10 Changes from the first to the second wave of the HFCS

10.1 Introduction

The HFCS has now been conducted twice in Austria. The field phase of the first wave lasted from September 2010 to May 2011. In the second wave, the fieldwork was carried out between June 2014 and February 2015. The second wave drew heavily on the experience gained in the first wave. After all, HFCS data are used for a broad range of research and by all relevant institutions in Austria.

This chapter offers a short, but comprehensive insight into the changes made between the two HFCS waves for those who already have experience with evaluating data from the first wave in Austria. Furthermore, this chapter provides the foundation for evaluations based on both waves of the HFCS in Austria, which require an understanding of how the two survey waves differ.

The structure of this chapter mirrors the structure of the documentation as a whole: Following an overview of the key changes made to the questionnaire (section 10.2) and to interviewer training and selection (section 10.3), we discuss editing measures (section 10.4), the multiple imputation process (section 10.5) and the sampling design (section 10.6). The final two sections deal with the construction of survey weights (section 10.7) and replicate weights (section 10.8). The user guide (see chapter 9) is not discussed here, since it was left broadly unchanged. The chapter is completed with the concluding remarks.

10.2 Questionnaire

10.2.1 Recording of household matrix

The second wave benefited from a more efficient, and hence faster, technical process for recording the demographics for all household members and identifying the financially knowledgeable person (and if necessary reference person).

Once interviewers had established the number of household members, a matrix appeared which allowed them to record the basic data for all people in the household: name (required for referencing while completing the questionnaire), age, gender, relationship with the reference person and financial affiliation with the household. This information was subsequently shown in the form of a list, to enable respondents to cross-check the data, with a view to making adjustments, deleting a person or adding another person. The interview did not start until the basic data for all household members had been confirmed. Compared with the first wave, this means that above all the questions about age and gender were moved from the first chapter of the survey to the household matrix, which helped save time.

10.2.2 Lists of predefined ranges

As outlined in section 2.6.2, all questions about amounts of money were asked in loops. If respondents were unable or unwilling to indicate specific amounts or individual ranges, they were asked to choose a range from a list of predefined ranges. While the same list of predefined ranges was used for all questions in the first wave, three different lists were used in the second wave, depending on the question at hand (see table 19).[1] The first column shows the list of ranges used in the first wave, and the other three columns the lists of ranges (A, B and C) used in the second wave.

[1] *See the online appendix for the showcards used during the interviews.*

Table 19

Different lists of predefined ranges in the first wave and the second wave (in EUR)

First wave (HFCS Austria 2010)	Second wave (HFCS Austria 2014)		
List of predefined ranges	List A	List B	List C
EUR			
A 1 – below 101	A 1 – below 101	A 1 – below 10,001	A 1 – below 1,001
B 101 – below 501	B 101 – below 201	B 10,001 – below 50,001	B 1,001 – below 2,501
C 501 – below 1,001	C 201 – below 301	C 50,001 – below 75,001	C 2,501 – below 5,001
D 1,001 – below 2,501	D 301 – below 401	D 75,001 – below 100,001	D 5,001 – below 7,501
E 2,501 – below 5,001	E 401 – below 501	E 100,001 – below 150,001	E 7,501 – below 10,001
F 5,001 – below 7,501	F 501 – below 751	F 150,001 – below 200,001	F 10,001 – below 15,001
G 7,501 – below 10,001	G 751 – below 1,001	G 200,001 – below 300,001	G 15,001 – below 20,001
H 10,001 – below 25,001	H 1,001 – below 1,501	H 300,001 – below 400,001	H 20,001 – below 25,001
I 25,001 – below 50,001	I 1,501 – below 2,001	I 400,001 – below 500,001	I 25,001 – below 30,001
J 50,001 – below 75,001	J 2,001 – below 3,001	J 500,001 – below 750,001	J 30,001 – below 35,001
K 75,001 – below 100,001	K 3,001 – below 5,001	K 750,001 – 1 million	K 35,001 – below 40,001
L 100,001 – below 250,001	L 5,001 – below 7,501	L more than 1 million – 3 million	L 40,001 – below 50,001
M 250,001 – below 500,001	M 7,501 – below 10,001	M more than 3 million – 5 million	M 50,001 – below 75,001
N 500,001 – 1 million	N 10,001 – below 25,001	N more than 5 million – 10 million	N 75,001 – below 100,001
O more than 1 million – 5 million	O 25,001 – below 50,001	O more than 10 million	O 100,001 – below 200,001
P more than 5 million – 10 million	P more than 50,000		P 200,001 – below 300,001
Q more than 10 million – 25 million			Q 300,001 – below 500,001
R more than 25 million – 50 million			R 500,001 – 1 million
S more than 50 million – 100 million			S more than 1 million
T more than 100 million			

Source: HFCS Austria 2010 und 2014, OeNB.

Table 20

Unweighted percentiles of selected variables in the HFCS Austria 2010 wave

Percentiles	Food-at-home consumption[1]	Current value of main residence[2]	Gross income from dependent employment[3]
	EUR		
P10	170	80,000	6,400
P20	200	113,000	11,100
P30	250	148,000	14,400
P40	300	169,000	17,500
P50	350	200,000	20,200
P60	400	231,000	24,000
P70	450	275,000	28,600
P80	500	342,000	34,800
P90	600	485,000	45,500

Source: HFCS Austria 2010, OeNB.

[1] Rounded to the nearest EUR 10.
[2] Rounded to the nearest EUR 1,000.
[3] Rounded to the nearest EUR 100.

The new lists were created based on the (unweighted) empirical distribution of the euro measures recorded for the corresponding variables.[2] See table 20[3] for the distribution of amount spent on food at home (as an example of list A), the distribution of the value of main residences at the time of the interview (as an example of list B) and the distribution of gross income from dependent employment (as an example of list C).

With some exceptions, all empirically observed distributions were found to fall into very few (mostly one or two) ranges from the list of ranges used in the first wave. The switch to three different lists of ranges in the second wave allowed us to capture a much more accurate range in which a given amount

[2] *The average of all implicates was used.*

[3] *This table was first published in Lindner et al. (2014).*

falls. For instance, while in the first wave the values for the household main residence at the time of the interview were essentially spread over three ranges from the list of ranges, they are now spread over six ranges. Conversely, those ranges that were less relevant for a particular question, for instance values of below EUR 10,000 for real estate property, were merged into one range. Respondents were also asked to confirm all predefined ranges that they had selected, as well as all amounts or individual ranges that they had specified (see section 2.6.2).

List A was used for questions about consumption expenditure and loan repayments. List B was used for questions related to properties and investment in self-employment businesses, and list C was typically used for outstanding loans and income. Questions about financial assets were either allocated to list A or list C, depending on the distribution of assets as observed in the first wave of the survey.[4] The predefined ranges referred to amounts in euro only.

10.2.3 Recording households owning a farm

The first wave of the HFCS showed that farmers found it particularly challenging to break down their assets so that they would fit the structure defined by the HFCS. In the first HFCS wave, support in this respect was limited to a few notes in the questionnaire and the training of interviewers. For the second wave, the questionnaire was improved for the group of farmers as outlined below (and in section 2.6.3).

Before the interview started, respondents were classified by the interviewer as running an "Agricultural business" or as running "No agricultural business." The classification was straightforward in all but a few cases. Even if interviewers misclassified a respondent, they still recorded all the relevant information.

Specifically, extra information was recorded on farmers as follows:

- Was it possible to differentiate housing assets (i.e. the household main residence) from business assets? [in the questionnaire chapter on the household main residence]
- If not, what percentage of the recorded value did respondents allocate to their main residence? [in the questionnaire chapter on the household main residence]
- Does the value recorded for investment in a self-employment business include the recorded main residence? [in the questionnaire chapter on investment in a self-employment business]

Moreover, for questions relating to the value of the main residence, the yes/no question on properties other than household main residence, as well as the question about investment in a self-employment business and its value, farmers received detailed instructions as to which components of their household balance sheet were to be recorded under which position.

This approach greatly facilitated in particular subsequent imputations (see also section 5.4.5) with a view to breaking down the assets into farmers' main residences and their agricultural businesses. In addition, the training of interviewers was improved to enable them to better handle interviews with farmers (see section 10.3).

[4] *See the questionnaire in the online appendix for a detailed overview of which ranges (list A, B or C) were used for which questions.*

10.2.4 Loans from relatives and friends

For the second wave, the recording of loans received from relatives and friends was adjusted in line with international standards. In the first wave, information on such liabilities had been collected within the set of questions about other uncollateralized loans. Information on private loans was collected from the question regarding which institution the loan was taken out with, which respondents could answer with "family" and "friends."

The first wave of the HFCS showed that loans among private individuals are a significant component of the household balance sheet and that less information is required to record private loans from other lenders. Therefore, the decision was made to record these two categories separately. As a result, the second wave of the HFCS in Austria utililized two loops[5] with up to three iterations, for both liabilities to "relatives and friends" and liabilities to other institutions ("other uncollateralized loans").

10.2.5 Recording inheritances and gifts

The results of the first HFCS wave (see e.g. Fessler and Schürz, 2015) show that transferring capital in the form of inheritances and gifts is a major channel through which households accumulate wealth. The HFCS addresses capital transfers in two sections. First, respondents are asked to indicate how they acquired the (part-) ownership of their main residence. Second, information about ownership transfer of all other wealth is collected using loops. In the second HFCS wave in Austria this loop is repeated up to five times (compared to three in the first wave). This was also necessary since in the loop the value of a gift/inheritance at the time of ownership transfer is collected, which makes collection of this information in summary questions difficult. However, the core dataset published by the ECB so far contains only three iterations of the loop question on inheritances and gifts.

10.2.6 Comment fields in the questionnaire

In the case of questions which proved difficult for respondents or were identified as essential in the first wave, the digital questionnaire for the second wave was expanded to include an additional box, allowing both numeric values and text to be entered. The information recorded in those boxes proved to be very relevant during the postinterview checks. Such explanations have often helped solve problems recording some information that would otherwise have required follow-up phone calls (see section 4.4.2).

10.3 Interviewers

With regard to the interviewers, the key improvements observed can be attributed to the experience interviewers gained in the first wave. About half of the interviewers employed had already worked on the HFCS in 2010/11. Moreover, the training for interviewers in the second wave was revised to accommodate the experience gained from the first wave. In particular, rather than offering separate theoretical and practical training sessions, the two components were integrated more closely. After an initial theory session, interviewers had a chance to run through some items of the questionnaire in a practical, interactive session. There-

[5] *See section 2.6.1. on the structure and navigation of loops.*

after, the training alternated between theory and practice. A mock interview session using a highly complex household owning an agricultural business as an example made sure that all aspects of the questionnaire were practiced and discussed in detail.

In the first wave interviewers who joined the interview team during fieldwork were trained in additional sessions conducted by the survey company on its own, without involving an OeNB HFCS staff member. The quality of interviews conducted by these interviewers was poorer than those of interviewers who had attended the regular training. Therefore additional training sessions by the survey company were ruled out for the second wave.

10.4 Consistency checks and editing

The consistency checks were intensified further in the second wave. On the one hand, the experience gained allowed for greater efficiency in checking, and on the other hand it was possible to increase the number of checks. In the second wave, the number of consistency checks programmed into the digital questionnaire was increased significantly.

The four-eyes principle for case-by-case reviews was retained; the answers of all households (whether interviewed successfully or not) were checked for internal consistency, grouped by interviewers. Moreover, the higher frequency with which the survey company forwarded data to the OeNB during the field phase further diminished the gap between the interviews and potential follow-up queries.

The inclusion of verbatim record fields for complex questions, as well as the expansion of the comment fields in which interviewers made postinterview comments on all households, facilitated the ex post editing measures. The verbatim records collected helped clarify numerous problems (see also chapter 4). Nonetheless, if required ex post queries were made by phone.

10.5 Multiple imputations

A major change in the HFCS imputation procedure made for the second wave was not to conduct any weighted regressions. While in the first wave weighted regressions were estimated using the final survey weights (in step 3 of the procedure), those weights were only used as model predictors in the second wave. This approach is in line with the current trend in imputation literature (see e.g. Frumento et al., 2012): The purpose of multiple imputations is to produce good forecasts of the missing values (and the corresponding degree of uncertainty); the weighting of households should not occur before the final analysis of the dataset and a general assessment of the population have been carried out. This change made for the second wave ought to help decrease the standard errors slightly, because the variance of the imputed values within every multiple imputation sample, as well as the variance of the imputed samples across the multiple imputation sample, is relatively small as a result of the nonweighting of regressions.

Further improvements of the HFCS imputation procedure are aimed at enhancing the consistency across the variables and improving the convergence as well as the evaluation of the convergence of the imputation model. We were able to increase the consistency between some quantitative variables by imputing the variables as a share of other variables (see section 5.4.5) rather than imputing each variable separately. As a case in point, outstanding consumer loans were imputed

as a share of the original size of the loan; or the size of a household's first consumer loan was imputed as a share of all loans taken out by the specific household.

The convergence of the HFCS imputation procedure was potentially improved in the second wave by increasing, from 6 to 10, the number of cycles which define how often regressions are to be re-estimated and how often imputed values are to be updated. The higher this figure is, the closer the imputed values should be to a draw of the common predictive a posteriori distribution of the variables with missing values (see section 5.4.9). As discussed in chapter 5, this relationship has been cross-checked in general in simulations, but it is yet to be underpinned with theoretical evidence (see section 5.3).

The criteria for evaluating convergence were expanded in the second wave to include not only graphical checks, but also the widely used Gelman-Rubin criterion (see e.g. Cowles and Carlin, 1996). Care is taken to keep the variance of the mean across the multiple imputation sample relatively small compared with the variance of the mean across the cycles (see section 5.4.9).

10.6 Sampling

Essentially, the sampling design was improved in three ways for the second wave of the HFCS in Austria: The gross samples were enlarged, the clusters of municipalities (other than Vienna) with more than 50,000 inhabitants were adjusted in size, and changes were made to the selection probability for PSUs.

In particular, the gross sample in Austria was increased from 4,436 to 6,308 households. As a result, it was possible to arrive at a significantly higher number of successfully interviewed households. Therefore the dataset generated by the second wave of the HFCS in Austria contains 2,997 households that were interviewed successfully (compared with 2,380 in the first wave). A larger sample should increase the precision of the estimators based on the HFCS and further facilitate the analysis of subgroups, given the larger amount of interviews that could be analyzed.

In the first wave, households were clustered to groups of 8 per PSU in Vienna (as opposed to 12 in the rest of Austria) to account for the smaller geographical distances between households in Vienna. In the second wave, this strategy was applied to all strata with municipalities of 50,000 and more inhabitants. Thus, all other major cities in Austria[6] now have a sampling design with a relatively small number of SSUs (i.e. households) per drawn PSU (i.e. enumeration district). The theoretical basis of the sampling suggests that this strategy reduces the variance of individual estimators[7] generated on the basis of the survey.

The biggest and most important change compared with the first wave is the improvement with regard to the selection probability for PSUs. In the first wave, each PSU within a stratum had the same probability of being drawn. In the second wave, the selection probability was tied to the number of households in a given PSU and was defined by the ratio of the number of households in a given PSU divided by the number of households in this stratum times the number of drawn PSUs. In the second wave the PSUs were also drawn with replacement. Some

[6] *Vienna, St. Pölten, Graz, Klagenfurt, Villach, Linz, Wels, Salzburg and Innsbruck.*

[7] *See also chapter 8 in Levy and Lemeshow (2008).*

PSUs could therefore be drawn more than once. This adjustment was made with a view to decreasing the variance of household weights (see also section 10.7).

The sampling design also changed slightly as a result of changes in the sampling frame resulting from mergers of municipalities or enumeration districts. The sampling was based on external data provided by Statistics Austria and the Austrian Post Office at the time of sampling. Following the merger of some municipalities in the time between the first and the second wave, enumeration districts underwent changes, as did the strata allocation of some municipalities. However, these changes should play a minor role in the evaluation of the data.

10.7 Construction of survey weights

The adjustments made to the sampling design also affected the calculation of the design weights. As in the first wave of the HFCS in Austria, the design weights were also based on the selection probability of households in the second wave, with the design weights being the inverse of the selection probabilities. For a detailed overview of the calculation of design weights, see section 7.2.1. The adjustments made to improve sampling and the resultant design weights influenced the distribution of the weights in the second wave.

See chart 7 and table 21[8] for a comparison of the distribution of design weights in the first and the second waves of the HFCS in Austria.

Table 21

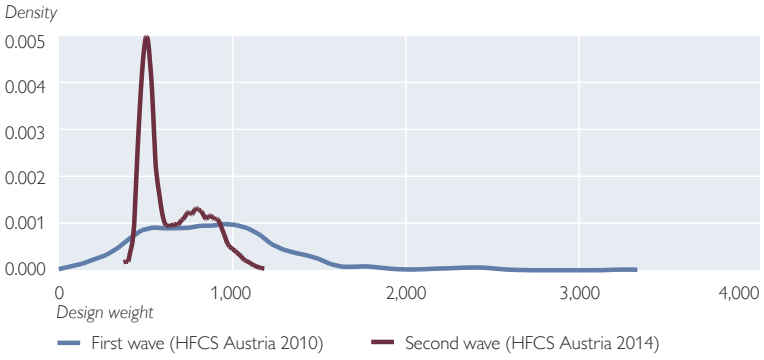
Comparison of descriptive statistics of design weights in HFCS Austria 2010 and 2014

	First wave (HFCS Austria 2010)	Second wave (HFCS Austria 2014)
Minimum	61	369
Median	857	553
Mean	884	642
Maximum	3,271	1,183
Standard deviation	434	177
Number of observations	4,436	6,308

Source: HFCS Austria 2010 und 2014, OeNB.

Chart 7

Comparison of the distribution of design weights in HFCS Austria 2010 and 2014



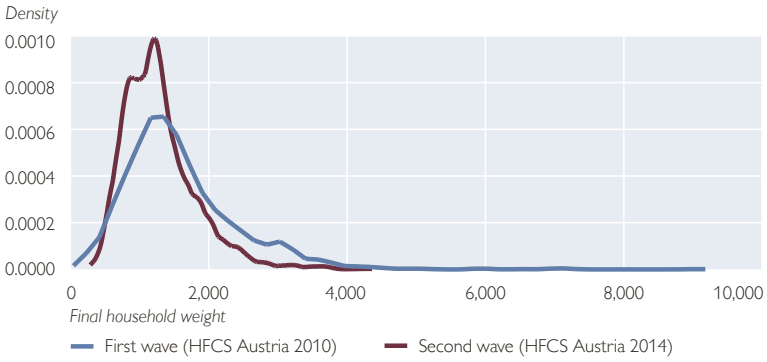
Source: HFCS Austria 2010 und 2014, OeNB.

It can be seen that the variance of the design weights has diminished significantly in the second wave due to the adjustments made to the sampling design. In particular, the standard deviation dropped from about 430 in the first wave to about 180 in the second wave. Furthermore, the range of the design weights was reduced significantly, to values between 369 and 1,183, in the second wave. As a result of the higher number of households in the gross sample, the mean of the design weights decreased from about 880 to about 640.

[8] *Both chart 7 and table 21 were published and discussed in Lindner et al. (2014).*

Chart 8

Comparison of the distribution of final household weights in HFCS Austria 2010 and 2014



Source: HFCS Austria 2010 und 2014, OeNB.

Table 22

Comparison of descriptive statistics of final household weights in HFCS Austria 2010 and 2014

	First wave (HFCS Austria 2010)	Second wave (HFCS Austria 2014)
Minimum	169	287
Median	1,429	1,207
Mean	1,586	1,289
Maximum	9,054	4,360
Standard deviation	834	527
Number of observations	2,380	2,997

Source: HFCS Austria 2010 und 2014, OeNB.

However, the adjustments are aimed at enhancing the final household weights, which goes hand in hand with changes in the design weights, where the changes appear to be more prominent. For details on the changes of the final household weights, see chart 8 and table 22.

The enhanced sampling procedure is also evident from the final household weights. In particular, the range of the final household weights was reduced from a range of about 170 to 9,050 in the first wave to about 290 to 4,360 in the second wave. It should be noted that nonresponse and poststratification adjustments cause the variability of the final household weights to be higher than that of the design weights.

The nonresponse adjustment method was also improved for the second wave. In both waves, this step in calculating the weights (see also section 7.2.3) is based on a nonlinear model designed to explain nonresponse. For the second wave, the model was expanded to include additional paradata on the interview, the place of residence and residential area, as well as data on the municipalities and political districts. Moreover, the response propensity classes applied were selected in line with the method proposed by Haziza and Beaumont (2007) so as to achieve an optimal tradeoff between bias and variance. Unlike the first wave, where households' response propensities were designed to be grouped into five classes (quintiles), the method used in the second wave is based on seven classes.

In the poststratification process (see section 7.2.4) additional information regarding main residences and second homes of the participating households allowed for a more granular poststratification of households officially registered at their main residence[9] on the basis of the relevant external data. Whereas in the first wave poststratification with the microcensus data from Statistics Austria was based on household size and municipality size, in the second wave it was based on province, household size and the tenure status of the household's main residence. This made it possible to poststratify the data, not only based on the shares of households in the individual groups, but also based on the number of households in those groups.

[9] *Households not officially registered at their main residence in the centralized residence registry were excluded from the poststratification process, i.e. in the absence of relevant external data they were given a poststratification weight of 1.*

10.8 Construction of replicate weights

Like in the first wave, the construction of replicate weights mirrors the computation of the final household weights.

The sampling error calculation model mimics the original sampling procedure as closely as possible. As in the first wave, all single-PSU strata are paired with a neighboring stratum. However, in the second wave, we made sure to pair single-PSU strata with the smaller of two or more neighboring strata (in terms of numbers of PSUs) wherever possible, in order to limit the number of aggregations required. As a result only 50 of the 185 strata had to be aggregated (compared with 81 out of 170 strata in the first wave). Aggregation is necessary as the construction of replicate weights requires at least two PSUs per stratum. With-replacement sampling of PSUs in the sampling error calculation model mirrors the sampling design and is also more in line with survey theory. However, for the purpose of the sampling calculation model we considered all PSUs within a stratum to have the same selection probability.

By enhancing the sampling procedure (and the way it is represented in the sampling error calculation model), we were able to produce more efficient replicate weights in terms of variance. Table 23 shows a comparison of the two survey waves.

Table 23

Comparison of selected replicate weights in HFCS Austria 2010 and 2014

	Minimum	Median	Mean	Maximum	Standard deviation	Number of observations
HFCS Austria 2010						
1st set of replicate weights	5	1,101	1,586	11,805	1,875	2,380
2nd set of replicate weights	4	1,089	1,586	14,345	1,941	2,380
3rd set of replicate weights	4	948	1,586	18,429	2,091	2,380
998th set of replicate weights	3	1,174	1,586	22,191	2,005	2,380
999th set of replicate weights	5	1,170	1,586	17,956	2,132	2,380
1,000th set of replicate weights	3	1,122	1,586	14,139	2,028	2,380
HFCS Austria 2014						
1st set of replicate weights	7	1,040	1,289	14,374	1,519	2,997
2nd set of replicate weights	10	989	1,289	11,418	1,472	2,997
3rd set of replicate weights	8	1,023	1,289	10,852	1,436	2,997
998th set of replicate weights	10	1,104	1,289	8,369	1,385	2,997
999th set of replicate weights	6	985	1,289	11,201	1,410	2,997
1,000th set of replicate weights	7	974	1,289	10,349	1,473	2,997

Source: HFCS Austria 2010 und 2014, OeNB.

It is evident that the standard deviation could also be reduced with regard to the replicate weights. At the same time, the range (minimum to maximum) of individual replicate weights can differ between the two waves for randomly drawn PSUs per resample, meaning there is no clear sign of a significant reduction in this respect.

10.9 Concluding remarks

This chapter provided a small but comprehensive overview of the differences between the first and second waves of the HFCS in Austria. For more detailed information on specific aspects of this publication, please see the relevant chapters or sections in this documentation.