which were not selected at the same time providing one has information concerning the eligibility of members of each panel at the time these were selected. It was decided first to combine SHP_I and SHP_II at longitudinal level and then to produce weights from this combined longitudinal panel. We therefore combined the samples after adjustment for non-response to the individual questionnaire for the longitudinals (steps 10 and 12, Diagram 19).

We combined both panels in order to form a new panel which includes both samples. The reference year of the combined panel is that of the most recent panel, i.e. 2004.
Individuals from the first panel, who left the target population at time t_{S2} (see Diagram 37), those belonging to the population $U_a^{(1)}$, are excluded from the panel combining both samples.
The combination is carried out by adjusting the weight for each unit according to the allocation factor of the panels (see paragraph 10.1), but the calculation of this factor is slightly more complex than in the case of the individual and household cross-sectional weightings.
Once the panels have been combined, the step of calibration adjusts these weights so that the totals according to certain control variables are the same as those of the Swiss population in 2004.
In this chapter, we describe steps 12, 17 and 6 of Diagram19.

12.1 Adjustment for non-response to the individual questionnaire for SHP_I, reference year 2004

The adjustment for non-response to the individual questionnaire presented in paragraph 9.1 (step 10), and carried out as part of the longitudinal weighting of the SHP_I alone, must be made again before we can combine the two panels. In fact, the longitudinal weighting of the SHP_I alone refers to the year 1999, whereas the combined longitudinal weighting of SHP_I and SHP_II together refers to year 2004. This makes a revision of the status of persons in the SHP_I necessary. The longst-comb variable is an update of the longst variable (see paragraph 6.3) and is defined as follows :

- for individuals from SHP_II $longst_comb = longst_{SHP_II}$
- for individuals from SHP_I:

$$longst_comb = \begin{cases} longst_{SHP_I} & \text{if } longst_{SHP_I} = 1,3 \text{ or } 4 \\ 1 & \text{if } longst_{SHP_I} = 2 \text{ and age in } 2004 \geq 14 \text{ years} \\ 2 & \text{if } longst_{SHP_I} = 2 \text{ and age in } 2004 < 14 \text{ years} \end{cases}$$

In summary, the individuals who had a child longitudinal status in the SHP_I but who have reached or exceeded the age of 14 in 2004, change status and "become" adult longitudinals for the combined weighting of the panels. It is this which makes it necessary to carry out the adjustment again for the non-response to the individual questionnaire by taking into account these new statuses of longst-comb (for the SHP_I only).
Procedure

We created a file containing adult longitudinal individuals (LONGST_comb=1) who are in households having responded to the grid or who are cross-sectionally out of scope (CODEREP=1,3,4,5,6 or 7), i.e. all those for whom P_NRGRIL > 0.
We defined:

REPQI = 0 Non-respondent to the individual questionnaire (CODEREP=3, 4 or 5)
1 Respondent to the individual questionnaire (CODEREP=6 or 7).

Cross-sectionally out-of-scope individuals (CODEREP=1) do not undergo adjustment for non-response, their adjustment factor is therefore equal to 1 (weight before adjustment = weight after adjustment) and the SAS macro for the segmentation `%aj_nr(…)` attributes them a RHG=. (missing).

Table 44 Distribution of longitudinal adults belonging to the cross-section target population and who belong to households responding to the grid according to the response to the individual questionnaire. Please note that we have 50 cross-sectionally out-of-scope longitudinals for the SHP_I by taking the reference year 2004. These latter keep their P_NRGRIL.

	SHP_I Wave 9
Response to I.Q. (REPQI)	Number of individuals (proportion as %)
0 - non-respondent to I.Q.	1195 (22.4)
1 - respondent to I.Q.	4129 (77.6)
Total	**5324**

Table 45 Descriptive statistics of the adjustment factor, f_aj_nrlqi, for the 4129 individuals responding to the individual questionnaire (those for whom REPQI=1, see Table 44). PL_NRQI = P_NRGRIL * f_aj_nrlqi. The RHG formed where the response rate reached 100% lead to an adjustment factor of 1.00 (i.e. no adjustment).
For the 50 cross-sectionally out-of-scope individuals, we have f_aj_nrlqi=1 (these latter are not included in the statistics in the Table below).

	SHP_I		SHP_I
Average	1.30	Min	1.00
Median	1.20	Max	2.63
Std. deviation	0.28	Range	1.63

The adjustment is made at the individual level. Data stemming from the grid of the **current wave** were used for the modelling.
The model obtained with segmentation analysis included 19 RHG for the SHP_I. A weighted response rate of at least 30% and a minimum number of 50 individuals were required in each RHG (see Appendix G).

The adjusted weight for non-response (individual questionnaire) is PL2_NRQI.

The (4129+50=) 4179 individuals from SHP_I who had a non-zero weight PL2_NRQI are those who will have a final combined longitudinal non-zero weight (F_PLONG=1) among the individuals from SHP_I.

12.2 Combination of the two panels: step 17

Therefore, an allocation factor p, where $p \leq 1$ et $p = \frac{n_1}{n_1 + n_2}$ is applied to the weights of longitudinal individuals. In fact, factor p is the proportion of the number of responding longitudinals from the oldest panel (n_1) to the total number of longitudinals ($n_1 + n_2$).

The factor is applied in the following way:

- p to individuals from SHP_I (oldest panel) who were eligible at the time of selection of the most recent panel (the SHP_II) at the moment t_{S2} (see Diagram 37).
- $1-p$ $1-p$ to individuals from SHP_II (the most recent panel) who belonged to the target population at the moment t_{S1}.
- 1 (i.e. no factor) to individuals from SHP_II who did not belong to the target population at the moment t_{S1}. The individuals belong to the population $U_c^{(2)}$ the "births" between 1999 and 2004 (see Diagram 37).

Thus the new panel obtained is representative of the same population as the SHP_II, i.e. the resident population in private households in Switzerland in 2004. This new panel is at its wave 4 in 2007.
In order to decide how to apply the panel allocation factor (p), it is essential to know the eligibility of the members of each of the panels. In order to do this, we created two new variables:

- PRES_P1: variable indicating whether the individual belonged to the target population at the moment of selection of the oldest panel (SHP_I, in 1999). This has a value of 1 if the individual belonged; 2 if they did not.
- PRES_P2: variable indicating whether the individual belonged to the target population at the moment of selection of the most recent panel (SHP_II, in 2004). This has a value of 1 if the individual belonged; 2 if they did not.

The variables PRES_P1 and PRES_P2 already had to be constructed for the combined longitudinal weighting of SHP_I and SHP_II 2005 (see [3]). In spite of this, a certain number of longitudinal individuals (longst_comb=1) responded to the individual questionnaire in 2007 but had not done so in 2005. There are, therefore, new cases for which PRES_P1 and PRES_P2 values have to be determined. For these remaining cases, we applied the algorithm described in Appendix J.

The samples are combined at the level of the major regions. An allocation factor is therefore calculated for each major region. The weight before combination is PL2_NRQI (step 12) for individuals from SHP_I and PL_NRQI (step 10) for those from SHP_II.
Calculation of combined weight:

- p is applied to individuals from SHP_I who have PRES_P2=1.

- $1 - p$ is applied to individuals from SHP_II who have PRES_P1=1.
- No factor is applied (therefore we apply 1) to individuals from the SHP_II who have PRES_P1=2.

Weight after combination: PLI_COMBPAN_REG

Table 46 presents the allocation factors calculated for each major region.

Table 46 Allocation factors by region. Household cross-sectional weighting of SHP_I combined with SHP_II.

REGION04 h	*Total number of individuals* n_h	*Number of individuals SHP_I:* n_{h1}	*Number of individuals SHP_II:* n_{h2}	*Allocation factor of the panels:* p_h
1	1129	746	383	0.66
2	1580	1035	545	0.66
3	895	629	266	0.70
4	1131	687	444	0.61
5	799	529	270	0.66
6	596	384	212	0.64
7	234	169	65	0.72
TOTAL	**6364**	**4179**	**2185**	

12.3 Calibration: step 18

The adjusted longitudinal weight for non-response to the individual questionnaire (PLI_COMBPAN_REG) represents the population of Switzerland at the moment of selection of the sample of SHP_II at wave 1. The calibration step adjusts these weights so that the totals according to certain control variables are the same as those of the Swiss population in 2004 according to the ESPOP. The totals can be found in Table 48.

Weight before calibration: PLI_COMBPAN_REG

Control totals

1. Sex* (14-24 years; 25-34 years; 35-44 years; 45-54 years; 55+ years)
2. Nationality (Swiss, foreign)
3. Region (Lake Geneva region & Central Switzerland & Ticino; Espace Mittelland & North-western Switzerland & Zurich & Eastern Switzerland)
4. Marital status (married; Unmarried)

Final weight: PLI_COMBREG_CAL[37]

Table 47 gives some characteristic numbers of the calibrations carried out.

[37] In the descriptives of weightings given to users, the final individual longitudinal weight variable, i.e. PLI_COMBREG_CAL, is renamed WP07L1P.

Table 47 Characteristics of the marginal calibration (including cross-sectionally out-of-scope individuals), combined individual longitudinal weighting of SHP_I and SHP_II. We have 6364 observations with a PLI_COMBREG_CAL>0.

	Calibration adjustment factor = PLI_COMBREG_CAL/PLI_COMBPAN_REG	PLI_COMBREG_CAL
Average	0.87	956
Median	0.82	769
Std. deviation	0.25	676
Min	0.55	213
Max	1.94	8260
Range	1.39	8046

Table 48 Estimated size of population before and after calibration. Combined individual longitudinal weight of SHP_I and SHP_II. Please note that these totals reflect the population aged 14 and over.

	Estimate with the weight before calibration (PLI_COMBPAN_REG)	Population ESPOP 2004 (and estimate with weight after calibration PLI_COMBREG_CAL)
Male 14-24 years	658891	469973
Male 25-34 years	399272	481790
Male 35-44 years	684968	611688
Male 45-54 years	580162	513613
Male 55 years and over	944735	888469
Female 14-24 years	498202	449624
Female 25-34 years	431335	491394
Female 35-44 years	788455	611586
Female 45-54 years	678971	509749
Female 55 years and over	1084101	1052980
Swiss	5960561	4841446
Foreign	788530	1239420
Geneva, Central and Ticino	2272801	1924056
Mittelland, N-W, Zurich, Eastern	4476290	4156810
Not married	2947102	2686477
Married	3801989	3394388
Total	**6749091**	**6080866**

13 Conclusion

The weightings described in this document have been transposed on to those of the previous year, which were, in turn, more or less replicas of those produced by the methodologists from Statistics Canada (see [3], [11]) for the data from the Swiss Household Panel (SHP) for the first 7 years. We have gone to great lengths to enhance the report with examples and explanations concerning the many demanding theoretical aspects used for the various steps of the production of weights.

This year, as for the previous one, various measures were taken to combat attrition (for particularities of 2007, see paragraph 3.1.3) and also to try and take these measures into consideration in the non-response models for the production of weightings. Attrition is a problem that grows as the SHP ages. In fact, the models to explain non-response have their limits and the marginal calibrations become more difficult. Thus we are often forced to Winsorise the weights in order to limit their variability. Nevertheless, as shown in figures 2 and 3, in 2007-08, the effect of attrition was offset by the great efforts made, and for the SHP_I we finished with more individual and household interviews carried out than for the previous year. Furthermore the SHP_II continued to deteriorate, but less than for its first waves.

The data from the files supplied are not weighted. It is the responsibility of data users to apply the appropriate weights depending on the analyses they wish to produce. Strong biases are a potential risk if weights are not applied.

13.1 Guidelines for the application of the weightings produced

Generally, the smaller the sample, the greater the variability of the sampling. Consequently, the estimates regarding the small sub-groups of the population tend to be less reliable than those made with regard to large (sub-)groups.
Bearing this in mind, there are three first steps to be considered:

(1) When using cross-sectional data (at individual or household level), one should always use the maximum amount of information available:
   - if a variable is available for both surveys (SHP_I and SHP_II), both samples should be utilised and the combined weightings used. Using information from either only SHP_I or SHP_II would result in an avoidable loss of precision in the results produced.

(2) When using longitudinal data (at individual level), one should always use the maximum amount of information available:
   - if an operation relies only on data collected since 2004, weightings should be used that allow longitudinal combination of both SHPs. It should be noted that in this case the population studied is the one residing in private households in Switzerland in 2004 (even for the data from SHP_I).
   - If the utilisation relies (even only partially) on data collected before 2004, then only the SHP_I can be used and only the longitudinal weight of the SHP_I should be applied.

(3) If an estimate relies on less than 30 observations (non-weighted number of observations), these should only be published with reservation, regardless as to

the value of its variation coefficient. In fact, the smaller the group of observations studied, the higher is the risk of bias.

The concepts of variance, coefficient of variation and confidence interval provide an indication of the degree of sampling variability. However, we remind users that the estimates of these variables, provided by the classic procedures of most of the current statistical software (SAS, SPSS, Stata, MS Excel ...), are subject to error when the analysed data come from complex surveys such as the SHP. This software presumes, amongst other things, that a simple random sampling has been carried out. That is to say that they do not take into account the special characteristics of the survey design of the SHP such as stratification, nor the corrections made for non-response, nor the post-stratification (marginal calibration). Even if the classical procedures from this software enable one to resort to weighting for the production of estimates, they do not properly take into account all the specifications of the SHP survey for the calculation of estimates of variance, an essential element for the majority of statistical tests. In fact, there is a tendency to underestimate the above-mentioned variance and consequently the confidence intervals linked to these estimations.

13.2 Recommendations on estimations of precision

Despite the fact that the variance is often underestimated when using the software mentioned in the preceding paragraph, we nevertheless recommend its calculation. To do this, however, use of weighting is essential. As a matter of fact, confidence intervals of estimations produced for example, will have a tendency to be too narrow. But if a difference or a variable already proves to be non-significant with such a calculation, it will not be necessary to investigate further.
Recent work has shown (see [23]) that it is possible, in particular with the SAS and SPSS software to use procedures which allow all specifications of the survey design to be taken into consideration. This concerns the SAS procedures SURVEY and those of the COMPLEX SAMPLE module from SPSS. They calculate the variance according to the Taylor series linearization method, and give very similar results (at least for the SHP) to those produced by re-sampling methods which are more difficult to implement, such as the bootstrap or the jack-knife. For this reason, for SAS or SPSS users, as far as is possible, we recommend the use of these procedures which will certainly provide much more accurate results than those provided by standard procedures.

14 Bibliography

[1] Graf, E (2008). *Pondérations du SILC pilote: SILC_I vague 2, SILC_II vague 1, SILC_I et SILC_II combinés*. Federal Statistical Office, N°338-0051.

[2] Cornali, A. et Vonlanthen, C. (2000). *Description de la pondération du Panel Suisse de Ménages*. Federal Statistical Office: internal document.

[3] Cauchon, C. et Latouche, M. (2007). *Pondératon du Panel Suisse de Ménages : PSM_I vague 7, PSM_II vague 2, PSM_I et PSM_II combinés ; description détaillée des tâches.* Swiss Household Panel.

[4] Cauchon, C. et Latouche, M. (2006). *Pondération du SILC, SILC vague 1, PSM II et SILC combinés, PSM I, PSM II et SILC combinés; description détaillée des tâches*. Swiss Household Panel.

[5] Graf, E. et Tillmann, R. (2005). *Comparaison du déroulement des enquêtes PSMI, PSMII et SILC (2004-2005)*. Working Paper 4_05, Swiss Household Panel.

[6] Graf, E. et Tillmann, R. (2006). *Comparaison du déroulement des enquêtes PSM et SILC (2005-2006) : l'exemple de la gestion des interviews.* Swiss Statistical Society bulletin, 53 (avril 2006), 9-13.

[7] J. Dufour, F. Gagnon, Y. Morin, M. Renaud, C.-E. Särndal (1998). *Measuring the impact of alternative weighting schemes for longitudinal data.* SAA proceedings, Survey Research Methods Section, pp. 552-557.

[8] Kass, G.V. (1980). *An exploratory technique for investigating large quantities of categorical data*. Applied Statistics, 29, 119-127.

[9] Lavallée, P. (2002). Le sondage indirect ou la méthode généralisée du partage des poids.

[10] Latouche, M. (2007). *Calage des poids et estimations de la variance pour le Panel Suisse de ménages.* Swiss Household Panel: internal document.

[11] Cauchon, C. et Latouche, M. (2007). *Pondération du Panel Suisse de Ménages, PSM_I vague 7, PSM II vague 2, PSM_I et PSM_II combinés; description détaillée des tâches*. Swiss Household Panel.

[12] Latouche, M. et Naud, J.-F. (2005). Panel suisse de ménages, stratégie de pondération en présence de plusieurs panels. Document interne du SHP.

[13] Naud, J.-F. (2005). Deux façons de combiner des panels. SHP internal document.

[14] Merkouris, T. (2001). *Estimation transversale dans le cas des enquêtes auprès des ménages à panels multiples.* Techniques d'enquête, vol. 27, pp. 189-200.

[15] Ardilly, P. et Lavallée, P. (2003). *La méthode du partage des poids et l'enquête européenne sur le revenu et les conditions de vie.* Recueil du Symposium de

Statistique Canada sur les Défis à la réalisation d'enquêtes pour la prochaine décennie.

[16] Lévesque, I. et Franklin, S. (2000). *Pondérations longitudinale et transversale de l'Enquête sur la Dynamique du travail et du revenu, Année de référence 1997*. Statistics Canada, 75F0002MIF-00004.

[17] Naud, J.-F. (2004). *Pondération longitudinale avec panels combinés, Enquête sur la dynamique du travail et du revenu, 1996-2002*. Statistics Canada, 75F0002MIF2004008.

[18] Latouche, M., Dufour, J. et Merkouris, T. (2000). *Pondération transversale : combinaison de deux panels ou plus*. Statistics Canada, 75F0002MIF-00006.

[19] Deville, J.-C. et Särndal, C. E. (1992). *Calibration Estimators in Survey Sampling.* Journal of the American Statistical Association, Vol87, No. 418 (June 1992), PP. 376-382.

[20] Tambay, J. L., Schiopu-Kratina, I., Mayda, J., Stukel, D. et Nadon, S. (1998). *Traitement de la non-réponse du cycle de deux enquêtes sur la santé de la population.* Techniques d'enquête, vol. 24, no 2, pp.159-169.

[21] Kalton, G. et Kasprzyk, D. (1986). *Le traitement des données d'enquête manquantes.* Techniques d'enquête, 12, pp. 1-17.

[22] Nakache, J.-P. et Confais, J. (2003). *Statistique explicative appliquée : analyse discriminante, modèle logistique, segmentation par arbre.* Editions Technip.

[23] Plaza, S. et Graf, E. (2007). *Recommandations et exemples pratiques concernant l'application des pondérations*. Swiss Household Panel.

[24] Graf, E. (2008). Pondérations du PSM: PSM_I vague 8, PSM_II vague 3, PSM_I et PSM_II combinés. Order number: 338-0054

[25] M.I.S. Trend SA (2008), Panel Suisse de Ménages 2007-2008, rapport technique.

APPENDIX A – Data files for SHP_I and SHP_II

Here is a concise description of the data files that were used to carry out the weightings for SHP_I and SHP_II. For more details, please refer to the internal working papers produced by the author.

MENAGES 2007 this file contains households that replied to the grid at wave 9 of SHP_I and at wave 4 of SHP_II.
The file contains 4436 registrations: 2893 for SHP_I; 1543 for SHP_II and 52 variables. Name: `MEN07.sas7bdat`, made by programme `EGR_SHP_men07_Ind07_20081113.sas`

Table A-1 List of variables from file MEN07:

IDHOUS07	household identifier
FILTER07	1= SHP_I, 0= SHP_II
HLINGU07	household language, 1=French, 2=German,3=Italian
REGION07	household region (7 regions)
COM2_07	Topology of municipality 2: 1999 for SHP_I households, 2004 for those from SILC and SHP_II (9 typologies), 1=centres, 2=suburban municipalities, 3=high income municipalities, 4=semi-urban municipalities, 5=tourist municipalities, 6=industrial and tertiary municipalities, 7=rural commuter municipalities, 8=mixed rural municipalities, 9=agricultural municipalities
NBPERS07	number of individuals in the household
NBADUL07	number of adults (18 years and +)
NBKID07	number of children (-18 years)
H07H15	original variable, type of building, 1= building with several apartments, 2=single family house, 3=other type of house/apartment.
H07H29	1=tenant, 2=owner, 3=rent-free occupant
H07H37	accommodation costs, in francs
H07H20	number of rooms in dwelling
FRANC	household questionnaire completed in French (1) or in another language (0)
ALLEM	household questionnaire completed in German (1) or in another language (0)
ITAL	household questionnaire completed in Italian (1) or in another language (0)
LAKE_GE	household region = Lake Geneva (1) or another (0)
MIDDLE	household region = Mittelland (1) or another (0)
NOREST	household region = North-east (1) or another (0)
ZURICH	household region = Zurich (1) or another (0)
EST	household region = East. Switzerland (1) or another (0)
CENTRAL	household region = Centr. Switzerland (1) or another (0)
TESSIN	household region = Ticino (1) or another (0)
CENTRES	municipality type: Agglomerations (1) or other (0)
SUBURB	municipality type: suburbs (1) or other (0)
WEALTH	municipality type: rich municipality (1) or other (0)
PERIPH	municipality type: peripheral municipality (1) or other (0)
TOURIST	municipality type: tourist municipality (1) or other (0)
INDUS	municipality type: Industrial or tertiary municipality (1) or other (0)
RURAL	municipality type: commuter towns(1) or other (0)
MIXED	municipality type: mixed rural municipality (1) or other (0)
SEULE	single-person households (1) or multi-person (0)

AD_SEUL	household with a single adult (with or without children) (1) vs. household with more than 1 adults (0)
NO_KID	household without children (1) or with child(ren) (0)
APPART	household living in an apartment in a building (1) or other (0)
DETACH	household living in a single family house (1) or other (0)
PROPRIO	owner (1) or other (0)
RAT	ratio, number of rooms/no. of persons
RAT1	1=first lowest third ratio no. rooms/no. of persons (<1.34) otherwise 0
RAT2	1=second lowest third ratio no. rooms/no. of persons (>1.34, <=2), otherwise 0
RAT3	1=third lowest third ratio no. rooms/no. of persons (<2) otherwise 0
LOY_1	first lowest third for accommodation costs (>=0, <=1000 CHF/month) (1) vs. other (0)
LOY_2	second lowest third for accommodation costs (>=0, <=1000, <=1500 CHF/month) (1) vs. other (0)
LOY_3	highest third for accommodation costs (>1500 CHF/month and +) (1) vs. other (0)
Q_INDIV	0=no individual questionnaire completed in the household, 1=at least one individual questionnaire completed in the household
Nbr_Qind_complete	number of individual questionnaires completed in the household
STATHH06	1=household questionnaire completed, 0=household questionnaire not completed or only partially completed
MAISINDEP	1= household lives in single family house, otherwise 0
MEN_RR	1= household has been through refusal call, otherwise 0
NOTEASYADR	1= SHP_I (or SHP_II) household with only the grid in W8 (or W3) and at least household questionnaire or 1st individual questionnaire in W7 (or W2) or total non-respondent household in W8 (or W3) and at least the household questionnaire or 1st individual questionnaire in W7 (or W2) or total non-respondent household in W8 (or W3) and only the grid in W7 (or W2)
P1ADRFROMW2_3_4	1=SHP_I household complete for the last time in W2, 2* and W4, otherwise 0
P1EASYADR	1= household complete in W8 or only the household questionnaire or only 1 individual questionnaire and the household belongs to SHP_I, otherwise 0
P2EASYADR	1= household complete in W3 or only the household questionnaire or only 1 individual questionnaire and the household belongs to SHP_II, otherwise 0

INDIVIDUS 2007 this file contains all individuals from households that replied to the grid at wave 9 of SHP_I and at wave 4 of SHP_II. They all belong to the cross-sectional target population. They can be respondents or non-respondents to the individual questionnaire.

Table A-2 List of variables from file PERS07:

IDPERS	Individual identifier
IDHOUS07	Household identifier
STATUS07	Mode of data collection 0=respondent individual questionnaire, 1=proxy, thus not eligible, 2=non-respondent to individual questionnaire
SEX	Sex
AGE07	Age in years at the end of 2007
BIRTHY	Year of birth

CIVSTA07	Marital status in 2007 1=single, 2=married, 3= separated, 4=divorced, 5=widow/er
EDUCAT07	Highest level of education reached (information from individual and grid questionnaire) 0=compulsory school not finished 1=compulsory school, elementary professional training 2=household training, 1 year of commercial school, short 3= general education school 4 = apprenticeship (CFC) 5 = full-time vocational school 6 = baccalaureate 7 = advanced vocational training with federal certificate of proficiency 8 = technical or vocational school 9 = higher professional school, ETS etc. 10 = university, HEP, HES
OCCUPA07	Current occupation
NAT_1_07	First nationality
NAT_2_07	Second nationality
NAT_3_07	Third nationality
HAB_CH07	Length of stay in Switzerland in years (for non-Swiss)
HOMME	male (1) or female (0)
AGE_30	age under or 30 (1) or over 30 (0)
CELIB	Single (1) or other (0)
MARIE	married (1) or other (0)
SEPARE	separated (1) or other (0)
DIVORCE	divorced (1) or other (0)
VEUF	widowed (1) or other (0)
NO_FORM	no special training (1) or other (0)
FORM_PRO	professional training (1) or other (0)
FORM_SUP	higher education or training (1) or other (0)
TRAVAIL	economically active persons (1) or other (0)
SWISS	Swiss nationality (1) or other (0)
AGE_1	age group 15 to 24 (1) otherwise (0)
AGE_2	age group 25 to 34 (1) otherwise (0)
AGE_3	age group 35 to 44 (1) otherwise (0)
AGE_4	age group 45 to 54 (1) otherwise (0)
AGE_5	age group 55 and over (1) otherwise (0)
FILTER07	**1**= SHP_I, **0**= SHP_II
PERMIC	Residence permit C (1) or not (0)
PERMIB	Residence permit B (1) or not (0)

10993 registrations, 7225 for SHP_I, 3768 for SHP_II, 33 variables. Name: `PERS07.sas7bdat`, made by programme `EGR_SHP_men07_Ind07_20081113.sas`

PARTIS 2007

This file contains the list of individuals who, at the time of the wave being weighted, no longer lived in the household in which they had been counted during the **previous wave**. The OSM variable indicates whether the person having left the household is an OSM[38] or not, according to the following coding:

1 = OSM
2 = child of OSM
3 = Non-OSM

Table A-3 List of variables from file PARTIS07:

FILTER07	**1**= SHP_I, **0**= SHP_II
IDPERS	Identification of person
IDHOUS07	Identification of household
IDSOUR07	Number of original household (for split-offs, it is different from IDHOUS06)
OSM	Individual status (1=OSM, 2=child of OSM, 3=Non-OSM
G07RNH	Reason why no longer living in household 1=deceased 2=hospitalisation, institutionalisation 3=other type of institutionalisation 4=divorce, separation, break-up (including child having followed one of his parents) 5=child who has left parental home "for good" 6=child who has left parental home "temporarily" 7=left to go abroad (professional or other) 8=other reason
SEX	Sex, information on master file 2007
AGE07	Age in 2007, information on master file 2007
G07CIV	Marital status in 2007 1 = single, 2 = married, 3 = separated, 4 = divorced, 5 = widowed
G07EDU	Highest degree of education completed, internal variable 0=NO education or training completed 1=compulsory school 2=basic professional training and education 3=apprenticeship (CFC level) 4=full-time professional school/profession 5=vocational baccalaureate 6=diploma degree or general culture school (admin. school) 7=gymnasium/college, teacher training school, baccalaureate 8=commercial school/general training/domestic training (au pair) 12= higher professional training with federal certificate 13=professional or technical school (2 years full-time/3 years part-time) 14=higher professional/technical school/ETS (3 years full-time/4 years part-time) 15=university (bachelor's degree., PhD) 30=not yet started school 31-non-compulsory pre-school 32=special school/school for disabled persons
G07OCC	Current occupation 1=full-time paid work (min 37 hours per week) 2=part-time paid work (5 to 36 hours per week) 2=part-time paid work (1 to 4 hours per week) 4=in schooling, apprenticeship 5=work in family business

[38] OSM = Original Sample Member, i.e. a longitudinal individual.

	6=work in sheltered workshop (for disabled persons) 7=child, woman, man at home (max. age 62/65 years) 8=retired (elderly persons) 9=other pensioned persons (disability etc.) 10=unemployed 11=other situation, unpaid leave
G07NAA	First nationality
G07NAB	Second nationality
G07NAC	Third nationality
HOMME	male (1) or female (0)
CELIB	single (1) or other (0)
MARIE	married (1) or other (0)
SEPARE	separated (1) or other (0)
DIVORCE	divorced (1) or other (0)
NO_FORM	no special training (1) or other (0) G07EDU= **0**,**1**,**2**,**8**,**30**,**31**,**32**
FORM_PRO	professional training (1) or other (0) G07EDU= **3**,**4**,**6**
FORM_SUP	higher education or training (1) or other (0) G07EDU= **5**,**7**,**12**,**13**,**14**,**15**
TRAVAIL	economically active persons (1) or other (0) G07OCC= **1**,**2**,**4**
SWISS	Swiss nationality (1) or other (0)
AGE_1	age group 15 to 24 (1) otherwise (0)
AGE_2	age group 25 to 34 (1) otherwise (0)
AGE_3	age group 35 to 44 (1) otherwise (0)
AGE_4	age group 45 to 54 (1) otherwise (0)
AGE_5	age group 55 and over (1) otherwise (0)

286 registrations in total, 181 for SHP_I, 105 for SHP_II, 29 variables. Name: `partis07.sas7bdat`, made by the programme *`EGR_partis07.sas`*

Table A – 4 Repartition de la variable OSM :

SHP_I wave 9	SHP_II wave 4
127OSM	80 OSM
6 = child of OSM	6 = child of OSM
48 non-OSM	19 non-OSM

REVENUS 2007

This file contains the list of individuals who left the household during the previous waves and who came back for the wave being weighted. **Caution:** only households where the other members continued to respond are taken into account! This does not, therefore, refer to households having completely disappeared one year and who reappear afterwards. The file contains the same variables as PARTIS07 (but with the variable G07RET[39] instead of G07RNH).
25 registrations in total, 19 for SHP_I, 6 for SHP_II, 29 variables.

[39] The variable G07RET indicates the return of former household members (1=yes, 2=no), it is made from the variables G07RET01to G07RET16 investigated for each person). These variables are only investigated from wave 4 or sometimes in wave 3 in special cases of household split-off.

Name: `revenus07.sas7bdat`, made by the programme `EGR_revenus07.sas`

LONGOUT07

This file is made from the files PARTIS00, PARTIS01, PARTIS02, PARTIS03, PARTIS04, PARTIS05, PARTIS06, PARTIS07 and the individual files (the most recent ones) from each wave. It contains the list of longitudinal individuals (OSM) who left their household between the wave 1 and the wave being weighted and who are not counted in any other household, thus, never returned (absent individuals in the individual file of the wave being weighted (= wave 9 for SHP_I, wave 4 for SHP_II), with the reason for departure.
This file only contains, therefore, longitudinals who are out-of-scope for the current year, as some of the individuals may live in Switzerland in a household which could not be contacted.

Table A – 5 List of variables from file LONGOUT07:

IDPERS	Identification number of the individual
RNH	Reason for departure *Meaning of codes, see file PARTIS07)*
WAVE	Wave during which the departure was recorded *Waves 2 to 9 (for PSM_II we take the number of the wave that the PSM-I is at)*
FILTER07	**1**= SHP_I, **0**= SHP_II
OSM	**1** = longitudinal adult, **2** = longitudinal child

1134 registrations in total, 873 for SHP_I, 261 for SHP_II, 5 variables.
Name: `longout07.sas7bdat`, made by the programme `EGR_longout07.sas`

LISTPERS07

For all households surveyed in 2007, contains the list of all persons constituting these households with their personal identification number (`IDPERS`) at the previous wave.
4441 registrations, 2893 for SHP_I, 1548 for SHP_II, 19 variables.
Name: `listpers07.sas7bdat`, made by the programme `EGR_listpers07.sas`

OSM07

Contains all persons from each wave with an indication of their OSM status. This file must be updated annually.
30,0288 registrations in total; two variables: IDPERS and OSM.
Name: *`OSM_2008_10_20.sas7bdat`, made by the programme `EGR_osm07.sas`*

APPENDIX B – Derivation of variables for the classification of individuals

The variables produced during step 2 (classification of individuals) were derived in the following way:

LONGST =
- **1** **Longitudinal - adult**
  (if in INDIVIDUS from wave 1, AGECORRO >=14).
- **2** **Longitudinal - child**
  (if in INDIVIDUS from wave 1, AGECORRO < 14).
- **3** **Initially present cohabitant**
  (if not in INDIVIDUS from wave 1 ET)
  (HAB_CH07 < 1999 (SHP_I)
  < 2004 (SHP_II)
  AND AGECORRO>=0)).
  For the HAB_CH07 < 0 (do not know, no response), we will assume that they were initially present.
- **4** **Initially absent cohabitant**
  (if not in INDIVIDUS from wave 1 ET)
  (HAB_CH07 >= 1999 (SHP_I)
  >= 2004 (SHP_II)
  OR AGECORRO<0)).

Note: longitudinal children (< 14 years) have no longitudinal weight. However, children who were 14 years or over at wave 9 of SHP-I or at wave 4 of SHP_II, have been dealt with as if they were longitudinals at the time of the cross-sectional weighting and they have a cross-sectional weight, even if the child is the only longitudinal member of the household.

CODEREP =
- **0** **Children,** AGECORRC< 14
- **1** **1 – Cross-sectionally out-of-scope**
  (LONGST=1 or 2 and in LONGOUT and RNH=1, 2, 3 or 7)
- **2** **2 – Total non-respondent**
  (AGECORRC>= 14 and not in INDIVIDUS 2007 and not in LONGOUT and with RNH=1, 2, 3 or 7)
- **3** **Grid respondent only**
  (AGECORRC >= 14 and in INDIVIDUS 2007 and STATUS06=1 or 2 and STATHH07 = 0)
- **4** **4 – Only grid and household questionnaire respondent, in a household where all individuals are non-respondent to individual questionnaire**
  (AGECORRC >= 14 and in INDIVIDUS 2007 and STATUS06=1 or 2 and STATHH07 = 1 and Q_INDIV=0)
- **5** **5 – Only grid and household questionnaire respondent, in a household where at least one individual is respondent to the individual questionnaire**
  (AGECORRC >= 14 and in INDIVIDUS 2007 and STATUS07=1 or 2 and STATHH07 = 1 and Q_INDIV=1)
- **6** **6 – Grid respondent, household and individual questionnaire**
  (AGECORRC >= 14 and in INDIVIDUS 2007 et STATUS07=0 et STATHH07=1)
- **7** **7 – Grid respondent and individual questionnaire only**
  (AGECORRC >= 14 and in INDIVIDUS 2007 et STATUS07=0 et STATHH07=1)

Note that all children (i.e. < 14 years at the current wave) are cross-sectionally out-of-scope.

Weight indicators

These variables indicate, for each individual, if he or she has a positive or zero weight.

Longitudinal weight (individual):
F_PLONG = 1 Positive longitudinal weight
(LONGST=1, CODEREP=1, 6 or 7)
0 Zero longitudinal weight (others)

Household cross-sectional weight:
F_PTRME = 1 Positive household cross-sectional weight:
(CODEREP=3, 5 or 6)
0 Zero cross-sectional household weight (others)

Individual cross-sectional weight:
F_PTRIN = 1 Positive individual cross-sectional weight:
(CODEREP=6 or 7).
0 Zero cross-sectional individual weight (others)

*Construction of age variables for **SHP_I wave 9***

IF age99 >= 0 THEN agecorro = age99;
ELSE IF age07 >= 0 THEN agecorro = age07-8;
ELSE IF year1 >= 1800 THEN agecorro = 1999-year1;
ELSE IF year9 >= 1800 THEN agecorro = 1999-year9;
→ Babies born after the first wave will have a negative AGECORR0.
Year 1, year 9: year of birth communicated in wave 1 or 9.

Imputation of missing ages
Ages imputed this year and also the 2 previous years:
IF idpers=2402101 THEN agecorro = 45;
IF idpers=2782101 THEN agecorro = 27;
IF idpers=9090101 THEN agecorro = 70;
Ages imputed this year but not the previous years: we took the age of the (presumed) partner
IF idpers=6233105 THEN agecorro = 42;

agecorrc= agecorro + 8

*Construction of age variables for **SHP_II wave 4***

IF age07 >= 0 THEN agecorro = age07-3;
ELSE IF age04 >= 0 THEN agecorro = age04;
ELSE IF annee4 >= 1800 THEN agecorro = 2004-annee4;
ELSE IF annee1 >= 1800 THEN agecorro = 2004-annee1;
→ Babies born after the first wave will have a negative AGECORR0.
Year 1, year 4: year of birth communicated in wave 1 or 4.

Imputation of missing ages

if IDPERS=20046101 then agecorro=57;
if IDPERS=20046102 then agecorro=57;
if IDPERS=20873102 then agecorro=19;
if IDPERS=20873103 then agecorro=21;
if IDPERS=21103102 then agecorro=59;
if IDPERS=21811101 then agecorro=27;
if IDPERS=21811102 then agecorro=27;
if IDPERS=22647102 then agecorro=56;
if IDPERS=23562102 then agecorro=44;
if IDPERS=24640102 then agecorro=68;
if IDPERS=20046103 then agecorro=39;
if IDPERS=22604104 then agecorro=22;
if IDPERS=22864111 then agecorro=38;
if IDPERS=23482102 then agecorro=50;
if IDPERS=23784101 then agecorro=44;

agecorrc= agecorro + 3

Note on the imputation of missing ages: an age is imputed only for adult individuals having responded to the household or individual questionnaire. The others are attributed an age of 9999 years so that they will be dealt with as adults while remaining easily identifiable.
The imputations of missing ages are treated case by case, manually, taking into consideration the social situation and the household of the person concerned.

APPENDIX C – Modelling of non-response to the grid questionnaire

Modelling of non-response – variables used

The variables used for modelling are dichotomous variables which describe individuals and households and which stem from data (grid questionnaires) from wave 1 (1999 for SHP_I, 2004 for SHP_II). Their description and meaning can be found in APPENDIX A.
Here is a list of variables which could potentially be part of the non-response model (in **bold**, those which have been selected for the model):

Household level

Variable (value in wave 1)	Used for SHP_I in 06	**Used for SHP_I in 07**	Used for SHP_II in 06	**Used for SHP_II in 07**
FRANC	NO	NO (see Diagram C-1)	NO	NO (see Diagram C-1)
ITAL	YES	**YES**	YES	YES
GERMANY	**YES**	**YES**	**YES**	YES
LAKE_GE	NO	NO (see Diagram C-1)	NO	NO (see Diagram C-1)
MIDDLE	**YES**	**YES**	**YES**	**YES**
NORTH EAST	**YES**	**YES**	**YES**	YES
EAST	**YES**	**YES**	YES	YES
CENTRAL	**YES**	YES	**YES**	YES
ZURICH	**YES**	YES	YES	YES
MIXED	YES	**YES**	**YES**	YES
CENTRES	**YES**	**YES**	**YES**	YES
SUBURB	**YES**	**YES**	YES	YES
WEALTH	YES	YES	YES	YES
PERIPH	YES	**YES**	YES	YES
TOURIST	YES	YES	YES	YES
INDUS	**YES**	YES	YES	YES
RURAL	YES	**YES**	**YES**	YES
ADULT ALONE	NO	NO (see Diagram C-1)	NO	NO (see Diagram C-1)
NO_KID	**YES**	**YES**	**YES**	YES
ALONE	YES	**YES**	**YES**	YES
LOY_1	**YES**	**YES**	YES	YES
LOY_2	**YES**	**YES**	**YES**	**YES**
LOY_3	YES	**YES**	**YES**	**YES**
RAT1	YES	NO (see Diagram C-1)	YES	NO (see Diagram C-1)
RAT2	**YES**	**YES**	**YES**	YES
APPART	**YES**	NO (see Diagram C-1)	**YES**	NO (see Diagram C-1)
DETACH	NO	NO (see Diagram C-1)	NO	NO (see Diagram C-1)
OWNER	**YES**	**YES**	**YES**	YES
REV_1[40]	**YES**	**YES**	NO	NO
REV_2	**YES**	**YES**	NO	NO
REV_3	**YES**	**YES**	NO	NO

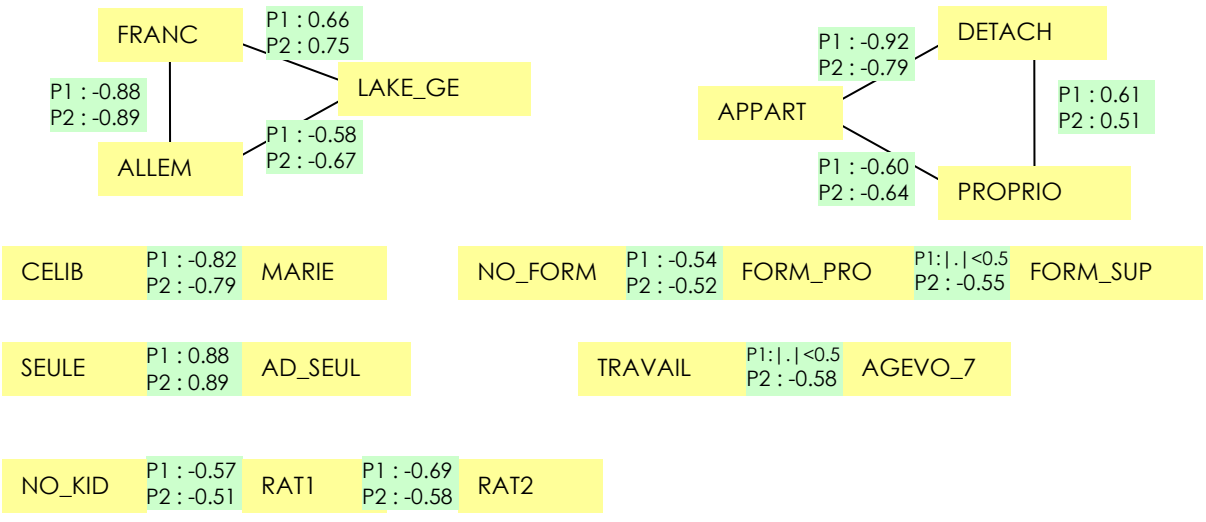
[40] REV_1 to REV_3 classifies the income of households responding to wave 1 of 1999 into three groups (first, second and third). These variables are only available for wave 1 of PSM_I.

Individual level

We introduced a new definition of age groups in 2007, by separating the 15 to 19 year olds from the 20 to 24 year olds. In fact, we observed that these two sub-groups of the population behaved in opposite ways as regards participation rates. We also added an age group for the 55 to 64 year olds.

Variable (value in wave 1)	Used for SHP_I in 06	**Used for SHP_I in 07**	Used for SHP_II in 06	**Used for SHP_II in 07**	**Description**
AGEV1_1	**YES**	YES	YES	YES	15-19 years (at wave 1)
AGEV1_2		YES		YES	20-24 years (at wave 1)
AGEV1_3	**YES**	**YES**	YES	YES	25-34 years (at wave 1)
AGEV1_4	**YES**	**YES**	YES	YES	35-44 years (at wave 1)
AGEV1_5	**YES**	**YES**	YES	YES	45-54 years (at wave 1)
AGEV1_6	**YES**	**YES**	**YES**	**YES**	55-64 years (at wave 1)
AGEV1_7		**YES**		YES	65+ years (at wave 1)
SINGLE	YES	NO (see Diagram C-1)	**YES**	NO (see Diagram C-1)	See appendix A.
DIVORCED	YES	YES	**YES**	YES	
SEPARATED	YES	YES	YES	YES	
MARRIE	**YES**	**YES**	**YES**	YES	
FORM_PRO	YES	YES	YES	NO (see Diagram C-1)	
FORM_SUP	**YES**	**YES**	**YES**	YES	
NO_FORM	**YES**	**YES**	**YES**	YES	
HOMME	**YES**	**YES**	YES	YES	
SWISS	**YES**	**YES**	YES	YES	
WORK	**YES**	**YES**	YES	NO (see Diagram C-1)	

Diagram C – 1 Dichotomous variables created to model non-response to the grid questionnaire which have Pearson correlation coefficients of > 0.5 in absolute value. The correlations observed are shown here.



Modelling of non-response – analysis by segmentation

SHP_I wave 9
The non-response model created with analysis by segmentation used 31 different variables (25 of which were common to the 2006 modelling) and lead to the creation of 64 homogenous response groups (GHR), 54 for longitudinal adults, 10 for longitudinal children (Diagram C-2).
Stopping criteria for the segmentation: 30% response rate and a minimum of 50 observations per GHR.

SHP_II wave 4
The non-response model created with analysis by segmentation used 4 different variables (all of which were also chosen for the 2006 modelling) and led to the creation of 6 GHR: 4 for longitudinal adults and 2 for longitudinal children (Diagram C-3).
Stopping criteria for the segmentation: 50% response rate and a minimum of 30 observations per GHR.

Figure C-2: Tree for longitudinal adults and children, SHP_I wave 9

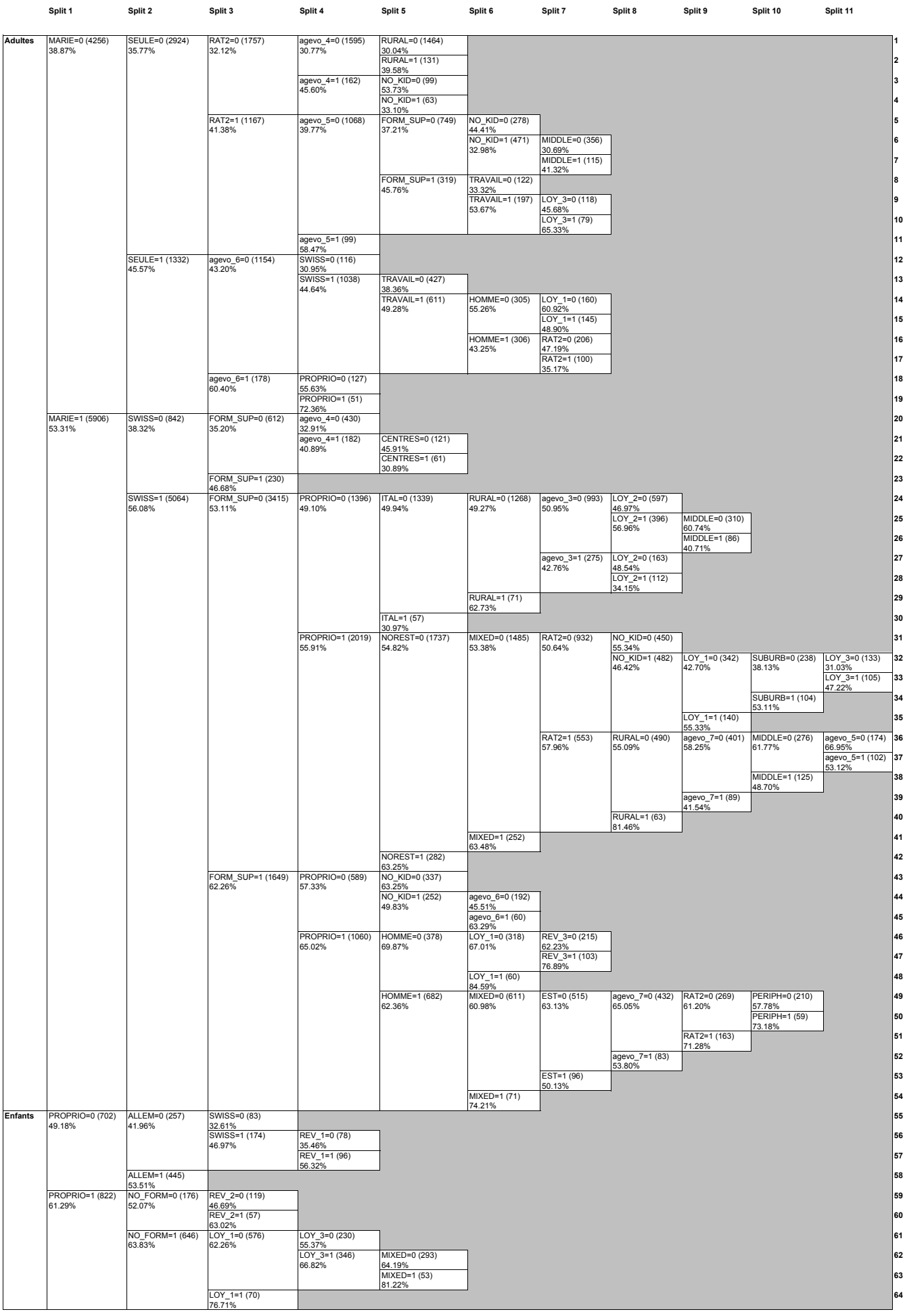


Diagram C-3 Tree for longitudinals, SHP_II wave 4. The tree obtained is very small due to the stopping criteria (50% and a minimum of 30 observations per GHR). As the global response rate of SHP_II was hardly more than 50%, the segmentation stopped very soon. For this year, we decided not to relax the constraints and kept 30% and 50 observations, as was the case for SHP_I, because in contrast to this year, the tree was becoming too big.

	Split 1	Split 2	Split 3	rhg
Adults	agevo_6=0 (4744) 50.51%			**1**
	agevo_6=1 (719) 59.45%	MIDDLE=0 (526) 56.45%	loy_3=0 (359) 52.04%	**2**
			loy_3=1 (167) 65.92%	**3**
		MIDDLE=1 (193) 68.14%		**4**
Children	loy_2=0 (197) 53.60%			**5**
	loy_2=1 (71) 69.04%			**6**

APPENDIX D – Modelling of non-response to the individual questionnaire, individual longitudinal weighting

The dichotomous variables used for the modelling of non-response stem from the grid of the current wave. For the description of variables, please see Appendix A.
The variables ADRFROMW5, REFW2W7CONV, REFW3W8CONV, INC_STAMPS, INC_LOTT, INC_GIFT present in 2006 were no longer available in 2007. However, the variables MAISINDEP, MEN_RR, NOTEASYADR, P1ADRFROMW2_3_4, P1EASYADR, P2EASYADR, HH1, HH2, HH3, HH4+, APPART, DETACH, PROPRIO, LOY_1, LOY_2, LOY_3, RAT1, RAT2, RAT3 have been added.

Household level

Variable	Used for SHP_I in 06	**Used for SHP_I in 07**	Used for SHP_II in 06	**Used for SHP_II in 07**
FRANC	NO	NO (see Diagram D-1)	NO	NO (see Diagram D-1)
ITAL	YES	YES	YES	YES
ALLEM	**YES**	**YES**	YES	YES
LAKE_GE	NO	NO (see Diagram D-1)	NO	NO (see Diagram D-1)
MIDDLE	YES	YES	YES	YES
NOREST	YES	YES	**YES**	YES
EST	**YES**	**YES**	YES	YES
CENTRAL	**YES**	YES	YES	YES
ZURICH	YES	YES	YES	YES
MIXED	YES	YES	YES	YES
CENTRES	YES	YES	YES	**YES**
SUBURB	YES	**YES**	YES	YES
WEALTH	YES	YES	YES	YES
PERIPH	YES	YES	YES	YES
TOURIST	YES	YES	YES	YES
INDUS	YES	YES	YES	YES
RURAL	YES	YES	YES	YES
AD_SEUL	NO	NO (see Diagram D-1)	NO	NO (see Diagram D-1)
NO_KID	**YES**	**YES**	YES	NO (see Diagram D-1)
SEULE	**YES**	**YES**	**YES**	**YES**
ADRFROMW5	**YES**	**--**	**--**	**--**
REFW2W7CONV	YES	**--**	**YES**	**--**
REFW3W8CONV	YES	**--**	YES	**--**
INC_STAMPS	**YES**	**--**	YES	**--**
INC_LOTT	YES	**--**	**YES**	**--**
INC_GIFT	YES	**--**	YES	**--**
MAISINDEP	**--**	YES	**--**	YES
MEN_RR	**--**	**YES**	**--**	**YES**
NOTEASYADR	**--**	**YES**	**--**	NO (see Diagram D-1)
P1ADRFROMW2_3_4	**--**	NO (see Diagram D-1)	**--**	--
P1EASYADR	**--**	NO (see Diagram D-1)	**--**	--
HH2	**--**	NO (see	**--**	YES

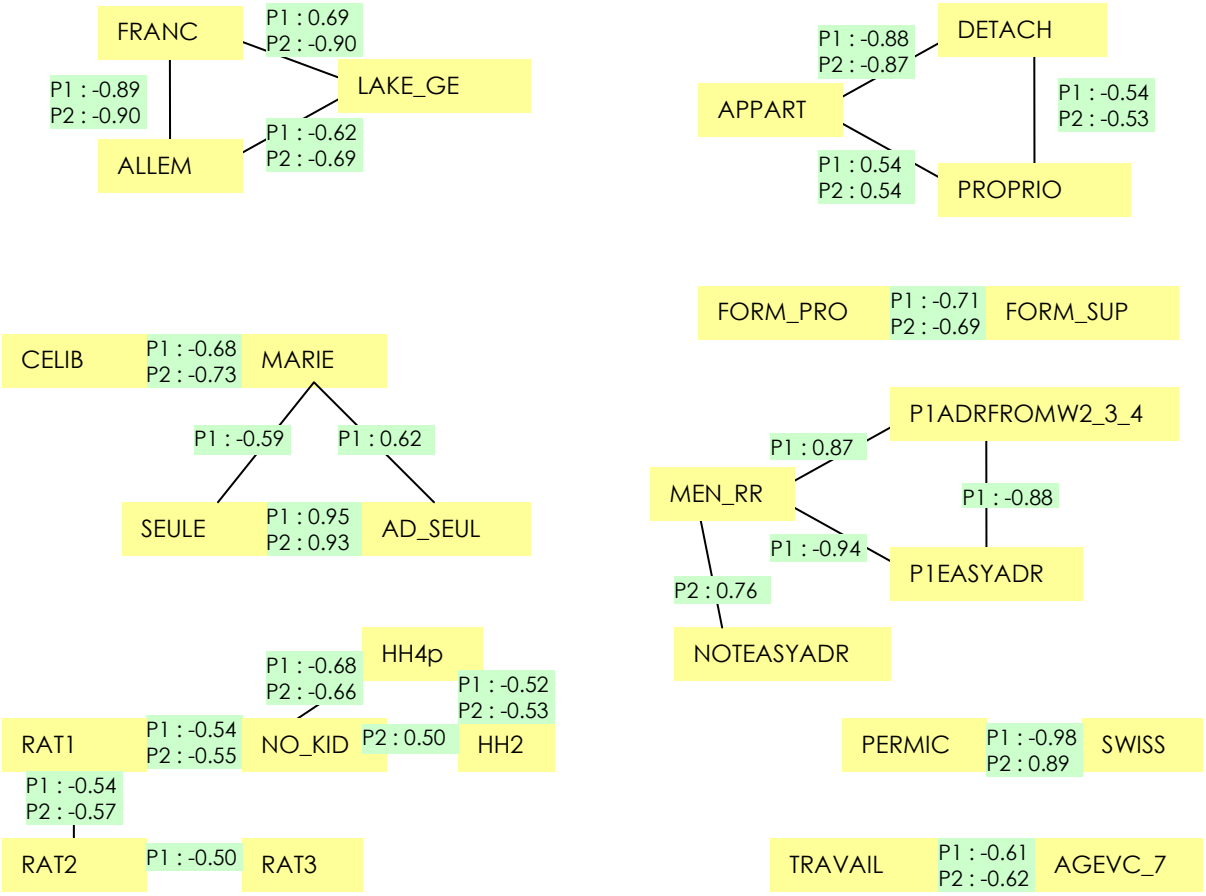
		Diagram D-1)		
HH3	--	YES	--	YES
HH4+	--	YES	--	NO (see Diagram D-1)
APPART		NO (see Diagram D-1)		NO (see Diagram D-1)
DETACH		NO (see Diagram D-1)		NO (see Diagram D-1)
PROPRIO		**YES**		**YES**
LOY_1		**YES**		**YES**
LOY_2		YES		**YES**
LOY_3		**YES**		**YES**
RAT1		NO (see Diagram D-1)		NO (see Diagram D-1)
RAT2		NO (see Diagram D-1)		**YES**
RAT3		**YES**		YES

Individual level

We introduced a new definition of age groups in 2007, by separating the 15 to 19 year olds from the 20 to 24 year olds. In fact, we observed that these two sub-groups of the population behaved in opposite ways as regards participation rates. We also added an age group for the 55 to 64 year olds.

Variable	Used for SHP_I in 06	**Used for SHP_I in 07**	Used for SHP_II in 06	**Used for SHP_II in 07**
AGEVC_1	**YES**	**YES**	YES	YES
AGEVC_2		YES		YES
AGEVC_3	YES	YES	YES	YES
AGEVC_4	YES	YES	**YES**	YES
AGEVC_5	YES	YES	**YES**	YES
AGEVC_6	YES	YES	YES	YES
AGEVC_7		NO (see Diagram D-1)		**YES**
CELIB	NO	YES	NO	YES
DIVORCE	YES	YES	YES	YES
SEPARE	YES	YES	YES	YES
MARIE	YES	NO (see Diagram D-1)	**YES**	YES
VEUF	YES	YES	YES	YES
FORM_PRO	NO	NO (see Diagram D-1)	NO	NO (see Diagram D-1)
FORM_SUP	**YES**	**YES**	**YES**	**YES**
NO_FORM	**YES**	**YES**	**YES**	**YES**
HOMME	**YES**	**YES**	**YES**	**YES**
SWISS	YES	**YES**	**YES**	**YES**
TRAVAIL	YES	**YES**	YES	NO (see Diagram D-1)
PERMIB	YES	NO (too few observations)	YES	YES
PERMIC	NO	NO (see Diagram D-1)	NO	NO (see Diagram D-1)

Diagram D – 1 Dichotomous variables created to model non-response to the grid questionnaire which have Pearson correlation coefficients of > 0.5 in absolute value. P1=SHP_I, P2=SHP_II.



Synthesis

11 different variables (8 in common with the modelling for the previous wave) were chosen for SHP_I, 13 for SHP_II (5 in common with the modelling for the previous wave), 9 of which feature in both models.

The model obtained with the analysis by segmentation includes 26 GHR for SHP_I and 20 for SHP_II. A weighted response rate of at least 30% and a minimum of 50 individuals were required in each GHR for the SHP_I; for SHP_II the constraints were 50% and 30 individuals per GHR.

Figure D – 2 Model for non-response to the individual questionnaire: individual longitudinal weighting of SHP_I.

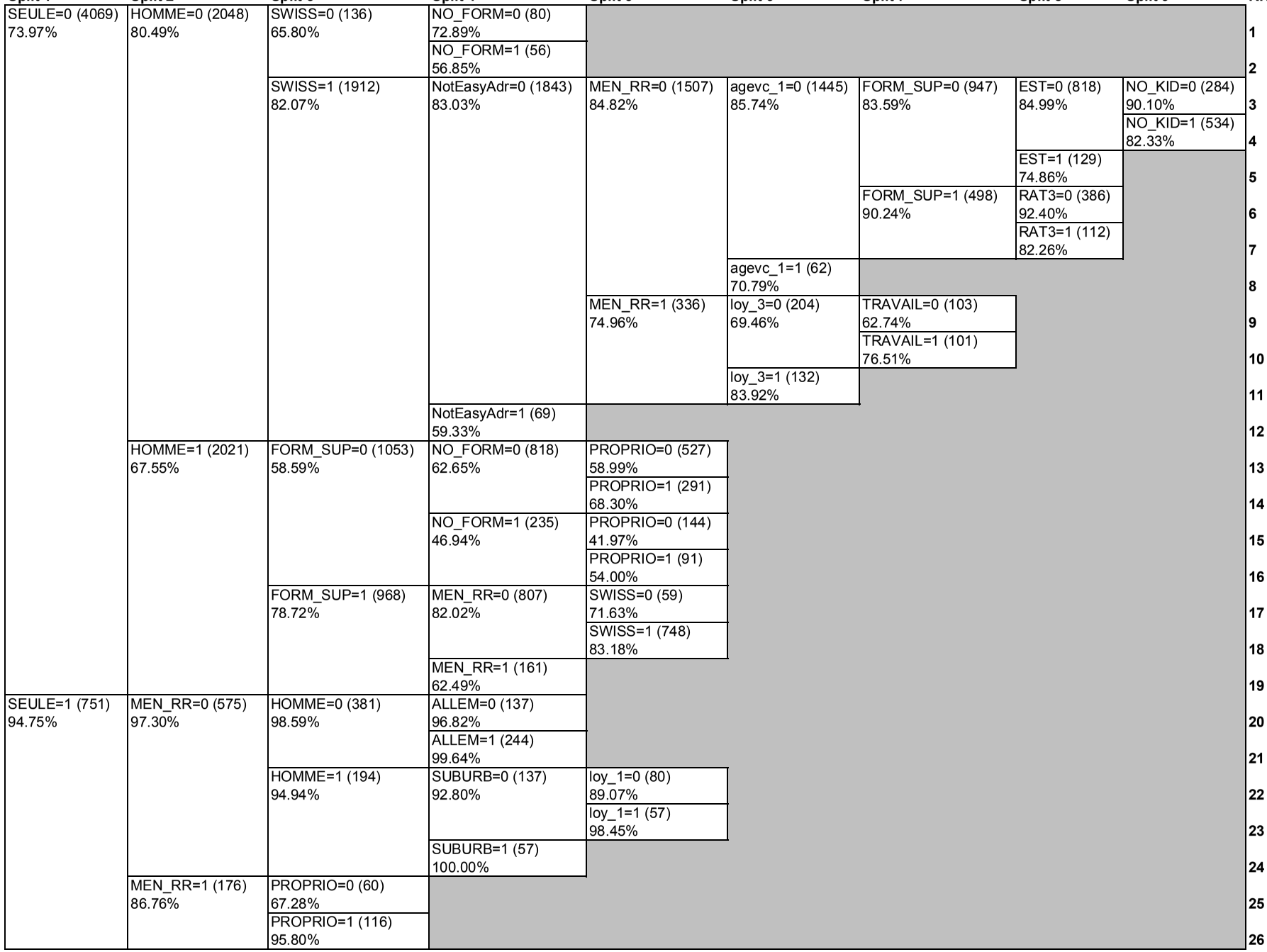
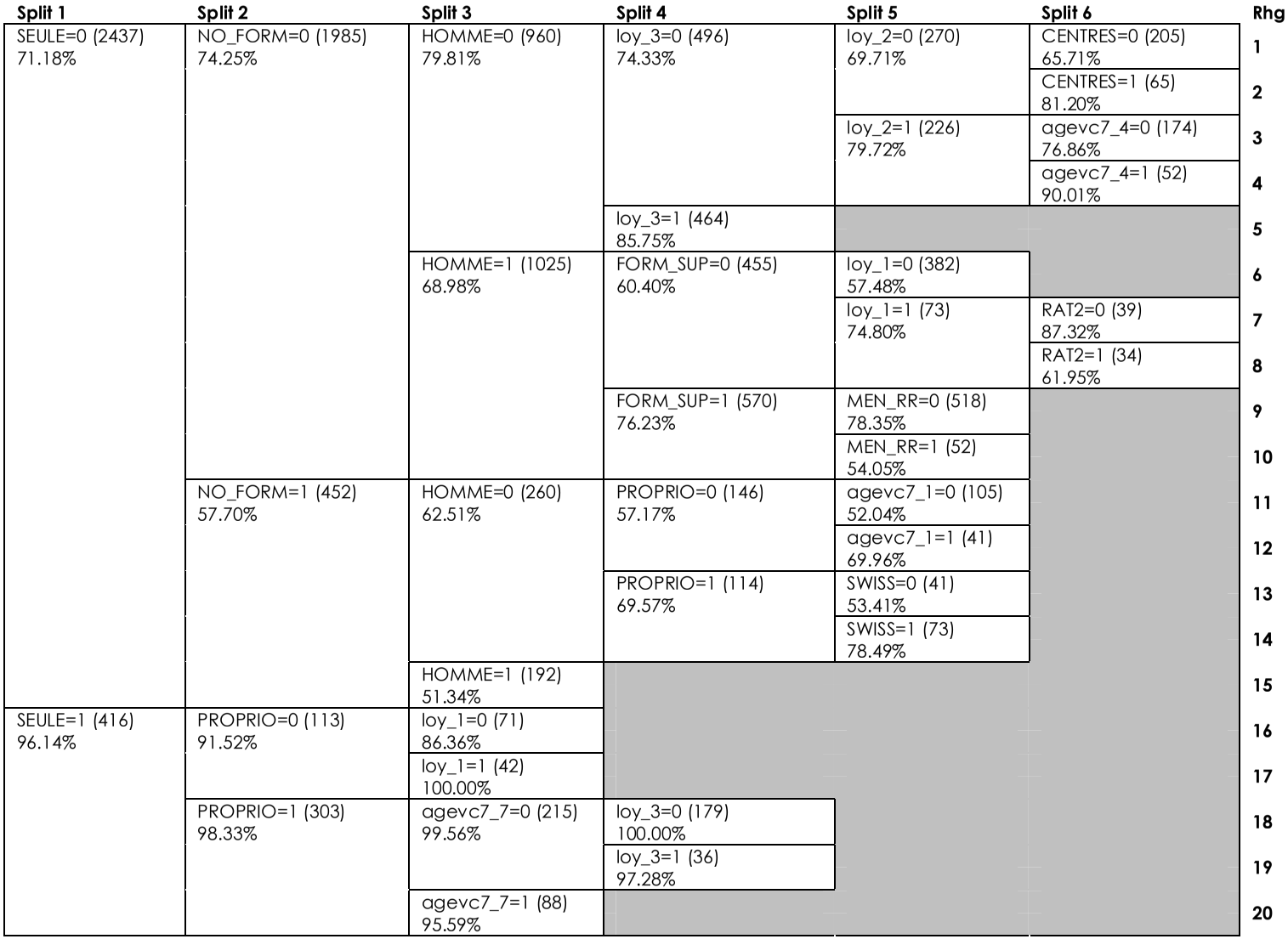


Diagram D – 3 Complete model for non-response to the individual questionnaire: individual longitudinal weighting of SHP_II The numbers between brackets are the non-weighted numbers of individuals. The percentages represent the weighted response rates.



APPENDIX E – Modelling of non-response to the individual questionnaire, individual cross-sectional weighting

The dichotomous variables used for the modelling of non-response stem from the grid of the current wave. For the description of variables, see Appendix A.

Household level

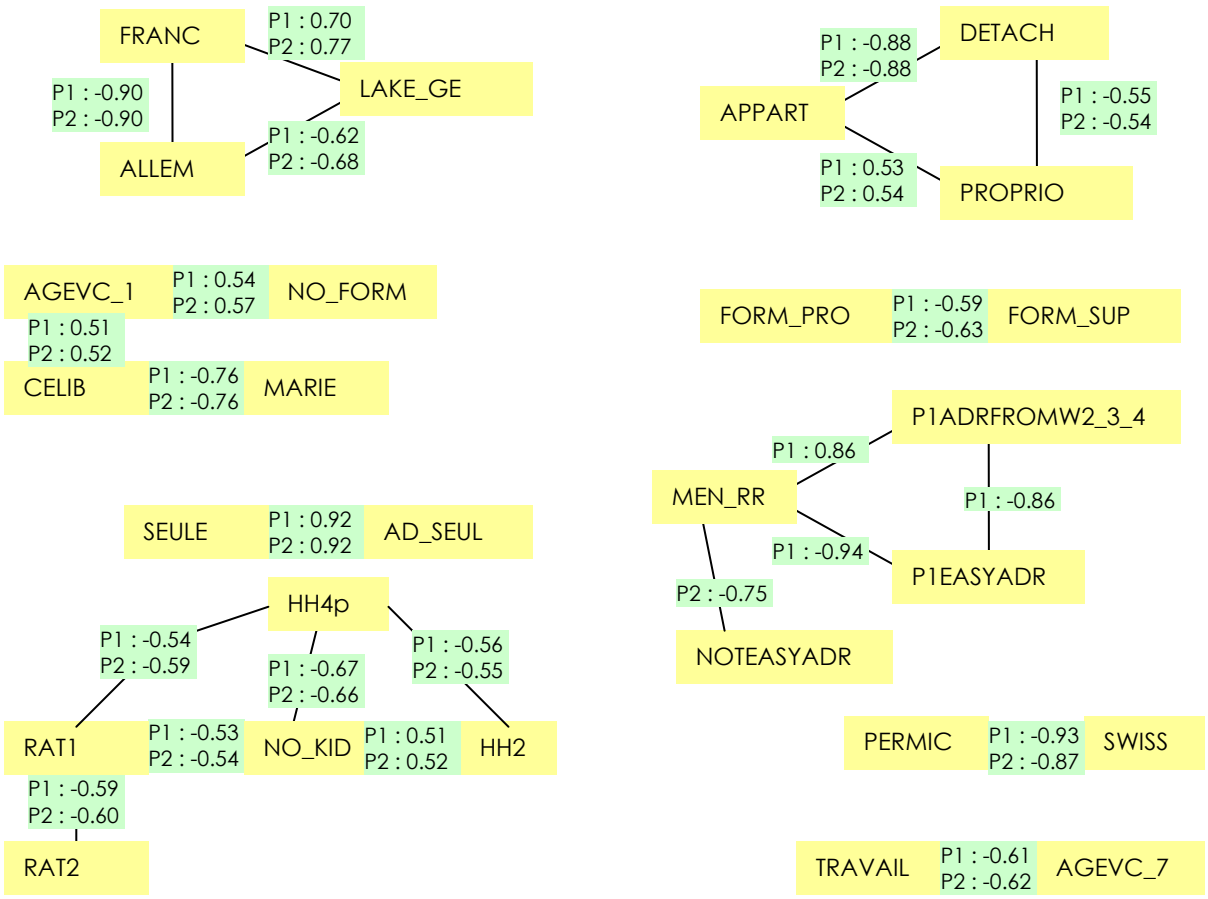
Variable	Used for SHP_I in 06	**Used for SHP_I in 07**	Used for SHP_II in 06	**Used for SHP_II in 07**
FRANC	NO	NO (see Diagram E-1)	NO	NO (see Diagram E-1)
ITAL	YES	YES	YES	YES
ALLEM	**YES**	**YES**	YES	YES
LAKE_GE	NO	NO (see Diagram E-1)	NO	NO (see Diagram E-1)
MIDDLE	YES	**YES**	YES	YES
NOREST	YES	**YES**	YES	YES
EST	**YES**	YES	YES	YES
CENTRAL	YES	YES	YES	YES
ZURICH	YES	**YES**	YES	YES
MIXED	**YES**	YES	YES	YES
CENTRES	**YES**	YES	**YES**	**YES**
SUBURB	YES	**YES**	YES	YES
WEALTH	YES	YES	YES	YES
PERIPH	YES	YES	**YES**	YES
TOURIST	YES	YES	YES	YES
INDUS	YES	YES	YES	YES
RURAL	YES	YES	YES	YES
AD_SEUL	YES	NO (see Diagram E-1)	**YES**	NO (see Diagram E-1)
NO_KID	**YES**	**YES**	YES	NO (see Diagram E-1)
SEULE	**YES**	**YES**	YES	**YES**
ADRFROMW5	**YES**			
REFW2W7CONV	YES		**YES**	
REFW3W8CONV	YES		**YES**	
INC_STAMPS	**YES**		**YES**	
INC_LOTT	YES		**YES**	
INC_GIFT	YES		**YES**	
Loy_1		**YES**		**YES**
Loy_2		**YES**		**YES**
Loy_3		**YES**		**YES**
RAT_1		NO (see Diagram E-1		NO (see Diagram E-1)
RAT_2		**YES**		**YES**
RAT_3		**YES**		YES
PROPRIO		**YES**		**YES**
MAISINDEP		YES		YES
APPART		NO (see Diagram E-1)		NO (see Diagram E-1)
DETACH		NO (see Diagram E-1)		NO (see Diagram E-1)
MEN_RR		**YES**		YES
NotEasyAdr		**YES**		NO (see

				Diagram E-1)
P1AdrFromW2_3_4		NO (see Diagram E-1)		
P1EasyAdr		NO (see Diagram E-1)		
P2EasyAdr				NO (=-MEN_RR)
HH2				YES
HH3				YES
HH4p				NO (see Diagram E-1)

Individual level

Variable	Used for SHP_I in 06	**Used for SHP_I in 07**	Used for SHP_II in 06	**Used for SHP_II in 07**
AGEVC_1	**YES**	**YES**	YES	**YES**
AGEVC_2		**YES**		YES
AGEVC_3	**YES**	YES	YES	YES
AGEVC_4	YES	**YES**	**YES**	**YES**
AGEVC_5	YES	YES	**YES**	**YES**
AGEVC_6	**YES**	**YES**	YES	YES
AGEVC_7		**YES**		**YES**
CELIB	NO	NO (see Diagram E-1)	NO	NO (see Diagram E-1)
DIVORCE	YES	YES	**YES**	YES
SEPARE	YES	YES	YES	YES
MARIE	**YES**	YES	**YES**	**YES**
VEUF	YES	YES	YES	YES
FORM_PRO	NO	NO (see Diagram E-1)	NO	NO (see Diagram E-1)
FORM_SUP	**YES**	**YES**	**YES**	**YES**
NO_FORM	**YES**	**YES**	**YES**	YES
HOMME	**YES**	**YES**	**YES**	**YES**
SWISS	**YES**	**YES**	**YES**	**YES**
TRAVAIL	YES	NO (see Diagram E-1)	YES	NO (see Diagram E-1)
PERMIB	YES	NO (see Diagram E-1)	YES	YES
PERMIC	NO	NO (see Diagram E-1)	NO	NO (see Diagram E-1)

Diagram E – 1 Dichotomous variables created to model non-response to the grid questionnaire which have Pearson correlation coefficients of > 0.5 in absolute value.



Synthesis

24 and 15 different variables were chosen for SHP_I and SHP_II respectively, 12 of which feature in both models. The model obtained with the analysis by segmentation includes 39 GHR for SHP_I and 23 for SHP_II. A weighted response rate of at least 30% and a minimum of 50 individuals were required in each GHR for the SHP_I; for SHP_II the constraints were 50% and 30 individuals per GHR.

Figure E-2 Model for non-response to the individual questionnaire: Individual cross-sectional weighting of SHP_I.

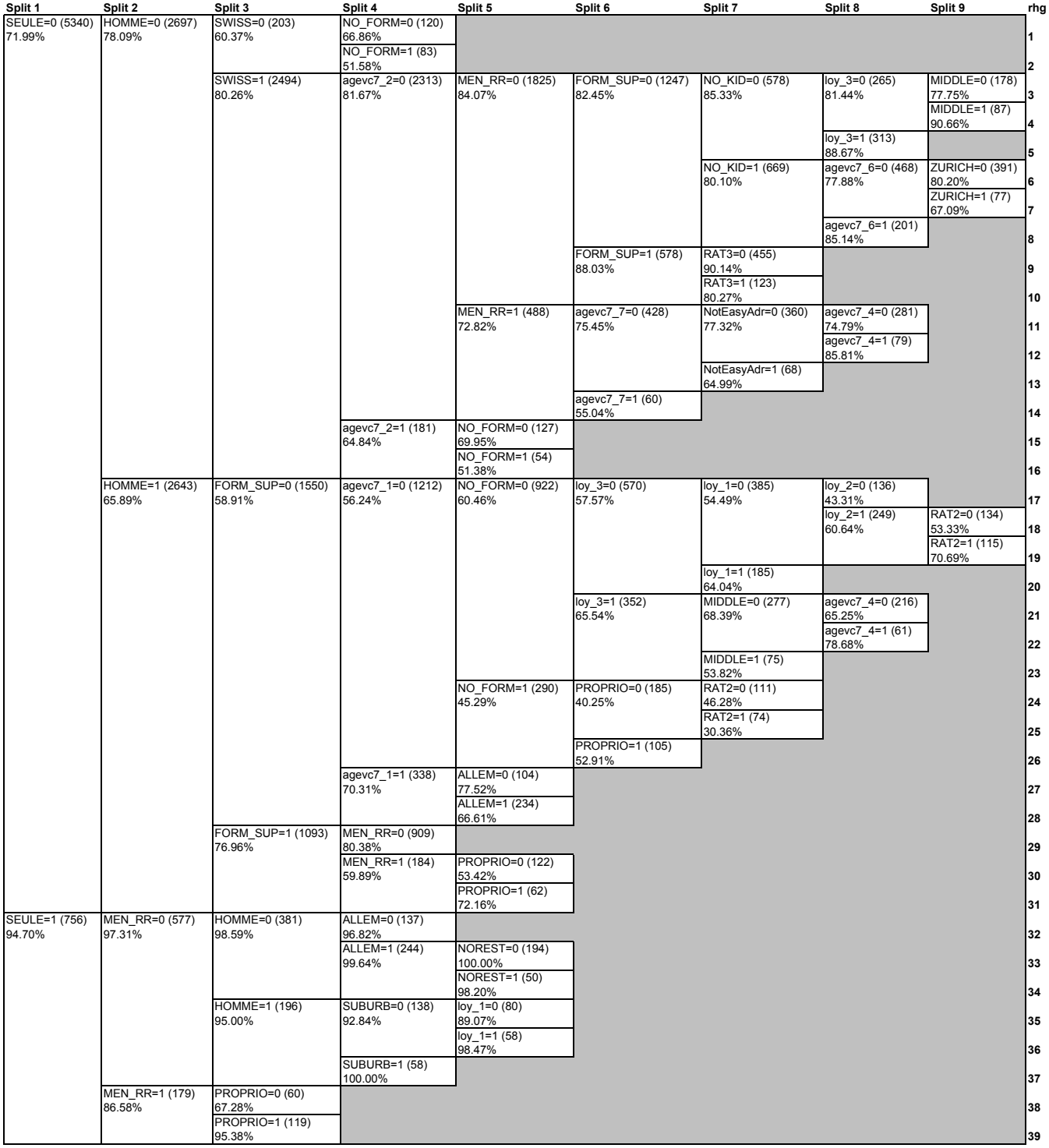


Diagram E – 3 Complete model for non-response to the individual questionnaire: Individual cross-sectional weighting of SHP_II.

Split 1	Split 2	Split 3	Split 4	Split 5	Split 6	Split 7	Rhg
SEULE=0 (2758) 70.22%	FORM_SUP=0 (1716) 65.77%	HOMME=0 (951) 71.45%	loy_3=0 (560) 66.58%	loy_2=0 (320) 60.61%	loy_1=0 (167) 52.67%		1
					loy_1=1 (153) 69.16%		2
				loy_2=1 (240) 74.33%	SWISS=0 (36) 54.20%		3
					SWISS=1 (204) 77.78%	agevc7_4=0 (166) 74.88%	4
						agevc7_4=1 (38) 91.14%	5
			loy_3=1 (391) 78.59%				6
		HOMME=1 (765) 58.81%	loy_1=0 (662) 56.83%	agevc7_1=0 (502) 53.23%	loy_2=0 (338) 50.12%		7
					loy_2=1 (164) 59.45%		8
				agevc7_1=1 (160) 68.94%	CENTRES=0 (128) 73.75%	loy_3=0 (55) 64.15%	9
						loy_3=1 (73) 81.27%	10
					CENTRES=1 (32) 50.62%		11
			loy_1=1 (103) 70.70%				12
	FORM_SUP=1 (1042) 77.96%	loy_3=0 (467) 72.65%	agevc7_5=0 (377) 75.24%				13
			agevc7_5=1 (90) 62.72%	HOMME=0 (41) 75.77%			14
				HOMME=1 (49) 51.86%			15
		loy_3=1 (575) 82.13%	MARIE=0 (181) 76.18%	RAT2=0 (81) 67.55%			16
				RAT2=1 (100) 83.82%			17
			MARIE=1 (394) 84.52%				18
SEULE=1 (416) 96.14%	PROPRIO=0 (113) 91.52%	loy_1=0 (71) 86.36%					19
		loy_1=1 (42) 100.00%					20
	PROPRIO=1 (303) 98.33%	agevc7_7=0 (215) 99.56%	loy_3=0 (179) 100.00%				21
			loy_3=1 (36) 97.28%				22
		agevc7_7=1 (88) 95.59%					23

APPENDIX F – Modelling of non-response to the household questionnaire, household cross-sectional weighting

The variables used for modelling are dichotomous variables which describe households and the individuals who live in them and which stem from data (grid questionnaires) from wave 1 (1999 for SHP_I, 2004 for SHP_II). Here is a list of variables which could potentially be included in the non-response model. For the description of these variables, refer to Appendix A.

Variable	Used for SHP_I in 06	**Used for SHP_I in 07**	Used for SHP_II in 06	**Used for SHP_II in 07**
FRANC	NO	NO (see Diagram F-1)	NO	NO (see Diagram F-1)
ITAL	YES	YES	YES	YES
ALLEM	YES	YES	YES	YES
LAKE_GE	NO	NO (see Diagram F-1)	NO	NO (see Diagram F-1)
MIDDLE	YES	YES	YES	YES
NOREST	YES	YES	YES	YES
EST	YES	YES	YES	YES
CENTRAL	YES	YES	YES	YES
ZURICH	YES	YES	YES	YES
MIXED	YES	YES	YES	YES
CENTRES	YES	YES	YES	YES
SUBURB	YES	YES	YES	YES
WEALTH	YES	YES	YES	YES
PERIPH	YES	YES	YES	YES
TOURIST	YES	YES	YES	YES
INDUS	YES	YES	YES	YES
RURAL	YES	YES	YES	YES
AD_SEUL	NO	YES	NO	NO (see Diagram F-1)
NO_KID	YES	YES	YES	YES
SEULE	YES	YES	YES	YES
ADRFROMW5	YES			
REFW2W7CONV	YES		YES	
REFW3W8CONV	YES		YES	
INC_STAMPS	YES		YES	
INC_LOTT	YES		YES	
INC_GIFT	YES		YES	
APPART	YES	NO (see Diagram F-1)	**YES**	NO (see Diagram F-1)
DETACH	NO	NO (see Diagram F-1)	NO	NO (see Diagr. F-1)
PROPRIO	**YES**	**YES**	**YES**	**YES**
RAT1	YES	NO (see Diagram F-1)	YES	NO (see Diagram F-1)
RAT2	YES	YES	YES	YES
RAT3	**YES**	NO (see Diagram F-1)	YES	NO (see Diagram F-1)
LOY_1	**YES**	**YES**	**YES**	**YES**
LOY_2	**YES**	**YES**	**YES**	**YES**
LOY_3	**YES**	**YES**	**YES**	**YES**
TOUS_HOMMES	YES	YES	YES	YES
TOUS_FEMMES	NO	NO (see Diagram F-1)	NO	YES
NON_MARIE	NO	NO (see	NO	NO (see

		Diagram F-1)		Diagram F-1)
max_no_educ	YES	YES	YES	YES
max_educ_sup	YES	**YES**	YES	YES
max_educ_pro	NO	NO (see Diagram F-1)	NO	YES
tous_suis	YES	YES	**YES**	YES
tous_etrang	NO	NO (see Diagram F-1)	NO	NO (see Diagram F-1)
agemax_1	YES	YES	YES	YES
agemax_2	YES	YES	YES	YES
agemax_3	YES	YES	YES	YES
agemax_4	YES	YES	YES	YES
agemax_5	**YES**	NO (see Diagram F-1)	**YES**	YES
MEN_RR		**YES**		**YES**
NotEasyAdr		**YES**		NO (see Diagram F-1)
P1AdrFromW2_3_4		NO (see Diagram F-1)		
P1EasyAdr		NO (see Diagram F-1)		
P2EasyAdr				NO (see Diagram F-1)
HH2		YES		YES
HH3		YES		YES
HH4p		NO (see Diagram F-1)		NO (see Diagram F-1)

Synthesis

7 and 5 different variables were chosen for SHP_I and SHP_II respectively, 5 of which feature in both models. The model obtained with the analysis by segmentation includes 10 GHR for SHP_I and 7 for SHP_II. A weighted response rate of at least 30% and a minimum of 50 individuals were required in each GHR for the SHP_I; for SHP_II the constraints were 50% and 30 individuals per GHR.

Diagram F – 1 Dichotomous variables created to model non-response to the household questionnaire which have Pearson correlation coefficients of > 0.5 in absolute value.

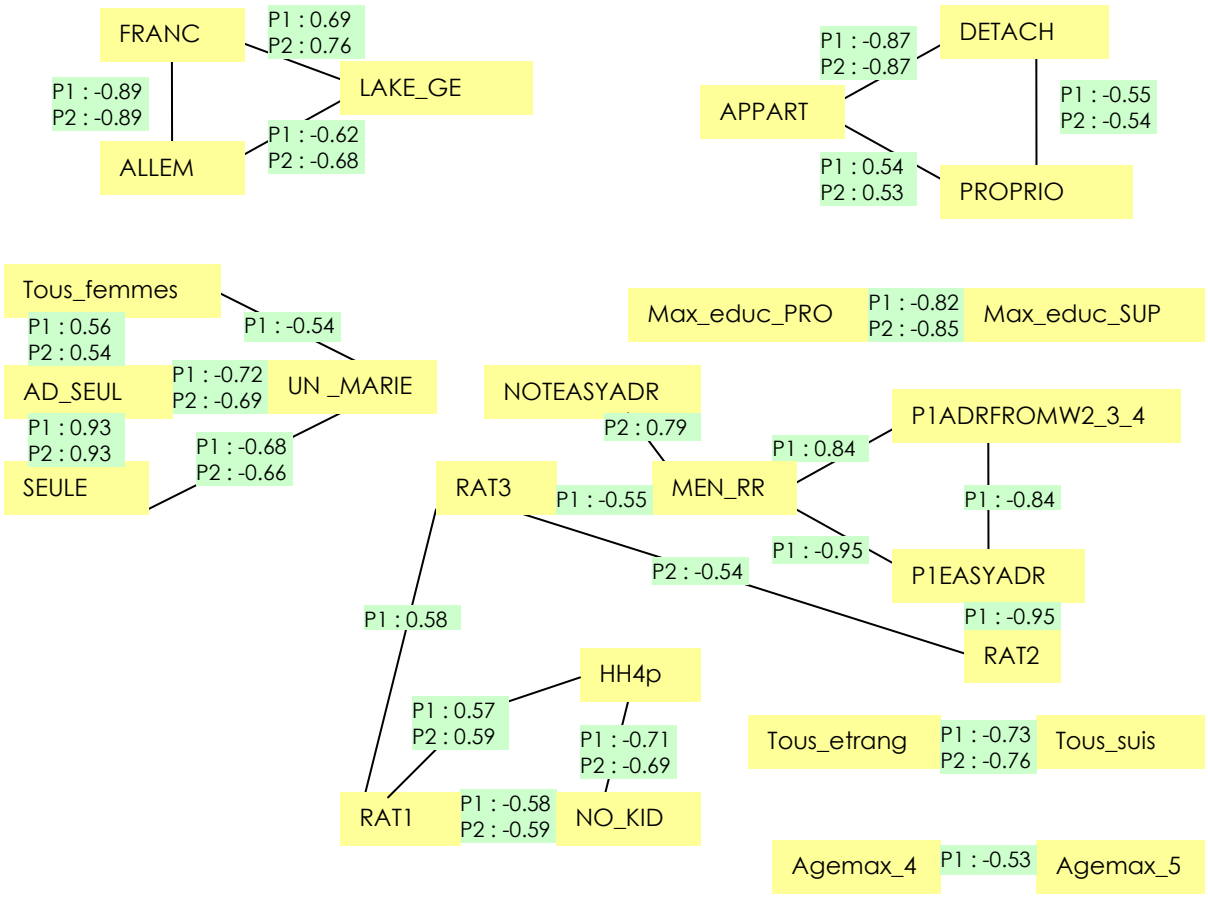


Diagram F – 2 Complete model for non-response to the household questionnaire: cross-sectional household weighting of SHP_I. The numbers between brackets are the non-weighted number of individuals. The percentages represent the weighted response rates.

Split 1	Split 2	Split 3	Split 4	Split 5	Split 6	Split 7	Rhg
PROPRIO=0 (1576) 94.31%	NotEasyAdr=0 (1514) 95.29%	MEN_RR=0 (1230) 97.16%	loy_3=0 (726) 95.36%	loy_1=0 (441) 92.34%	loy_2=0 (228) 85.46%	max_educ_sup=0 (111) 80.73%	1
						max_educ_sup=1 (117) 90.26%	2
					loy_2=1 (213) 100.00%		3
				loy_1=1 (285) 100.00%			4
			loy_3=1 (504) 100.00%				5
		MEN_RR=1 (284) 87.28%	loy_3=0 (196) 82.38%	loy_1=0 (132) 74.30%			6
				loy_1=1 (64) 100.00%			7
			loy_3=1 (88) 100.00%				8
	NotEasyAdr=1 (62) 69.39%						9
PROPRIO=1 (1317) 100.00%							10

Diagram F – 3 Complete model for non-response to the household questionnaire: household cross-sectional weighting of SHP_II. The numbers between brackets are the non-weighted numbers of individuals. The percentages represent the weighted response rates.

<table>
<tr><th>Split 1</th><th>Split 2</th><th>Split 3</th><th>Split 4</th><th>Split 5</th><th></th></tr>
<tr><td rowspan="5">MEN_RR=0
(1374)
97.99%</td><td rowspan="3">PROPRIO=0
(684)
96.05%</td><td rowspan="2">loy_3=0
(401)
93.35%</td><td>loy_1=0
(271)
89.99%</td><td>loy_2=0
(143)
80.11%</td><td>1</td></tr>
<tr><td></td><td>loy_2=1
(128)
100.00%</td><td>2</td></tr>
<tr><td></td><td>loy_1=1
(130)
100.00%</td><td></td><td>3</td></tr>
<tr><td>loy_3=1
(283)
100.00%</td><td></td><td></td><td>4</td></tr>
<tr><td>PROPRIO=1
(690)
100.00%</td><td></td><td></td><td></td><td>5</td></tr>
<tr><td rowspan="2">MEN_RR=1
(169)
83.38%</td><td>PROPRIO=0
(85)
68.80%</td><td></td><td></td><td></td><td>6</td></tr>
<tr><td>PROPRIO=1
(84)
100.00%</td><td></td><td></td><td></td><td>7</td></tr>
</table>

APPENDIX G – Modelling of non-response to the individual questionnaire, individual longitudinal weighting of SHP_I and SHP_II combined.

The dichotomous variables used for the modelling of non-response stem from the grid of the current wave. This modelling is only concerned here with longitudinal individuals in relation to the year 2004 of the SHP_I. In fact, the modelling must be done again because for the combined longitudinal weighting, the reference year is 2004 and no longer 1999, which necessitates a redefinition of the longitudinals in relation to the year 2004 (see paragraph 12.1). For the description of variables, please see Appendix A.

Household level

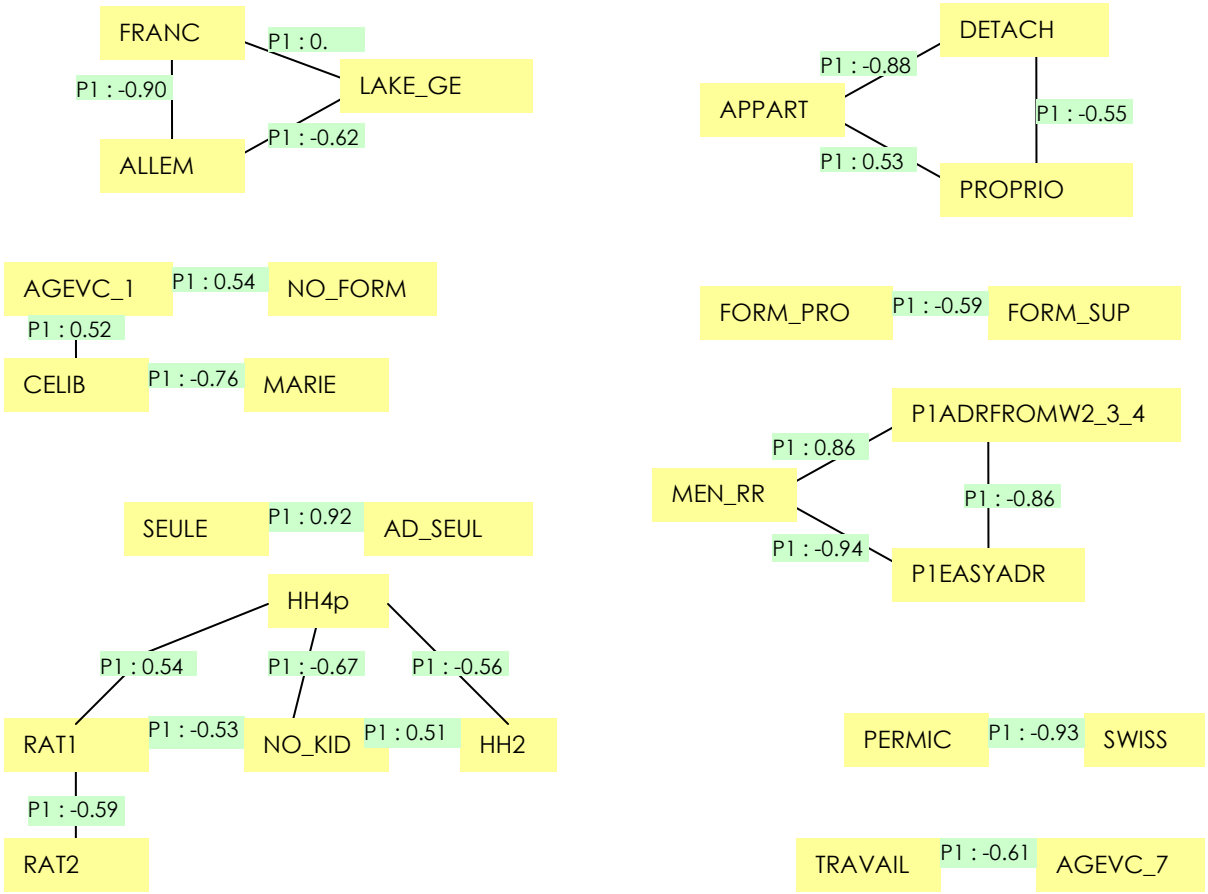
Variable	Used for SHP_I in 06	**Used for SHP_I in 07**
FRANC	NO	NO (see Diagram G-1)
ITAL	YES	YES
ALLEM	**YES**	**YES**
LAKE_GE	NO	NO (see Diagram G-1)
MIDDLE	YES	YES
NOREST	YES	YES
EST	**YES**	**YES**
CENTRAL	**YES**	YES
ZURICH	YES	YES
MIXED	YES	YES
CENTRES	**YES**	**YES**
SUBURB	YES	**YES**
WEALTH	YES	YES
PERIPH	YES	YES
TOURIST	YES	YES
INDUS	YES	YES
RURAL	YES	YES
AD_SEUL	NO	NO (see Diagram G-1)
NO_KID	**YES**	NO (see Diagram G-1)
SEULE	**YES**	**YES**
ADRFROMW5	**YES**	
REFW2W7CONV	YES	
REFW3W8CONV	YES	
INC_STAMPS	**YES**	
INC_LOTT	YES	
INC_GIFT	YES	
Loy_1		**YES**
Loy_2		**YES**
Loy_3		**YES**
RAT_1		YES
RAT_2		NO (see Diagram G-1)
RAT_3		**YES**
PROPRIO		**YES**
MAISINDEP		YES
APPART		NO (see

Individual level

Variable	Used for SHP_I in 06	**Used for SHP_I in 07**
AGEVC_1	**YES**	YES
AGEVC_2		**YES**
AGEVC_3	YES	**YES**
AGEVC_4	YES	YES
AGEVC_5	YES	YES
AGEVC_6	YES	YES
AGEVC_7		YES
CELIB	NO	NO (see Diagram G-1)
DIVORCE	YES	YES
SEPARE	YES	YES
MARIE	YES	YES
VEUF	YES	YES
FORM_PRO	NO	NO (see Diagram G-1)
FORM_SUP	**YES**	**YES**
NO_FORM	**YES**	NO (see Diagram G-1)
HOMME	**YES**	**YES**
SWISS	YES	**YES**
TRAVAIL	YES	NO (see Diagram G-1)
PERMIB	NO	NO (see Diagram G-1)
PERMIC	NO	NO (see Diagram G-1)

		Diagram G-1)
DETACH		NO (see Diagram G-1)
MEN_RR		**YES**
NotEasyAdr		YES
P1AdrFromW2_3_4		NO (see Diagram G-1)
P1EasyAdr		NO (see Diagram G-1)
P2EasyAdr		NO (see Diagram G-1)
HH2		YES
HH3		YES
HH4p		NO (see Diagram G-1)

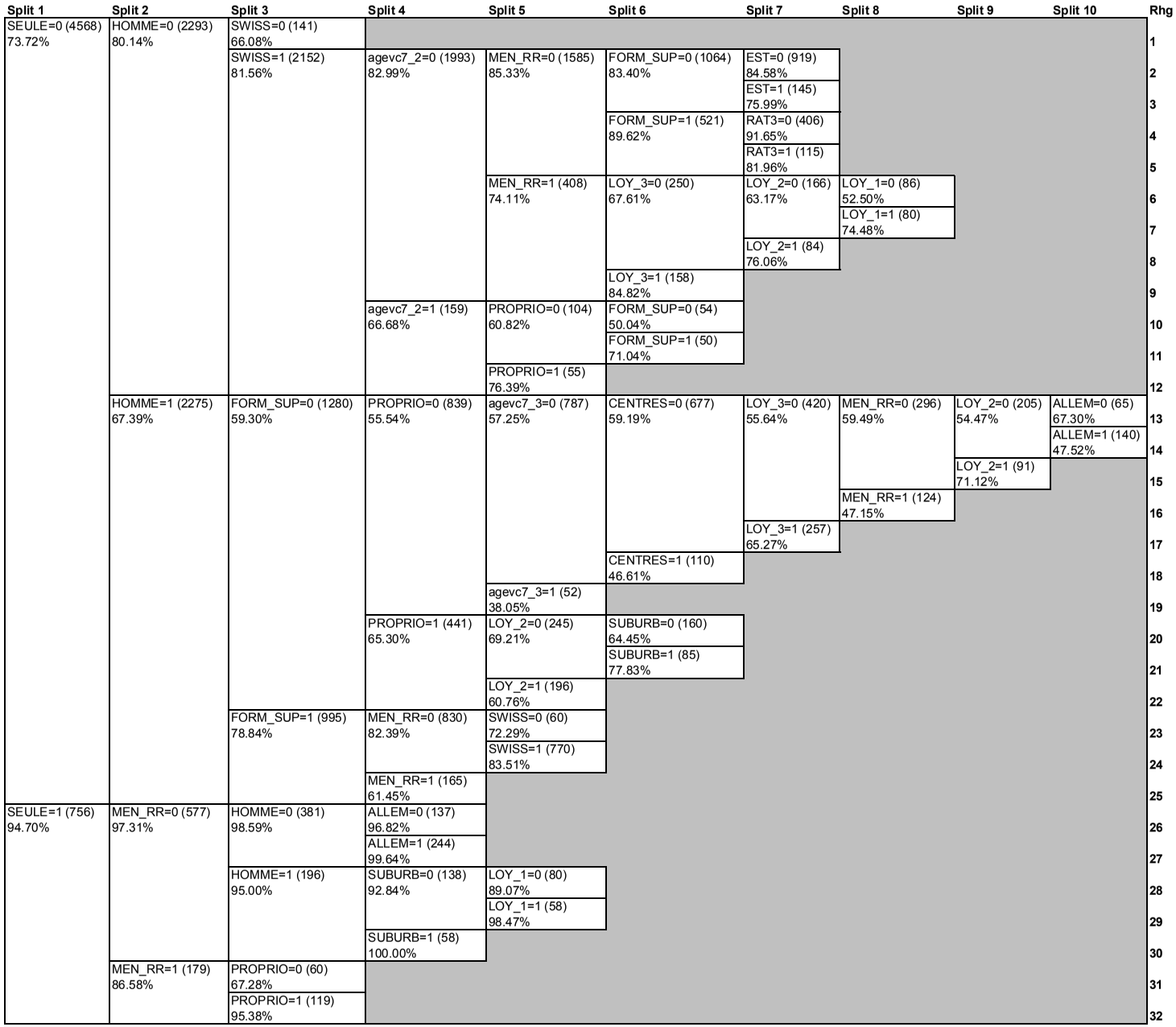
Table G – 1 Dichotomous variables created to model non-response having Pearson correlation coefficients of > 0.5 in absolute value.



Modelling of non-response – analysis by segmentation

The model obtained with analysis by segmentation includes 19 GHR for the SHP_I. We used 16 variables, 7 of which are in common with the previous year. A weighted response rate of at least 30% and a minimum number of 50 individuals were required for each GHR.

Figure G-2 Complete model for non-response to the individual quest. for longitudinals from SHP_I (individual longitudinal weighting for SHP_I and SHP_II combined).



APPENDIX H – Characteristics of distributions of produced weightings

Table H – 1 Longitudinal individual weighting for SHP_I wave 9 (2007). For each set of weights, we calculated all the statistics on the sub-group of weights > 0. We also made a comparison with the same statistics for the weightings of the previous year.

PTI_NRQI and PL_NRQI weight corrected for non-response to the individual questionnaire
POIDTRIN final individual cross-sectional weight
POIDLONG final individual longitudinal weight

Name of set of weights	PTI_NRQI 06	**PTI_NRQI 07**	POIDTRIN 06	**POIDTRIN 07**	PL_NRQI 06	**PL_NRQI 07**	POIDLONG 06	**POIDLONG 07**
N	4088	4626	4088	4626	3617	3959	3617	3959
Total of weights	6462474	6554824	6196348	6554824	5771759	5771759	5998212	5998212
Average	1581	1417	1516	1417	1595	1458	1658	1515
Standard deviation	678	651	1023	651	699	660	968	912
Variance	459568	424409	1047053	424409	488682	435770	937888	831666
Range	5993	5971	10861	5971	5416	4869	9741	7917
Interquartile interval	751	644	716	644	839	726	1034	956
Max/min ratio	32	24	80	24	13	12	27	22
Ratio P99 /P1	6	8	11	8	8	8	11	10
Ratio P95 / P5	4	4	5	4	4	4	5	5
Quantiles								
100% max	6185	6231	10999	10736	5875	5329	10121	8291
99%	3840	3974	5705	6339	4128	3790	5214	5020
95%	2926	2688	3620	3380	2867	2741	3580	3282
90%	2482	2227	2521	2345	2484	2318	2713	2578
75% Q3	1863	1639	1664	1432	1945	1732	2020	1841
50% median	1408	1255	1207	1034	1428	1298	1439	1259
25% Q1	1112	996	948	817	1106	1006	986	885
10%	943	827	790	681	929	827	795	728
5%	836	715	679	599	713	675	694	670
1%	599	514	500	432	501	501	490	502
0% min	192	260	138	167	460	460	381	374

Table H – 2 Cross-sectional household weighting SHP_I wave 9 (2007). For each set of weights, we calculated all the statistics on the sub-group of weights > 0. We also made a comparison with the same statistics for the weightings of the previous year.

P_NRGRIL weight adjusted for non-response to the grid questionnaire
PTM_NRQM weight adjusted for non-response to the household questionnaire
POIDTRME_INT1 weight after first calibration
POIDTRME_INT weight after second calibration
POIDTRME_INT_WIN final cross-sectional household weight (Winsorised at the 1st and 99th percentiles). Winsorisation is carried out on the set of weights POIDTRME_INT on the file at individual level. This is why the numbers given here, calculated at household level, do not tally exactly when the total, the P1, P99 of the weights _INT are compared with the weights _INT_WIN.

Name of set of weights	P_NRGRIL 06	**P_NRGRIL 07**	PTM_NRQM 06	**PTM_NRQM 07**	POIDTRME_ INT1 06	**POIDTRME_ INT1 07**	POIDTRME_ INT 06	**POIDTRME_ INT 07**	POIDTRME_ INT_WIN 06	**POIDTRME_ INT_WIN 07**
N	2028	2301	2537	2817	2537	2817	2537	2817	2537	2817
Total of weights	2587667	2666800	3140124	3201321	3056981	3103928	3053584	3103183	3064134	3109426
Average	1276	1159	1238	1136	1205	1102	1204	1102	1208	1104
Standard deviation	399	383	395	390	730	751	730	751	708	725
Variance	159500	146983	155922	152384	532354	563588	532847	564514	500612	526284
Range	2110	2300	2819	2720	7625	7558	7286	7187	4088	4240
Interquartile interv.	625	481	573	459	631	589	633	584	637	588
Max/min ratio	4	5	17	13			223	73	26	20
Ratio P99 /P1	3	4	5	5	15	15	15	15	15	15
Ratio P95 / P5	3	3	3	3	5	6	5	6	5	6
Quantiles										
100% max	2742	2862	2995	2955	7321	7206	7318	7287	4254	4463
99%	2301	2207	2269	2215	4079	4199	4107	4202	4138	4231
95%	2025	1946	1975	1892	2596	2615	2593	2620	2613	2638
90%	1884	1759	1807	1692	1994	1917	1985	1907	2000	1919
75% Q3	1576	1351	1511	1335	1390	1264	1390	1259	1400	1267
50% median	1168	1056	1166	1058	1018	862	1013	862	1021	868
25% Q1	951	870	938	876	759	675	757	674	763	679
10%	826	771	797	730	583	525	580	522	584	526
5%	780	704	719	596	474	443	472	443	476	446
1%	704	616	478	404	266	281	268	280	270	282
0% Min	632	562	176	234	**-304**	**-353**	32	100	166	223

Table H – 3 Longitudinal individual weighting for SHP_II wave 4 (2007). For each set of weights, we calculated all the statistics on the sub-group of weights > 0. We also made a comparison with the same statistics for the weightings of the previous year.

PTI_NRQI and PL_NRQI corrected weights for non-response to the individual questionnaire
POIDTRIN final individual cross-sectional weight
POIDLONG final individual longitudinal weight

Name of set of weights	PTI_NRQI 06	**PTI_NRQI 07**	POIDTRIN 06	**POIDTRIN 07**	PL_NRQI 06	**PL_NRQI 07**	POIDLONG 06	**POIDLONG 07**
N	2568	2347	2568	2347	2453	2185	2453	2185
Total of weights	7158158	7250411	6196348	6280768	6953021	6956714	6080866	6080866
Average	2787	3089	2413	2676	2834	3184	2479	2783
Standard deviation	1125	1039	1295	1508	1062	973	1203	1284
Variance	1265490	1080017	1677224	2274016	1127007	947179	1448394	1648812
Range	7794	8058	12451	12050	6894	8287	11242	10637
Interquartile interval	1429	1211	1331	1346	1427	1284	1370	1309
Max/min ratio	12	21	25	48	8	9	18	18
Ratio P99 /P1	6	5	9	10	5	5	7	6
Ratio P95 / P5	3	3	5	4	3	3	4	4
Quantiles								
100% max	8503	8464	12966	12307	7874	9267	11886	11265
99%	6420	6010	6917	8642	6101	6090	6537	7289
95%	4915	5061	5017	5741	4795	4993	4979	5332
90%	4250	4595	4143	4390	4261	4488	4052	4392
75% Q3	3387	3634	2877	3078	3456	3802	2990	3248
50% median	2565	2879	2027	2232	2612	2926	2179	2476
25% Q1	1958	2423	1546	1732	2029	2517	1620	1939
10%	1596	1987	1230	1468	1703	2175	1305	1569
5%	1434	1596	1104	1283	1566	1980	1191	1451
1%	1028	1181	792	835	1260	1203	958	1239
0% min	710	406	515	257	980	980	644	629

Table H – 4 Cross-sectional household weighting SHP_II wave 4 (2007). For each set of weights, we calculated all the statistics on the sub-group of weights > 0. We also made a comparison with the same statistics for the weightings of the previous year.

POIDINIT2 initial corrected weights in wave 1 for non-reponse to grid questionnaire
P_NRGRIL weight adjusted for non-response to the grid questionnaire
PTM_NRQM weight adjusted for non-response to the household questionnaire
POIDTRME_INT final cross-sectional household weight

Name of set of weights	POIDINIT2 06	**POIDINIT2 07**	P_NRGRIL 06	**P_NRGRIL 07**	PTM_NRQM 06	**PTM_NRQM 07**	POIDTRME_INT 06	**POIDTRME_INT 07**	**POIDTRME_INT_WIN 07**
N	1660	1439	1431	1246	1684	1489	1684	1489	1489
Total of weights	2094703	1814983	3023814	3030473	3543600	3574007	3183038	3199947	3208510
Average	1262	1261	2113	2432	2104	2400	1890	2149	2155
Standard deviation	278	277	638	570	749	700	856	982	963
Variance	77299	76856	407126	324461	560613	490027	732642	965169	928199
Range	1466	1466	3745	3401	6118	6667	6422	7900	5290
Interquartile interval	252	249	775	550	817	647	1007	1165	1171
Max/min ratio	3	3	5	3	13	20	30	168	13
Ratio P99 /P1	2	2	3	3	6	5	7	8	8
Ratio P95 / P5	2	2	2	2	3	3	4	4	4
Quantiles									
100% max	2436	2436	4740	4823	6627	7010	6644	7947	5723
99%	2147	2147	4078	4242	4375	4480	4735	5409	5434
95%	1832	1815	3441	3560	3555	3662	3500	4188	4208
90%	1749	1749	3030	3307	3141	3332	2990	3370	3385
75% Q3	1339	1336	2406	2608	2443	2665	2281	2643	2655
50% median	1173	1173	1991	2292	1971	2279	1713	1961	1970
25% Q1	1087	1087	1631	2057	1626	2018	1275	1478	1484
10%	1015	1019	1447	1944	1352	1756	1038	1168	1174
5%	1000	1000	1347	1707	1122	1188	890	958	963
1%	980	980	1205	1531	749	990	635	638	641
0% min	970	970	995	1422	510	343	222	47	433

Table H – 5 Longitudinal individual weighting for SHP_I and SHP_II combined wave 4 (2007). For each set of weights, we calculated all the statistics on the sub-group of weights > 0. We also made a comparison with the same statistics for the weightings of the previous year.

PL2_NRQI weight corrected for non-response to the individual questionnaire
PLI_COMBPAN_REG individual weight adjusted for non-response, combined at the major region level,
PLI_COMBPAN_CAL final individual longitudinal weight

Name of set of weights	PL2_NRQI 06	**PL2_NRQI 07**	PLI_COMBPAN _REG 06	**PLI_COMBPAN _REG 07**	PLI_COMBPAN _CAL 06	**PLI_COMBPAN _CAL 07**
N	6244	6364	6244	6364	6244	6364
Total of weights	13'146'565	13'137'147	6774804	6749091	6080866	6080866
Average	2105	2064	1085	1061	974	956
Standard deviation	1022	1116	520	525	641	676
Variance	1044813	1245897	270481	275689	411087	456374
Range	7399	8791	5839	6060	9870	8046
Interquartile interval	1236	1529	538	496	513	490
Max/min ratio	16	19	21	20	49	39
Ratio P99 /P1	7	8	7	8	10	12
Ratio P95 / P5	4	5	3	3	5	5
Quantiles						
100% max	7874	9267	6130	6372	10075	6364
99%	5417	5324	3095	3214	3417	6080866
95%	4120	4283	1975	1917	2148	956
90%	3553	3660	1660	1601	1609	676
75% Q3	2579	2697	1283	1232	1110	456374
50% median	1873	1756	963	936	805	8046
25% Q1	1343	1168	745	737	597	490
10%	1045	947	624	613	486	39
5%	972	852	568	552	428	12
1%	768	669	440	411	341	5
0% min	476	476	291	311	205	6364

Table H – 6 Longitudinal individual weighting for SHP_I and SHP_II combined wave 4 (2007). For each set of weights, we calculated all the statistics on the sub-group of weights > 0. We also made a comparison with the same statistics for the weightings of the previous year.

PTI_NRQI_COMB	weight corrected for non-response to the individual questionnaire
PTI_COMBPAN_REG	individual weight adjusted for non-response, combined at the major region level,
PTI_COMBREG_CAL	cross-sectional individual weight after calibration
PTI_COMBREG_CAL _WIN	final cross-sectional individual weight (Winsorised at the 1st and 99th percentiles)

Name of set of weights	PTI_NRQI _COMB 06	**PTI_NRQI _COMB 07**	PTI_COMBPAN _REG 06	**PTI_COMBPAN _REG 07**	PTI_COMBREG _CAL 06	**PTI_COMBREG _CAL 07**	**PTI_COMBREG _CAL_WIN 07**
N	6656	6973	6656	6973	6656	6973	6973
Total of weights	13'620'632	13'805'235	6744736	7101803	6196349	6280768	6280768
Average	2046	1980	1013	1018	931	901	901
Standard deviation	1056	1127	430	533	575	714	639
Variance	1115407	1269486	184788	283894	331128	509698	407930
Range	8311	8204	3487	7168	6104	9124	3893
Interquartile interval	1292	1513	512	455	479	441	447
Max/min ratio	44	33	30	53	73	94	16
Ratio P99 /P1	8	10	6	9	10	16	16
Ratio P95 / P5	5	5	4	4	5	6	6
Quantiles							
100% max	8503	8464	3606	7306	6189	9222	4161
99%	5550	5297	2427	3172	3152	4104	4161
95%	4116	4283	1844	1894	2085	2138	2168
90%	3491	3567	1590	1583	1613	1514	1535
75% Q3	2556	2631	1216	1161	1059	978	991
50% median	1761	1595	907	889	761	695	705
25% Q1	1264	1118	703	706	579	537	544
10%	1014	898	597	571	479	433	439
5%	896	783	526	488	420	383	388
1%	681	552	386	351	319	264	268
0% min	192	260	119	138	85	98	268

Table H – 7 cross-sectional household weighting for SHP_I and SHP_II combined wave 4 (2007). For each set of weights, we calculated all the statistics on the sub-group of weights > 0. We also made a comparison with the same statistics for the weightings of the previous year.

PTM_NRQM_COMB weight adjusted for non-response to the household questionnaire

PTM_COMBPAN_REG cross-sectional household weight adjusted for non-response, combined at the major region level,

PTM_COMBREG_CAL2 cross-sectional household weight after 2nd calibration (=final weight in 2007)

PTM_COMBREG_CAL2_WIN final cross-sectional household weight (Winsorised at the 1st and 99th percentiles). Winsorisation is carried out on the set of weights PTM_COMBREG_CAL2 on the file at individual level. This is why the numbers given here, calculated at household level, do not tally exactly when the total, the P1, P99 of the weights _INT are compared with the weights _CAL2_WIN.

Name of set of weights	PTM_NRQM _COMB 06	**PTM_NRQM _COMB 07**	PTM_COMBPAN _REG 06	**PTM_COMBPAN _REG 07**	PTM_COMBREG _CAL2 06	**PTM_COMBREG _CAL 07**	**PTM_COMBREG _CAL2 07**	**PTM_COMBREG _CAL2_WIN 07**
N	4221	4306	4221	4306	4221	10651	4306	4306
Total weight	6683724	6775328	3305120	3411552	3111109	7369590	3163047	3166716
Average	1583	1573	783	792	737	692	735	735
Standard deviation	705	794	267	313	389	504	480	439
Variance	497383	630480	71404	97965	151108	254094	230136	192994
Range	6451	6776	2529	4122	3646	5598	5407	2684
Interquartile interv.	862	1124	326	302	399	384	389	393
Max/min ratio	38	30	26	35	90		138	17
Ratio P99 /P1	7	9	5	7	10	18	13	13
Ratio P95 / P5	4	5	3	3	5	7	6	6
Quantiles								
100% max	6627	7010	2632	4242	3687	5494	5447	2853
99%	3945	4017	1588	1980	2135	2838	2605	2633
95%	2964	3075	1283	1297	1500	1694	1635	1652
90%	2458	2608	1140	1158	1214	1167	1199	1212
75% Q3	1927	2090	926	909	882	799	852	861
50% median	1475	1307	738	738	640	548	605	612
25% Q1	1065	965	600	608	483	415	463	468
10%	841	791	494	497	377	316	352	356
5%	778	674	441	389	312	258	297	300
1%	545	447	298	284	222	154	200	202
0% min	176	234	103	121	41	**-104**	40	168

APPENDIX I – Response rates

Diagram I – 1 SHP_I wave 9.

Gross sample	**3976**	**100%**				
Eligibility status unknown (nobody answers, address contacted but unknown eligibility status)	68	2%				
Out of scope of sample (wrong number, modem, fax, deceased)	414	10%				
Ineligible addresses (secondary residence, nursing home, institution, medico-social establishment, prison, professional addresses)	18	0%				
Eligible addresses (eligibility status known)	3475	88%		**100%**		
	Current grid appointment		95	3%		
	Refusal, problems at grid level		483	14%		
	Grids completed		**2898**	**83%**		**100%**
			Refusal, problems at household level		76	2%
			Household grid + quest. completed		**2822**	**98%**

Individual level

Number of individuals in households eligible for individual questionnaire		5974	100%
	Current individual appointment	8	0%
	Refusal, problems at individual level	1332	22%
	Individual questionnaires completed	**4634**	**78%**

Diagram I – 2 SHP_II wave 4.

Gross sample	**2037**	**100%**
Eligibility status unknown (nobody answers, address contacted but unknown eligibility status)	34	2%
Out of scope of sample (wrong number, modem, fax, deceased)	98	5%
Ineligible addresses (secondary residence, nursing home, institution, medico-social establishment, prison, professional addresses)	5	0%
Eligible addresses (eligibility status known)	1900	93%

		100%
Current grid appointment	53	3%
Refusal, problems at grid level	298	16%
Grids completed	**1549**	**81%**

		100%
Refusal, problems at household level	53	3%
Household grid + quest. completed	**1495**	**97%**

Individual level

Number of individuals in households eligible for individual questionnaire	3108	100%
Current individual appointment	11	0%
Refusal, problems at individual level	746	24%
Individual questionnaires completed	**2351**	**76%**

APPENDIX J – Algorithm for belonging to target population

Algorithm which is referred to in paragraph 12.2:

1)…for individuals from SHP_I
- if OSM[41]=1.2 then PRES_P1=1
- if OSM=1.2 and CR_V6[42]≠1 then PRES_P2=1
- if OSM=1.2 and CR_V6=1 (cross-sectionally out of scope in 2004) then PRES_P2=2

2)…for individuals from SHP_II
- if OSM=1.2 then PRES_P2=1
- if OSM=1.2 and CR_V1=1 (cross-sectionally out of scope in 2004) then PRES_P2=2

For cases where, after 1) and 2), PRES_P1 or PRES_2 are still missing (applied to individuals from SHP_I as well as those from SHP_II):
- if HAB_CH07[43]=-5 then PRES_P1=1 (have always lived in Switzerland)
- if 0 < HAB_CH07 < 1999 then PRES_P1=1 (already living in Switzerland in 1999)
- if 0 < HAB_CH07 > 1999 then PRES_P1=1 (not living in Switzerland in 1999)
- if HAB_CH07=-5 then PRES_P2=1 (have always lived in Switzerland)
- if 0 < HAB_CH07 < 2004 then PRES_P2=1 (already living in Switzerland in 2004)
- If HAB_CH07 > 2004 then PRES_P2=2 (not living in Switzerland in 2004)
- if year of birth > 2004 then PRES_P2=2

For the remaining non-resolved cases, the eligibility status of the household was given[44]. Lastly, we allowed the remaining 20 cases not-covered to be eligible.

[41] Original Sample Member, see paragraph 3.3.2.
[42] Code response to wave 6, see paragraph 6.3.
[43] HAB_CH07: variable from individual questionnaire indicating since when the person has lived in Switzerland if this was not always the case.
[44] A household is said to be eligible if at least one member of the household is eligible.

Methodenberichte des Dienstes Statistische Methoden des BFS
Rapports de méthodes du Service de méthodes statistiques de l'OFS
Methodology reports published by the SFSO's Statistical Methods Unit

Graf, E. (2009). Weightings of the Swiss Household Panel: SHP_I wave 9, SHP_II wave 4, SHP_I et SHP_II combined. Order number: 338-0058

Graf, E. (2009). Pondérations du Panel Suisse de Ménages: PSM_I vague 9, PSM_II vague 4, PSM_I et PSM_II combinés. Numéro de commande: 338-0057-05

Qualité, L., Tillé, Y. (2009). Estimation de la précision d'évolutions dans l'enquête sur la valeur ajoutée. Numéro de commande: 338-0056

Renaud, A., Panchard, C. et Potterat, J. (2008). Statistique de l'emploi. Révision 2007 : méthodes d'estimation. Numéro de commande: 338-0055

Graf, E. (2008). Pondérations du PSM: PSM_I vague 8, PSM_II vague 3, PSM_I et PSM_II combinés. Numéro de commande: 338-0054

Andrade, B., Graf, M. (2008). Enquête suisse sur la structure des salaires 2006. Aspects méthodologiques du modèle des salaires "Salarium".
Numéro de commande: 338-0053

Renaud, A. (2008). Statistique de l'emploi. Révision 2007 : cadre de sondage et échantillonnage. Numéro de commande: 338-0052

Graf, E. (2008). Pondérations du SILC pilote: SILC_I vague 2, SILC_II vague 1, SILC_I et SILC_II combinés. Numéro de commande: 338-0051

Kilchmann, D. (2008). Statistik der sozialmedizinischen Institutionen 1999-2004 und Krankenhausstatistik 1999-2002. Einsetzungen für fehlende Daten. Bestellnummer: 338-0050

Renaud, A. (2008). Technologies de l'information et de la communication. Estimations sur la base de la statistique de la valeur ajoutée. Numéro de commande: 338-0049

Assoulin, D. (2007). Wertschöpfungsstatistik. Einsetzungsversuche für fehlende Antworten grosser Unternehmen. Bestellnummer: 338-0048

Kilchmann, D. (2007). Beherbergungsstatistik Campingplätze.Stichprobenrahmen und Schätzverfahren 2005/06. Bestellnummer: 338-0047

Gabler, S., Häder, S. (2007). Haushalts- und Personenerhebungen. Machbarkeit von Random Digit Dialing in der Schweiz. Bestellnummer: 338-0046

Ferrez, J., Graf, M. (2007). Enquête suisse sur la structure des salaires. Programmes R pour l'intervalle de confiance de la médiane. Numéro de commande: 338-0045

Renaud, A. (2007). Harmonisation de la scolarité obligatoire en Suisse (HarmoS). Design général de l'enquête et échantillon des écoles. Numéro de commande: 338-0044

Potterat, J. (2007). Betriebszählung 2005. Statistische Methoden zur Schätzung der provisorischen Ergebnisse. Bestellnummer: 338-0043

Hulliger, B. (2006). Umweltschutzausgaben der Unternehmen 2003, Stichprobenplan, Datenaufbereitung und Schätzverfahren. Bestellnummer: 338-0042

Renfer, J.-P. (2006). Enquête sur les chiffres d'affaires du commerce de détail. Plan d'échantillonnage et méthodes d'estimation. Numéro de commande: 338-0041

Salamin, P.-A. (2006). Statistique de l'aide sociale dans le domaine de l'asile. Plan de sondage et extrapolations pour l'enquête pilote 2005. Numéro de commande: 338-0040

Renaud, A. (2006). Statistique suisse des bénéficiaires de l'aide sociale. Pondération des communes 2004. Numéro de commande: 338-0039

Graf, M. (2006). Swiss Earnings Structure Survey 2002-2004. Compositional data in a stratified two-stage sample: Analysis and precision assessment of wage components. Order number: 338-0038

Potterat, J. (2006). Pensionskassenstatistik 2004. Statistische Methoden zur Schätzung der provisorischen Ergebnisse. Bestellnummer: 338-0037

Potterat, J. (2006). Kosten und Nutzen der Berufsbildung aus Sicht der Betriebe im Jahr 2004. Stichprobenplan, Gewichtung und Schätzverfahren. Bestellnummer: 338-0036

Kilchmann, D. (2006). Vierteljährliche Wohnbaustatistik. Stichprobenplan, statistische Datenaufarbeitung und Schätzverfahren 2005. Bestellnummer: 338-0035

Kilchmann, D. (2006). Erhebung über Forschung und Entwicklung in der schweizerischen Privatwirtschaft 2004. Bereinigung der Stichprobe, Ersatz fehlender Werte und Schätzverfahren. Bestellnummer: 338-0034

Kilchmann, D., Eichenberger, P., Potterat, J. (2005). Volkszählung 2000. Statistische Einsetzungsverfahren Band 2. Bestellnummer: 338-0033

Kilchmann, D., Eichenberger, P., Potterat, J. (2005). Volkszählung 2000. Statistische Einsetzungsverfahren Band 1. Bestellnummer: 338-0032

Graf, M., Matei, A. (2005). Enquête suisse sur la structure des salaires 2002. La précision du salaire brut standardisé médian. Numéro de commande: 338-0031

Graf, E., Renfer, J.-P. (2005). Enquête suisse sur la santé 2002. Plan d'échantillonnage, pondération et estimation de la précision. Numéro de commande: 338-0030

Potterat, J. (2005). Mietpreis-Strukturerhebung 2003. Gewichtung und Schätzverfahren. Bestellnummer: 338-0029

Potterat, J. (2005). Landwirtschaftliche Betriebszählung 2003. Schätzverfahren für die Zusatzerhebung. Bestellnummer: 338-0028

Renaud, A. (2004). Coverage estimation for the Swiss population census 2000. Estimation methodology and results. Order number: 338-0027

Kilchmann, D. (2004). Revision des Schweizerischen Lohnindex. Schätzmethoden der Lohnindices und deren Varianzschätzer. Bestellnummer: 338-0026

Graf, M. (2004). Enquête suisse sur la structure des salaires 2002. Plan d'échantillonnage et extrapolation pour le secteur privé. Numéro de commande: 338-0025

Renaud, A. (2004). Analyse de données d'enquêtes. Quelques méthodes et illustration avec des données de l'OFS. Numéro de commande 338-0024

Renaud, A., Potterat, J. (2004). Estimation de la couverture du recensement de la population de l'an 2000. Echantillon pour l'estimation de la sous-couverture (P-sample) et qualité du cadre de sondage des bâtiments. Numéro de commande: 338-0023

Graf, M. (2004). Fusion de données. Etude de faisabilité. Numéro de commande: 338-0022

Potterat, J. (2003). Mietpreis-Strukturerhebung 2003. Entwicklung des Stichprobenplans und Ziehung der Stichprobe. Bestellnummer: 338-0021

Potterat, J. (2003). Landwirtschaftliche Betriebszählung 2003. Stichprobenplan der Zusatzerhebung. Bestellnummer: 338-0020

Renaud, A. (2003). Estimation de la couverture du recensement de la population de l'an 2000. Echantillon pour l'estimation de la sur-couverture (E-sample). Numéro de commande: 338-0019

Hulliger, B. (2003). Erhebung über Forschung und Entwicklung in der schweizerischen Privatwirtschaft 2000. Bereinigung der Stichprobe, Ersatz fehlender Werte und Schätzverfahren. Bestellnummer: 338-0018

Renfer, J.-P. (2003). Enquête 2000 sur la recherche et le développement dans l'économie privée en Suisse. Plan d'échantillonnage. Numéro de commande: 338-0017

Potterat, J. (2003). Kosten und Nutzen der Berufsbildung aus Sicht der Betriebe. Schätzverfahren. Bestellnummer:
338-0016

Graf, M., Matei, A. (2003). Stratégie de choix des modèles de désaisonnalisation. Application aux séries de l'emploi total. Numéro de commande : 338-0015

Potterat, J., Salamin, P.A. (2002). Betriebszählung 2001. Methoden für die Datenbereinigung. Bestellnummer:
338-0014

Renaud, A. (2002). Programme international pour le suivi des acquis des élèves (PISA). Plans d'échantillonnage pour PISA 2000 en Suisse. Numéro de commande: 338-0013

Renfer, J.-P. (2002). Enquête 2001 sur les coûts et l'utilité de la formation des apprentis du point de vue des établissements. Plan d'échantillonnage. Numéro de commande: 338-0012

Potterat, J., Salamin, P.A. (2002). Betriebszählung 2001. Stichprobenplan und Schätzverfahren für die provisorischen Ergebnisse. Bestellnummer: 338-0011

Graf, M. (2002). Enquête suisse sur la structure des salaires 2000. Plan d'échantillonnage, pondération et méthode d'estimation pour le secteur privé. Numéro de commande: 338-0010

Renaud, A., Eichenberger P. (2002). Estimation de la couverture du recensement de la population de l'an 2000. Procédure d'enquête et plan d'échantillonnage de l'enquête de couverture. Numéro de commande: 338-0009

Kilchmann, D., Hulliger, B. (2002). Stichprobenplan für die Obstbaumzählung 2001. Bestellnummer: 338-0008

Graf, M. (2002). Passage du concept établissement au concept entreprise. Numéro de commande: 338-0007

Salamin, P.A. (2001). La technique de la double enquête pour la statistique du transport routier de marchandise. Numéro de commande: 338-0006

Peters, R., Renfer, J.-P. et Hulliger, B. (2001). Statistique de la valeur ajoutée 1997-1998. Procédure d'extrapolation des données. Numéro de commande: 338-0005

Potterat, J., Hulliger, B. (2001). Schätzung der Sägereiproduktion mit der Sägerei-Erhebung PAUL. Bestellnummer: 338-0004

Graf, M. (2001). Désaisonnalisation. Aspects méthodologiques et application à la statistique de l'emploi. Numéro de commande: 338-0003

Hüsler, J., Müller, S. (2001). Schlussbericht Betriebszählung 1995 (BZ 95), Mehrfach imputierte Umsatzzahlen. Bestellnummer: 338-0002

Renaud, A. (2001). Statistique suisse des bénéficiaires de l'aide sociale. Plan d'échantillonnage des communes. Numéro de commande: 338-0001

Hulliger, B., Eichenberger, P. (2000). Stichprobenregister für Haushalterhebungen: Umstellung auf Telefonnummern ohne Namen und Adressen, Abläufe für Erstellung und Stichprobenziehung. Bestellnummer: 338-0000

de Rossi, F.-X. (1998). Méthodes statistiques pour le compte routier suisse.

Hulliger, B., Kassab, M. (1998). Evaluation of Estimation Methods for the Survey on Environment Protection Expenditures of Swiss Communes.

Salamin, P.A. (1998). Etablissement d'une clef de passage pondérée entre l'ancienne (NGAE 85) et la nouvelle nomenclature (NOGA 95) générale des activités économiques.

Peters, R. (1998). Extrapolation des données de l'enquête de structure sur les loyers.

Bender, A., Hulliger, B. (1997). Enquête suisse sur la population active: rapport de pondération pour 1996.

Salamin, P.A. (1997). Evaluation de la Statistique de l'emploi.

Peters, R. (1997). Etablissement du plan d'échantillonnage pour l'enquête 1996 sur la recherche et le développement dans l'économie privée en Suisse.

Peters, R. (1997). Enquête 1996 sur la structure des salaires en Suisse: établissement du plan d'échantillonnage.

Peters, R. (1996). Pondération des données de l'enquête sur la famille en Suisse.

Comment, T., Hulliger, B., Ries, A. (1996). Gewichtungsverfahren für die Schweizerische Arbeitskräfteerhebung (1991-1995).

Hulliger, B. (1996). Haushalterhebung Familie 1994: Stichprobenplan, Stichprobenziehung und Reservestichproben.

Peters, R., Hulliger, B. (1996). Schätzverfahren für die Lohnstruktur-Erhebung 1994 / Procédure d'estimation pour l'enquête de 1994 sur la structure des salaires.

Peters, R. (1996). Schéma de pondération des indices PAUL.

Hulliger, B., Peters, R. (1996). Enquête sur le comportement de la population suisse en matière de transport en 1994: plan d'échantillonnage et pondération.

Hulliger, B. (1996). Gütertransportstatistik 1993: Schätzverfahren mit Kompensation der Antwortausfälle.

Salamin, P.A. (1995). Estimation des flux pour le module II des comptes globaux du marché de travail.

Peters, R. (1995). Enquête de structure sur les loyers: établissement d'un plan d'échantillonnage stratifié.

Hulliger, B. (1995). Konjunkturelle Mietpreiserhebung: Stichprobenplan und Schätzverfahren.

Schwendener, P. (1995). Verbrauchserhebung 1990 - Vertrauensintervalle.

Peters, R., Hulliger, B. (1994). La technique de pondération des données: application à l'enquête suisse sur la santé.

Hulliger, B., Peters, R. (1994). Enquête sur la structure des salaires en Suisse: stratégie d'échantillonnage pour le secteur privé.

Publikationsprogramm BFS

Das Bundesamt für Statistik (BFS) hat – als zentrale Statistikstelle des Bundes – die Aufgabe, statistische Informationen breiten Benutzerkreisen zur Verfügung zu stellen.

Die Verbreitung der statistischen Information geschieht gegliedert nach Fachbereichen (vgl. Umschlagseite 2) und mit verschiedenen Mitteln

FSO Publications

The Federal Statistical Office (FSO) is the central and official purveyor of statistical information to the Swiss Government. It is officially mandated to supply this information to a wide range of users.
This statistical data is organized and disseminated on the basis of a subject-matter classification (see inside coverpage).

Diffusionsmittel	*Kontakt* *Contact*	*Distribution medium*
Individuelle Auskünfte	+41 (0)32 713 60 11 info@bfs.admin.ch	Individual information
Das BFS im Internet	www.statistik.admin.ch	The FSO on the Internet
Medienmitteilungen zur raschen Information der Öffentlichkeit über die neusten Ergebnisse	www.news-stat.admin.ch	to the latest results to the latest results
Publikationen zur vertieften Information (zum Teil auch als Diskette/CD-Rom)	+41 (0)32 713 60 60 order@bfs.admin.ch	Purchase of printed and electronic publications and CD-Rom
Online Datenrecherche (Datenbanken)	www.statdb.bfs.admin.ch	Online data search

Nähere Angaben zu den verschiedenen Diffusionsmitteln im Internet unter der Adresse www.statistik.admin.ch→ Dienstleistungen→Publikationen Statistik Schweiz.

For information regarding the various means of distribution over the Internet, visit: www.statistics.admin.ch→Services→Swiss statistical publications.

Methodenberichte des Dienstes Statistische Methoden
Rapports de méthodes du Service de méthodes statistiques
Methodology Reports by the Statistical Methods Unit

Die Methodenberichte beschreiben die mathematischen und statistischen Methoden, die den Resultaten und Analysen der öffentlichen Statistik zu Grunde liegen. Sie enthalten ausserdem die Evaluation und Entwicklung von neuen Methoden im Hinblick auf eine zukünftige Anwendung. Diese Publikationen sollen einerseits die verwendeten Methoden dokumentieren, um Transparenz und Wissenschaftlichkeit sicher zu stellen, und sie sollen andererseits die Zusammenarbeit mit den Hochschulen und der Wissenschaft fördern.

Zur Illustration der beschriebenen mathematischen Konzepte, werden im Bericht numerische Resultate aufgeführt. Diese sind allerdings nicht als offizielle Resultate der betreffenden Erhebungen zu verstehen. Ebenfalls können die tatsächlich angewendeten Methoden leicht von den hier beschriebenen abweichen.

Die Methodenberichte sind auf der Internetseite des BFS in elektronischer Form verfügbar.

The methodology reports describe the mathematical and statistical methodology which forms the basis of the results and analyses of public statistics. They also present the evaluation and development of new methods with regard to their use in the future.
These publications serve to document the methodology used or planned in order to ensure transparency and scientific integrity. They also encourage collaboration with the universities and world of science.

Although the numerical results presented in the methodology reports are used to illustrate the mathematical concepts described, they are not the official results of the surveys concerned. Similarly, the methodology actually used may differ slightly from that described in these reports.

The methodology reports are available in electronic form at the FSO internet site.

The document has the following structure: the way data were collected is briefly described for each of the panels, then, after having presented the response rates achieved, a description is given of the different concepts related to the weighting. Next, the steps of the weighting are presented and then described one by one for each of the panels. A distinction is made between steps common to all weightings and steps peculiar to each one. Weightings at household level and individual level are also dealt with separately. The specific features of cross-sectional weightings as compared to longitudinal weightings are also noted. The final sections deal with the combined weightings of SHP_I wave 9 and SHP_II wave 4. At the end of the report, the interested reader will find a collection of appendices that cover in detail some of the steps only briefly mentioned in the main text.

Order number:
338-0058

Orders: +41
(0)32 713 60 60 Fax: (0)32 713 60 61 Email: order@bfs.admin.ch

Price:
Free of charge ISBN 978-3-303-00425-8