

7 Construction of Survey Weights

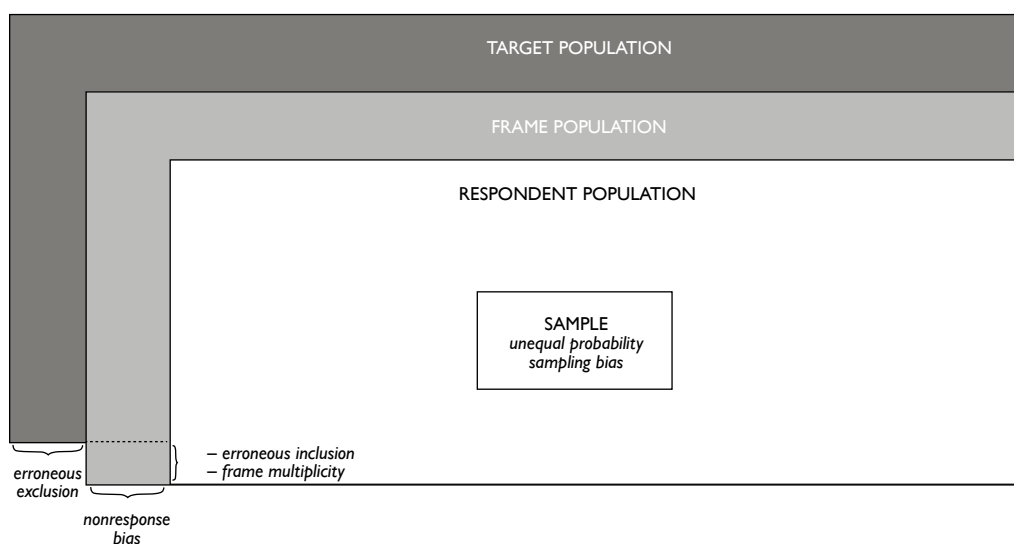
7.1 Introduction

Survey weights are usually constructed for two reasons: first, to make the sample representative of the target population and second, to reduce sampling variance.

The target population of the HFCS consists of all households in Austria, with a household being defined as a person or a group of people who live together in the same private dwelling and share expenses.¹ However, the sample may contain several types of biases that may cause a misrepresentation of this target population: the unequal probability sampling bias, the frame bias (due to an imperfect frame population), and the nonresponse bias (chart 6).

Chart 6

Misrepresentation of the Target Population in the Sample



Source: Adapted from Biemer and Christ, 2008

The unequal probability sampling bias is due to the fact that not every household has the same probability of being selected into the sample. For example, given the sampling design of the HFCS, households in small primary sampling units (PSUs) have a higher probability of being selected into the sample than households in large PSUs. Another example is the oversampling of Viennese households in the HFCS. To correct these misrepresentations, we construct design weights, which will be explained below. For further details about the sampling design of the HFCS see chapter 6.

The frame bias is caused by imperfections in the survey frame from which the sample is drawn. In the HFCS, the frame population is a list of all personal postal addresses in Austria (chapter 6). Erroneous exclusion of households could imply an imperfection with respect to the target population. In other words, we may have excluded, households without a postal address, for example, one-person households living together in residential communities and sharing an address that contains only one of these households. Then these types of households would be

¹ Some special types of households, like those living in care residences (retirees, people in need of care), prisoners, etc., are excluded from this definition. For more details on the definition of the target population, see chapter 6.

underrepresented. Another imperfection of the frame could be erroneous inclusion, that is, including addresses not belonging to households, for example, those of companies² or households in care residences. Finally, there is a third type of imperfection called frame multiplicity, which means that households may be duplicated because they have two (or more) addresses, for example multiple domiciles of commuters. Depending on its type, the frame bias can be reduced by constructing design weights³ (in case of erroneous inclusion and frame multiplicity) or by using poststratification weights (in case of erroneous exclusion). In this chapter, we explain the construction of these weights.

The nonresponse bias is caused by the fact that only a subset of the households included in the gross sample is willing to participate in the survey. Certain groups of households have a lower probability of participating in the HFCS than other groups, which is commonly found in the literature (see e.g. Kennickell and McManus, 1993). Thus, estimates for the entire frame population would be biased with respect to these group characteristics even though they are unbiased for the responding population. Using nonresponse weights can correct this bias (section 7.2.3).

Furthermore, as mentioned at the beginning, survey weights can help to reduce sampling variance, and, hence, increase the precision of the estimators. Ideally, the precision of the estimators should be improved via stratification prior to sampling. However, some variables (e.g. household size) that would have been very good for stratification and, thus, for improving the precision of our estimators, have not been available until after the sample had been drawn and the sample households had been contacted. Some of the gain in precision that would have been realized had these variables been used for stratification can be realized by poststratifying by these variables when constructing the poststratification weights. These poststratification weights were also employed for correcting erroneous exclusion (7.2.4).⁴

In sum, the construction of survey weights is very important in the HFCS. The following sections will explain how design, nonresponse and poststratification weights were constructed and how a final set of survey weights was derived from these weights. Finally, we will present some descriptive results that take these weights into account.

7.2 Construction of Survey Weights

7.2.1 Weight Components

We aim to construct a final survey weight w_i for each household i that is relatively small for households that are overrepresented in the sample compared to the target population and relatively large for households that are underrepresented. However, as already mentioned in the introduction, there are various reasons why households may misrepresent the target population. Depending on the type of misrepresentation, a specific adjustment by applying weights is required. In the HFCS, three types of weights are used: design weights w_{Di} , nonresponse weights

² Although the frame population was cleaned of addresses of companies, there may still be some erroneously left.

³ Sometimes referred to as noncoverage weights.

⁴ Poststratification weights can also correct for a third type of sample-specific bias: the target population may be randomly misrepresented by the specific households drawn into the sample.

w_{NRI} , and poststratification weights w_{PSi} . The product of these three weights yields the final survey weight w_i :

$$w_i = w_{Di} \cdot w_{NRI} \cdot w_{PSi}$$

Although some HFCS variables are asked at person level instead of household level, we do not construct weights for persons because the main focus of the survey is on the household level.⁵

7.2.2 Design Weights

Design weights help to reduce the unequal probability sampling bias, to repair erroneous inclusion and to correct for frame multiplicities. In the HFCS, we construct the design weights in three steps:

- Step 1: Constructing a preliminary set of design weights to eliminate the unequal probability sampling bias
- Step 2: Calibrating this preliminary set of design weights in such a way that they sum up to the frame population total
- Step 3: Resetting the calibrated design weights to zero for those observations that were either erroneously included or that were duplicated

In order to eliminate the unequal probability sampling bias, in step 1 we had to put proportionally more weight on households with a lower probability of being selected into the sample than other households and vice versa. Therefore, we had to calculate the probability of selection for each household and then get the preliminary design weight by inverting this probability. For example, a household with a probability of selection equal to 0.001 got a preliminary design weight equal to $1,000 = 1/0.001$, which is much higher than that used for a household with a probability of selection equal to 0.009, which would be $111 = 1/0.009$.

When calculating households' selection probabilities, we had to keep in mind the two-stage sample design: in the first stage the smallest territorial units, enumeration districts (Zählsprenkel), were drawn, and in the second stage households within the enumeration district were drawn (chapter 6). The probability that the i^{th} household in the j^{th} enumeration district is selected into the sample is the product of the enumeration district selection probability and the probability that the household is selected given that the household's enumeration district is selected. The inverse of this product is the preliminary design weight.

Due to this household reweighting, the estimated frame population total would no longer correspond to the actual frame population total. Therefore in step 2, the preliminary set of design weights was calibrated according to the actual frame population total.

Finally, although the frame population had been carefully prepared and cleaned before sampling, there were still some ineligible (see box 2) or duplicated observations in our sample (see also section 4.6.2.12), for example addresses belonging to companies or to care residences, or multiple domiciles of commuters. We flagged all such cases detected during the field work as ineligible or duplicated in our sample by setting the calibrated design weights equal to zero.

⁵ In order to carry out analyses at the individual level, additional weights would have to be constructed. This is planned as a future research project in the HFCS.

Table 11 shows some statistics of the obtained HFCS design weights across Austria's federal provinces. Vienna is the federal province with the lowest median weight, which makes perfect sense as households living in this province are oversampled and would have created a bias had they not been re-weighted downward by the design weights.

The value of a household's design weight can be interpreted as the number of households in the frame population that is represented by this household. For example, the median Viennese household represents 693 households in the frame population. Thus, the size of the frame population can be estimated by summing up all design weights, which in our case is 3,773,956 households.

Table 11

HFCS Design Weights by Federal Province

	Mean	Median	Minimum	Maximum
Vienna	747	693	0	2,554
Lower Austria	931	882	0	2,437
Burgenland	916	1,050	0	1,379
Styria	902	756	0	3,271
Carinthia	874	864	0	2,229
Upper Austria	844	857	0	2,394
Salzburg	905	969	0	1,712
Tyrol	872	879	0	1,852
Vorarlberg	895	849	0	1,560
All	851	833	0	3,271

Source: HFCS Austria 2010, OeNB.

Box 2

Unit Nonresponse in the HFCS in Austria

In the HFCS in Austria, successful interviews were conducted with 2,380 households of the gross sample, which comprised 4,436 addresses. The remaining 2,056 addresses were either classified as unit nonresponse (1,888 households), ineligible addresses (162 addresses) or addresses of unknown eligibility (6 addresses).

The unit nonresponse cases are households as defined in the HFCS that were not interviewed successfully. There were several reasons for this. First, and most importantly, households actively refused to take part in the survey, either by refusing to be interviewed or by breaking off the interview or by failing to keep the appointment for the interview and refusing to be contacted again. This applied to a total of 1,693 households. Second, no contact at all could be established with 75 households. Third, there were other reasons for nonresponse, such as illness, language barriers or ex post exclusion of interviews due to a high number of missing or unreliable values. This applied to 120 households.

In addition, 162 addresses were classified as ineligible because they were not part of the target population, for instance addresses of companies, empty buildings or second homes of households that were reached via their main residence address. Furthermore, there were another six addresses whose eligibility status was impossible to ascertain as the interviewers were unable to reach or find them. In accordance with the distribution of the eligibility status of the rest of the observed addresses in the sample, one of the six addresses was randomly chosen to be ineligible and the remaining five to be eligible.

The eligibility rate in the HFCS sample finally came to 96.3% and the nonresponse rate of the eligible households amounted to 44.3%. This means that successful interviews were conducted with 55.7% of the eligible households in the HFCS sample. 39.6% of the eligible households actively refused to take part in the survey.

7.2.3 Nonresponse Weights

As described in box 2, not all households participated successfully in the survey. If household characteristics correlate with nonresponse, the respondent population is not a random subsample of the frame population and the sample is nonresponse biased (chart 6). In the HFCS, this is indeed the case, as can be seen in table 12.

Table 12

Response Propensity Estimates Based on a Logit Regression Model

Covariates	Coefficients
Dwelling rating low	-0.260 (0.233)
Dwelling location countryside	-0.077 (0.0837)
Dwelling location Vienna	-1.091*** (0.0924)
Female interviewer	-0.138* (0.0738)
Interviewer's age	-0.005 (0.00359)
Interviewer's second-stage tertiary education	-0.550*** (0.0906)
Interviewer's working experience (in months)	0.00172*** (0.000512)
Number of interviewer contacts	-0.022 (0.0376)
Constant	0.883*** (0.197)
Observations ¹	4,273

Source: HFCS Austria 2010, OeNB.

¹ The remaining 163 observations in the data set are ineligible and are therefore not included in the regression.

Note: Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

The table shows a logit regression of household participation in the survey (1 if the household participated, 0 otherwise) on a set of variables that explain participation in the survey. The results show on the one hand that households living in Vienna and households contacted by interviewers with a high level of education have a lower probability of participating. On the other hand, households contacted by experienced interviewers achieve a significantly higher probability of responding than households contacted by other interviewers. This suggests that nonresponse is not random.

This bias can be corrected by using nonresponse weights, i.e. by attaching a higher nonresponse weight to households with a low probability of responding than the weight attached to households with a high probability of responding. To calculate the response propensities and the corresponding nonresponse weights we combine the weight-

ing class adjustment method with the model-based adjustment method (Biemer and Christ, 2008). The algorithm can be summarized in the following three steps:

- Step 1: Using the logit regression model shown in table 12 to estimate the probability of response for each household (given that the household was selected into the sample)
- Step 2: Grouping these households' response propensities into 5 classes (quintiles) and calculating the average response propensity for each class (unweighted total number of respondent households/unweighted total number of households)⁶
- Step 3: Getting the nonresponse weight of each class by inverting the average response propensity of each class

The advantage of this approach is that this method stabilizes the nonresponse weights because the response propensities predicted by the regression model vary widely and may contain extreme values.⁷ Information collected through interviewer questionnaires (e.g. level of education and experience of the interviewer)

⁶ The average response propensity is unweighted (with respect to the design weights) for efficiency reasons. See Little and Vartivarian (2003) for details.

⁷ Another problem of the simple use of logistic regression modelling, pointed out by Iannacchione et al. (1991), is that it does not ensure that the weighted sample marginal distributions conform to the population marginal distribution.

Table 13

HFCS Nonresponse Weights by Response Propensity

Quintile	Predicted response propensity %	Weight
First quintile	0 to 40	2.837
Second quintile	40 to 61	2.110
Third quintile	61 to 64	1.580
Fourth quintile	64 to 67	1.471
Fifth quintile	67 to 100	1.457

Source: HFCS Austria 2010, OeNB.

was found to correlate strongly and statistically significantly with households' response propensity and was therefore used in step 1.

The HFCS nonresponse weights are shown in table 13. They took five different values, one for each response class, and, by construction, those households with a high response propensity got a lower weight than those with a low response propensity. Nonrespondent households got a nonresponse weight equal to zero.

7.2.4 Poststratification Weights

Erroneous exclusion may be – as mentioned in the introduction – an imperfection in the HFCS frame with respect to our target population. We may have missed households without postal addresses, which implies that these types of households would be underrepresented. If there existed an external data set that covered these households and all the other ones of our target population, we could use it to adjust our sample to this external data set; we could put more weight on households that are similar to those without postal addresses in such a way that the estimated size of the target population in the HFCS would be the same as the one in the external data set.

Unfortunately, such a data set does not exist in Austria. Similar ones, like the EU-SILC (EU Statistics on Income, Social Inclusion and Living Conditions) or the Austrian Micro census, target a different population of households, which is due to their specific household definition. While the target population of the HFCS includes all households (according to the above definition), the EU-SILC and the Austrian Micro census include only households living in a dwelling officially registered in the Austrian residence registry (Zentrales Melderegister) as their main residence. This definition excludes a subset of households which are included in the HFCS household definition, namely all households living in dwellings that are not registered as a main residence or not registered at all. There are various reasons why households' actual main residences are not registered as such. For instance, students that studying away from home may keep their main residence at their parents' address even though they are already a household on their own by HFCS definition; others may just forget to register the address at which they actually live as their main residence. These and other problems when using main residence addresses for sampling households via the Austrian residence registry are also acknowledged by Statistics Austria.⁸

Given that these data sets also suffer from erroneous exclusion, it does not make sense to reweight the sample according to the target population *size* of these data sets. In addition, the erroneous exclusion bias in the HFCS sample is likely to be very small, as the vast majority of households do have postal addresses. How-

⁸ For the Micro census see Statistics Austria (2011a), page 10; for the EU-SILC see Statistics Austria (2011b), page 27.

ever, it appears to make sense to reweight the HFCS sample according to the *structure* of the target population of the external data sets. In particular, this may deliver a better picture with regard to the proportions of households in the Austrian provinces as the Austrian Micro census uses a much larger sample than the HFCS. We constructed poststratification weights that put more weight on households with a low probability of being included in the frame and less weight on households with a high probability; we did not change the HFCS frame population size, but we adjusted the proportions of various household types. This increases the comparability between the HFCS and the Micro census and at the same time reduces the erroneous exclusion bias. Furthermore, poststratification weights can also reduce the sampling variance and, hence, increase the precision of the estimators; moreover, they can eliminate sample-specific random misrepresentations of the target population (section 7.1).

The HFCS poststratification weights are constructed following the poststratification cell adjustment method (Biemer and Christ, 2008) and using the Austrian Micro census data available during the field phase of the HFCS in Austria (2010 Q4). The procedure was as follows:

- Step 1: Choosing good predictors for a household being included in the HFCS frame and cross-tabulate these variables to build the poststratification cells; our poststratification variables are municipality size and household size.
- Step 2: Calculating the average inclusion-in-frame propensity for each cell

$$\text{Total HFCS frame population} \cdot \frac{\text{HFCS frame population in the cell}}{\text{Micro census frame population in the cell}} \cdot \frac{\text{Total Micro census frame population}}{\text{Total Micro census frame population}}$$

- Step 3: Getting the poststratification weight by inverting the average inclusion-in-frame propensity for each cell.

We grouped household size into two very broad classes: 1 to 4 persons, and 5 or more persons. This way larger households are not underrepresented in the HFCS sample and, at the same time, we did not create a bias in our target population toward the Micro census' target population, because we are aware of the fact that in the HFCS target population the proportion of smaller households is higher (as it

includes households living in a dwelling without official main residence status in the residence registry). Moreover, we assigned the households to seven municipality sizes.

The HFCS poststratification weights are shown in table 14. They take 14 different values, one for each municipality size class and household size class. We can see that large households were underrepresented in the HFCS frame population because they tend to have higher poststratification weights. Table 14 also shows that large municipalities have lower poststratification weights and were therefore overrepre-

Table 14

HFCS Post-Stratification Weights for Municipality Size Class and Household Size Class

Municipality size (number of inhabitants)	Household size (number of persons)	
	1 to 4	5 or more
Up to 2,000	0.974	1.133
2,001 to 3,000	1.169	1.055
3,001 to 5,000	0.963	1.447
5,001 to 10,000	1.053	1.813
10,001 to 20,000	0.915	2.012
20,001 to 1 Mio	0.911	1.913
Over 1 million	0.977	1.249

Source: HFCS Austria 2010, OeNB.

sented (the average weight of the last three size classes is lower than that of the first four ones).

7.2.5 Final Weights

Three different weights were constructed to account for the different reasons why a household may misrepresent the target population. As we have seen, each one of these weights can be interpreted as an inverted probability. By taking their product we obtain a new inverted probability, which is the HFCS final weight w_i :

$$w_i = W_{Di} \cdot W_{NRi} \cdot W_{PSi}$$

$$w_i = \frac{1}{P(i \text{ is selected}) \cdot P(i \text{ responds} | i \text{ is selected}) \cdot P(i \text{ is included in the frame} | i \text{ is selected and responds})}$$

$$w_i = \frac{1}{P(i \text{ is selected and } i \text{ responds and } i \text{ is included in the frame})}$$

The final weight w_i incorporates all three adjustments and can be interpreted as the inverted probability that household i is in the net sample. Households with a high probability of being in the net sample have a lower final weight and represent fewer households in the target population than households with a low probability of being in the net sample.

Table 15 summarizes the HFCS nonresponse and poststratification adjustments, which we obtain by taking the product of the nonresponse weights (table 13) and the poststratification weights (table 14). These adjustments take 48 different values depending on household size, municipality size and predicted response propensity classes.

Table 15

Nonresponse and Poststratification Adjustments

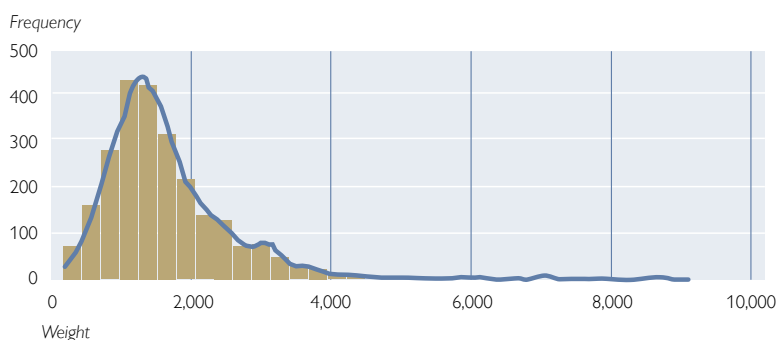
Poststratification cells		Nonresponse cells				
		0 to 40	40 to 61	61 to 64	64 to 67	67 to 100
Up to 2,000	1 to 4		2.055	1.539	1.432	1.419
	5 or more		2.392	1.791	1.667	1.651
2,001 to 3,000	1 to 4		2.468	1.848	1.720	1.703
	5 or more		2.227	1.668	1.552	1.537
3,001 to 5,000	1 to 4		2.033	1.523	1.417	1.403
	5 or more		3.053	2.286	2.128	2.107
5,001 to 10,000	1 to 4		2.222	1.665	1.549	1.534
	5 or more		3.826	2.866		2.641
10,001 to 20,000	1 to 4		1.931	1.446	1.346	1.333
	5 or more			3.181	2.960	2.931
20,001 to 1 million	1 to 4		1.923	1.440	1.340	1.327
	5 or more			3.024	2.814	2.787
Over 1 million	1 to 4	2.771	2.061			
	5 or more	3.544				

Source: HFCS Austria 2010, OeNB.

Note: Empty cells denote that there are no households of this size in the sample.

Chart 7

Distribution of Final HFCS Weights



Source: HFCS Austria 2010, OeNB.

Table 16

Comparison of Weighted and Unweighted Means of Selected HFCS Variables (Imputed)

	Mean	
	unweighted	weighted
Household size (number of persons)	2.11	2.13
	<i>Percentage share of all households</i>	
Vienna	28.3	23.3
Lower Austria	17.3	16.0
Burgenland	3.0	4.1
Styria	13.0	15.4
Carinthia	6.2	6.8
Upper Austria	15.1	15.3
Salzburg	5.7	6.6
Tyrol	7.6	8.4
Vorarlberg	3.8	4.1
	<i>EUR</i>	
Estimated household monthly net income	2,287	2,306
Value of main residence ¹	257,666	258,072

Source: HFCS Austria 2010, OeNB.

¹ The variable DA1110 was used here.

Finally, when taking into account also the design weights, we obtain the HFCS final weights, whose distribution is shown in chart 7. The HFCS final weights range from 169 to 9,054, with the mean being 1,586 and the median 1,429. Their distribution is slightly skewed to the right, which is not atypical for unequal probability sample designs: Households with a higher probability of selection (design weights below average) dominate the sample. This effect is reinforced by the further weight adjustments.

7.3 Selected Results

Table 16 shows the impact that the HFCS final weights have on estimations by comparing selected weighted and unweighted mean values of HFCS variables. For example, we can see that Viennese households were strongly down-weighted from 28.3% to 23.3%. This means that despite their high non-response rate, in sum, Viennese households were clearly overrepresented in the sample with respect to the target population. We can also see from the table that households with more expensive primary residences were underrepresented in the unweighted sample, which can probably be explained by these households' higher nonresponse rates.

The use of the final HFCS weights is sufficient when calculating e.g the weighted statistics shown in table 16.

To calculate the corresponding correct

variances or standard errors of these estimators, however, resampling weights, which are described in chapter 8, are necessary.

7.4 Concluding Remarks

We constructed a set of final weights to correct imperfections in the unweighted HFCS sample with respect to the HFCS target population. These imperfections are the unequal probability sampling bias, erroneous inclusion, frame multiplicity, and erroneous exclusion.

While the weighted HFCS sample enables unbiased population estimates, it also increases the variance of the population estimates, which makes them less

precise. According to the Unequal Weighting Effect (UWE) statistic developed by Kish (1995), the variance of HFCS population estimates may be increased by a maximum of 27.7% ($UWE=1+cv^2=1.2768$) as a result of weighting, which is a relatively small effect. Therefore, it is not necessary to apply weight trimming methods. Furthermore, we prefer to have much less bias at the cost of a small increase in variance in order to avoid distorted results being wrongly considered significant too often.

An explanation of how to correctly use the weights in Stata[®] is provided in chapter 9, User Guide