7 Construction of survey weights

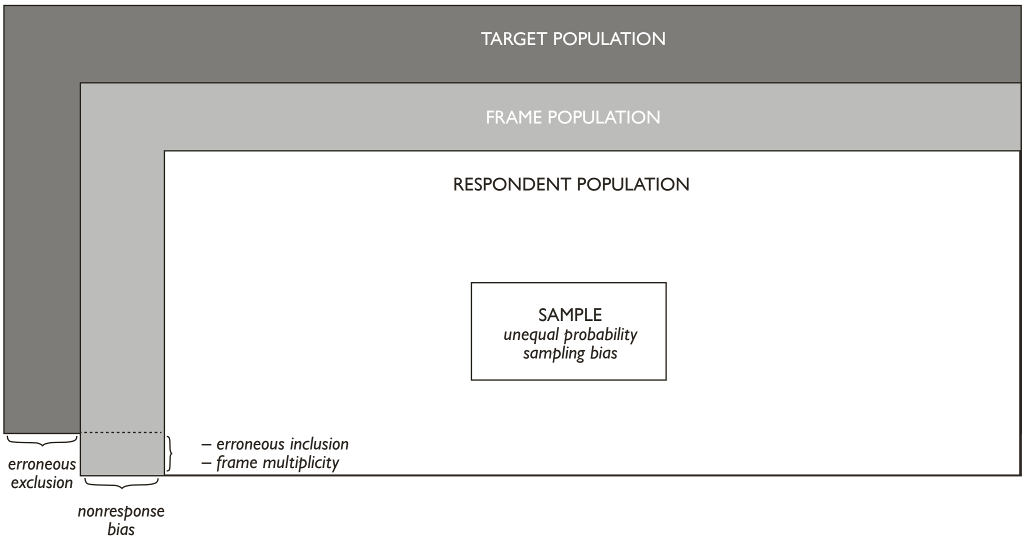
7.1 Introduction

Survey weights are usually computed 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 an individual 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: unequal probability sampling bias, frame bias and nonresponse bias (see chart 5).

Chart 5

Misrepresentation of the target population in the sample



Source: Adapted from Biemer and Christ (2008).

As mentioned above, the unequal probability sampling bias is due to the fact that not every household has the same probability of being selected into the sample, reflecting the fact that the number of primary sampling units (PSUs) to be drawn per stratum is fixed by the HFCS sampling design. Another example is the oversampling of households in metropolitan areas (like Vienna) in the HFCS sample, which is used to address the known problem of the relatively low survey participation propensity of urban households. To correct these misrepresentations, we computed design weights, which will be explained in section 7.2.2. Further details about the HFCS sampling design can be found in chapter 6.

Imperfections in the survey frame from which the sample is drawn can lead to frame bias. In the HFCS, the sampling frame is a list of all personal postal addresses in Austria (see chapter 6). Erroneous exclusion of households could imply an imperfection with respect to the target population. In other words, there is the possibility that households without a postal address, for example, one-person

[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.*

households living together in residential communities and sharing an address that contains only one of these households, were excluded. These types of households would then be underrepresented. Another imperfection of the frame could be caused by erroneous inclusion, that is, the inclusion of addresses not belonging to households, for example, those of companies[2] or individuals living 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 using design weights[3] (to address erroneous inclusion and frame multiplicity) or poststratification weights (to address erroneous exclusion). We explain the computation of these weights in more detail in sections 7.2.2 and 7.2.4.

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 – a phenomenon widely corroborated in literature (see e.g. Kennickell and McManus, 1993). Thus, estimates for the sampling frame would be biased with respect to these group characteristics, even though they are unbiased for the participating population. Using nonresponse weights can correct this bias (section 7.2.3).

Furthermore, as mentioned above, survey weights can help to reduce sampling variance, and, hence, to increase the precision of the estimators. Ideally, the precision of the estimators should be improved by 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 the estimators, were not 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 possible by using these variables for stratification can be achieved by using these variables for poststratification. These poststratification weights were also utilized for correcting erroneous exclusion (see chapter 7.2.4).[4]

The construction of survey weights is very important for the HFCS. The following sections will explain how design, nonresponse and poststratification weights were computed and how the final set of survey weights was derived from these weights. Moreover, 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 every 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, households may misrepresent the target population for various reasons. Therefore a specific adjustment using weights

[2] *Although the sampling frame was cleaned of addresses of companies, some may still be erroneously included.*

[3] *Sometimes referred to as noncoverage weights.*

[4] *Poststratification weights can, moreover, correct a third type of sample-specific bias: the target population may be accidentally misrepresented by the specific households drawn into the sample.*

is required for every type of misrepresentation. In the HFCS, three types of weights are used: design weights w_{Di}, nonresponse weights 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 the individual level rather than the household level, no weights were constructed for individuals because the main focus of the survey is the household level.

7.2.2 Design weights

Design weights help reduce the unequal probability sampling bias as well as rectify erroneous inclusion and frame multiplicities. In the HFCS, we compute the design weights on the basis of two-stage cluster sampling and the selection probabilities of the primary sampling units (PSUs) and the secondary sampling units (SSUs). In the first stage, the smallest territorial units, the so-called enumeration districts (PSUs), are drawn; then in the second stage, the households (SSUs) within these enumeration districts are drawn (see section 6). The probability that the i^{th} household in the j^{th} enumeration district is selected in the sample is the product of the selection probability for the enumeration district and the selection probability for the household, under the condition that the household's enumeration district is selected. The inverse of this product is the preliminary design weight. The calculation of the design weight mirrors the two steps of the sampling procedure:

Step 1: Calculate the probability that a certain PSU is selected. As described in section 6, this sampling probability is defined depending on the relative number of households in a PSU. The probability that PSU j will be selected in stratum h is

$$PSU(h,j) = \frac{M_{hj} m_h}{N_h},$$

where M_{hj} represents the number of households in this enumeration district *(h,j),* m_h the number of PSUs to be drawn in this stratum, and N_h the number of households in this stratum.

Step 2: Calculate the probability that an SSU is selected. Under the condition that a PSU is chosen, each household in this enumeration district has the same probability of being chosen. Thus, the probability of being selected is given by

$$\frac{m_{hj}}{M_{hj}},$$

where m_{hj} is the number of households to be drawn in the PSU (i.e. 8 in a stratum with a population of over 50,000 and 12 in the rest of Austria). As above, M_{hj} is the number of households in this enumeration district.

Overall, the ex ante selection probability *Prob(i)* for every household *i* is given by multiplying the two partial probabilities. This probability may be shown as:

$$Prob(i) = \frac{m_{hj} m_h}{N_h} = \frac{1}{w_{Di}}.$$

The design weight *(*w_{Di}*)* is calculated by inverting this probability. For example, a household with a probability of selection equal to 0.001 has a preliminary design weight of *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*.

This procedure ensures that every household that has the same probability of selection within a stratum on account of the sample design also has the same design weight. The design weights vary across the strata, due to the differing assumptions on the willingness of a household to participate, which determine the SSUs to be drawn, and the different size of the strata as a result of the number of households.

Finally, although the sampling frame was carefully prepared and cleaned before sampling, our sample still included some ineligible (see box 2) or duplicated observations (see also section 4.6.2.13), for example company addresses, addresses of care homes or secondary residences. We flagged all such cases detected during the fieldwork as ineligible or duplicated in our sample by setting the design weights equal to zero. As a result, the design weight total decreased from about 4.1 million to 3.9 million.

Table 12

HFCS design weights by federal province

	Mean	Median	Minimum	Maximum
Vienna	494	513	0	707
Lower Austria	757	779	0	1,099
Burgenland	729	694	0	1,008
Styria	641	636	0	1,146
Carinthia	613	484	0	1,183
Upper Austria	626	632	0	1,004
Salzburg	600	512	0	925
Tyrol	651	624	0	1,085
Vorarlberg	792	833	0	1,045
Total	614	547	0	1,183

Source: HFCS Austria 2014, OeNB.

Table 12 shows some statistics of the obtained HFCS design weights across Austria's provinces. Vienna and Salzburg are the provinces with the lowest median weights, which is plausible, as households living in these regions were oversampled because of their low willingness to participate during the first HFCS wave in Austria, which would have created a bias had they not been reweighted downward using the design weights.

The value of a household's design weight can be interpreted as the number of households in the sampling frame that is represented by this household. For example, the median household in Vienna represents 513 households in the sampling frame.

Box 2

Unit nonresponse in the HFCS in Austria

In the HFCS in Austria, successful interviews were conducted with 2,997 households from the gross sample, which comprised 6,308 addresses. The remaining 3,311 addresses were classified either as unit nonresponse (2,997 households), ineligible addresses (284 addresses) or addresses of unknown eligibility (30 addresses).

The unit nonresponse cases are households as defined in the HFCS that were not interviewed successfully for several reasons. The most common reason was that households actively refused to take part in the survey, either by refusing to be interviewed, breaking off the interview or by failing to keep the interview appointment and being subsequently unavailable for contact. This applied to a total of 2,657 households. Another reason was that no contact at all could be established with 136 households. The remaining 204 nonrespondents specified other reasons, such as illness, language barriers; or they resulted from ex post exclusion of interviews due to a high number of missing or unreliable values.

In addition, 284 addresses were classified as ineligible because they were not part of the target population, as they were, for instance, addresses of companies, empty buildings or second homes of households that could be reached via their main residence address. Finally, the eligibility status of another 30 addresses was impossible to ascertain, as the interviewers were unable to reach or find them. In accordance with how the eligibility status of the rest of the observed addresses in the sample was distributed, one of the 30 addresses was randomly chosen to be ineligible and the remaining 29 to be eligible.

The eligibility rate in the HFCS sample ultimately came to 95% and the nonresponse rate of the eligible households amounted to 50.2%. This means that successful interviews were conducted with 49.8% of the eligible households in the HFCS sample. Just 44.1% 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 sampling frame and the sample is nonresponse biased (see chart 5). In the HFCS, this is indeed the case, as can be seen in table 13. The table shows a logit regression of household participation in the survey (1 if the household participated, otherwise 0) on a set of variables that explain participation in the survey. The results show on the one hand that households living in apartments or in municipalities with higher personal incomes or with higher unemployment rates have a lower probability of participating. On the other hand, households contacted by older interviewers exhibit a significantly higher probability of responding than households contacted by younger interviewers. Moreover, households that were contacted by interviewers with a university degree, or households that live in neighborhood without graffiti or in municipalities with a higher average population age had an increased probability of participating. This suggests that nonresponse is not random.

Table 13

Response propensity estimates based on a logit regression model

Covariates	Coefficients
Paradata on the interview, place of residence and neighborhood	
Household interview order	0.00446***
	(0.000545)
Building characteristics (reference group: detached single-family house)	
Semi-detached single-family house	0.108
	(0.166)
Single-family townhouse	−0.0911
	(0.190)
Residential farm building	−0.279
	(0.178)
Apartment in a (high-rise) apartment building	−0.419***
	(0.0797)
Student dormitory/rented room	−0.601
	(0.429)
Other type of building	1.827***
	(0.444)
Building design characteristics (reference group: premium)	
Very good	−0.0765
	(0.187)
Medium	−0.102
	(0.189)
Basic	−0.168
	(0.205)
Very basic	0.120
	(0.287)
Location characteristics (reference group: city center)	
Between the city center and suburbs	0.225***
	(0.0796)
Suburbs and city outskirts	0.130
	(0.0844)
Countryside	−0.121
	(0.104)
Graffiti in the neighborhood (reference group: many)	
Location – graffiti = 2, some	0.879**
	(0.375)
Location – graffiti = 3, few	0.485
	(0.361)
Location – graffiti = 4, none at all	0.601*
	(0.359)
Sample design variables	
Design weight	0.000915***
	(0.000202)
Interviewer characteristics	
Female interviewer	−0.311***
	(0.0590)
Interviewer's age	0.0141***
	(0.00273)
Interviewer's second-stage tertiary education	0.302***
	(0.0677)
Interviewer's working experience in months	−0.00168***
	(0.000367)
Variables at the municipality level	
Average per capita income per municipality in 2011	−0.0000299**
	(0.0000119)
Share of employees in the primary sector per municipality in 2011	−4.090***
	(1.134)
Share of university-trained population per municipality in 2012	−1.199**
	(0.474)
Unemployment rate per municipality in 2011	−6.096***
	(1.408)
Average age of population per municipality in 2011	0.00454**
	(0.00221)
Variables at the district level	
Average crime rate per district in 2009 and 2010	0.000160
	(0.000360)
Constant	−0.325
	(0.609)
Observations[1]	6,023

Source: HFCS Austria 2014, OeNB.

[1] *The remaining 285 observations in the dataset are ineligible and are therefore not included in the regression.*

*Note: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.*

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 to households with a high probability of responding. To calculate the response probabilities and the corresponding nonresponse weights, the weighting class adjustment method is combined with the model-based adjustment method (see Biemer and Christ, 2008). The weighting classes are chosen optimally using the method described by Haziza and Beaumont (2007). The algorithm can be summarized in the following three steps:

Step 1: The logit regression model shown in table 13 was used to estimate the probability of response for each household (assuming that the household was selected in the sample)

Step 2: These households' response probabilities were grouped into seven classes. The number of classes and their resultant sizes are chosen optimally in line with Haziza and Beaumont (2007). To this end, a k-means algorithm is used to cluster households into a prespecified number of response classes with low variance and similar size. Next, class indicators are used as explanatory variables for the response propensity based on an ordinary least squares (OLS) regression from the logit regression model estimated in step 1. Beginning with one class, the number of classes is increased in an iterative process until the adjusted R^2 of this OLS regression exceeds 95%. This is the case for seven classes in the second wave of the HFCS in Austria. Finally, the average response propensity for each class was calculated (unweighted total number of respondent households/unweighted total number of households).[5]

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

Table 14

HFCS nonresponse weights by response propensity

Classes	Predicted response propensity	Weight
	%	
I	0 to 33	3.417
II	33 to 41	2.694
III	41 to 48	2.375
IV	48 to 55	1.853
V	55 to 64	1.664
VI	64 to 75	1.517
VII	75 to 100	1.216

Source: HFCS Austria 2014, OeNB.

Step 3: The nonresponse weight of a class is obtained by inverting the average response propensity of the respective class.

The advantage of this approach is that it stabilizes the nonresponse weights because the response propensities predicted by the regression model vary widely and can contain extreme values.[6] Information collected through interviewer surveys (see section 3.8), e.g. their level of education and experience, was found to correlate strongly and statistically significantly with households' response propensity and was therefore used in step 1. Additionally, sample design information and municipal or district-level information was used, which may also explain willingness to participate with statistical significance.

The HFCS nonresponse weights are shown in table 14. A value was calculated for each of the seven response groups and, by design, households with a high response propensity were assigned a lower weight than those with a low response propensity. Nonrespondent households were assigned a nonresponse weight equal to zero.

7.2.4 Poststratification weights

Erroneous exclusion may – as mentioned above – be an imperfection in the HFCS frame with respect to the target population. We may have missed households without postal addresses, which means that these types of households would be underrepresented. If an external dataset covering these households and all others in our target population existed, we could use it to adapt our sample to this external dataset accordingly; we could put more weight on households without postal addresses so that the estimated size of the target population in the HFCS would be the same as the one in the external dataset.

Unfortunately, such a dataset does not exist in Austria. Similar surveys, like the EU SILC (EU Statistics on Income and Living Conditions) or the Austrian microcensus, target a different population of households 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 microcensus only include households living in a dwelling officially registered in the central residence registry as their main residence. This definition excludes a subset of households 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 as to why in some cases households' actual main residences are not registered as such. For instance, students studying away from home

[6] *Another problem of the use of simple logit regression models, as highlighted by Iannacchione et al. (1991), is that such modeling does not ensure that the weighted sample marginal distributions conform to the population marginal distribution.*

may keep their main residence at their parents' address even though they are already a household of their own according to the HFCS definition; others may have just forgotten to register the address where they actually live as their main residence. Statistics Austria also acknowledges these problems and others when using main residence addresses for sampling households via the Austrian residence registry.[7]

Given that these datasets also suffer from erroneous exclusion, it does not make sense to reweight the entire sample according to the target population size of these datasets. However, during the second wave of the HFCS in Austria, the question the residence is registered as a main residence or not was recorded in the interview, so it would appear to make sense to reweight this group of households to the microcensus. In particular, this may deliver a better picture with regard to the proportions of households in the Austrian provinces, as the Austrian microcensus uses a much larger sample than the HFCS. For the small group of remaining households in the HFCS sample that are not registered at their main residence, reweighting the microcensus does not seem sensible. Yet the erroneous exclusion bias in the HFCS sample is likely to be very small in this case, as the vast majority of households do have postal addresses. We constructed poststratification weights that put more weight on households with a lower probability of being included in the frame and less weight on households with a higher probability. We adjusted the HFCS frame population size only for households registered at their main residence according to the microcensus. Households not registered at their main residence are then added.[8] This increases comparability between the HFCS and the microcensus in the second wave 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 (see section 7.1).

The HFCS poststratification weights are computed following the poststratification cell adjustment method (Biemer and Christ, 2008) and using the Austrian microcensus data (2014 Q4) available during the field phase of the HFCS in Austria. The procedure was as follows:

Step 1: Choose suitable predictors for including a household in the HFCS frame and cross-tabulating these variables to compute the poststratification cells. Different poststratification cells were defined depending on the registration status. For households registered at their main residence, the province, the tenure status of the main residence and household size serve as poststratification variables. No poststratification is performed for all other households, because, as described above, they are not included in the external dataset.

[7] *For the microcensus, see Haslinger und Kytir (2006), p. 512 f; for the EU SILC, see Statistics Austria (2015), p. 45.*

[8] *Before the poststratification adjustment, the HFCS frame population encompassed 3,875,337 households, consisting of households registered at their main residence (3,802,620) and households not registered at their main residence (72,717 or about 2% of households). After the poststratification adjustment of households registered at their main residence, the population of these households comes to 3,789,808, which corresponds to the household population according to the 2014 Q4 microcensus. As a result, the final HFCS household population amounts to 3,862,525 (= 3,789,808 + 72,717).*

Step 2: Calculate the average propensity to be included in the frame population for each cell:

$$\frac{\textit{HFCS frame population in the cell}}{\textit{Microcensus frame population in the cell}}$$

Step 3: In each cell, the propensity was adjusted by a constant factor, thus adjusting the external dataset total.

Step 4: Obtain the poststratification weight by inverting the average inclusion propensity for each cell.

Households registered at their main residence were grouped according to size, one group containing households with one to four individuals and one containing those with five or more individuals.[9] This ensures that larger households are not underrepresented in the HFCS sample. Additionally, households registered at their main residence were grouped into (part) owners[10] and tenants. Moreover, the households were assigned to the nine provinces.

Table 15 shows the HFCS poststratification weights – 36 values, i.e. one value per combination of registration status, province, tenure status of the main residence and household size. The table shows, for example, that large households are underrepresented in the HFCS sampling frame, as they tend to exhibit larger poststratification weights.

Table 15

HFCS poststratification weights for registration status of main residence, province, tenure status and household size by response propensity

	Official main residence				Other households			
	Household size (number of persons)				Household size (number of persons)			
	1 to 4	5 or more	1 to 4	5 or more	1 to 4	5 or more	1 to 4	5 or more
	Homeowners		Tenants		Homeowners		Tenants	
Vienna	1.068	0.978	0.964	1.394	1			
Lower Austria	1.418	2.463	0.556	1.945				
Burgenland	0.870	2.425	1.101	0.335				
Styria	0.968	1.251	0.930					
Carinthia	1.115	1.347	0.685	1.608				
Upper Austria	1.270	1.367	0.941	1.052				
Salzburg	0.743	0.478	2.330	3.142				
Tyrol	0.956	1.455	1.118	1.635				
Vorarlberg	1.480	0.824	0.408	0.562				

Source: HFCS Austria 2014, OeNB.

7.2.5 Final weights

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

[9] *Given the very low number in a poststratification cell, for main residence tenants in Styria the household size cells were aggregated.*

[10] *Includes free users of main residences.*

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

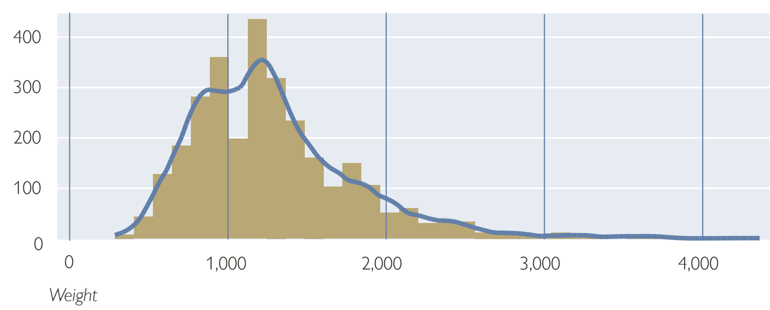
$$w_i = \frac{1}{Prob(i\ is\ selected) \cdot Prob(i\ responds \mid i\ \ is\ selected) \cdot Prob(i\ is\ included\ in\ the\ frame) \mid i\ is\ selected\ and\ responds)}$$

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

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.

Chart 6

Distribution of HFCS final weights



Source: HFCS Austria 2014, OeNB.

The combination of nonresponse weights and poststratification weights results in 252 different weight adjustment cells based on registration status, province, tenure status of the main residence, household size and the response propensity classes described above. Each household is represented in precisely one of these cells.

Finally, once we have taken the design weights into account, we obtain the HFCS final weights, whose distribution is shown in chart 6. The HFCS final weights range from 287 to 4,360, with the mean being 1,289 and the median 1,207. Their distribution is slightly skewed to the right, which is not atypical for unequal probability sample designs. After all, households with a higher probability of selection (below average design weights) dominate the sample. This effect is reinforced by the further weight adjustments. In addition, a slight bipolarity can be identified in the distribution of the weights as a result of the nonresponse weights.[11]

7.3 Selected results

Table 16 shows the impact of the HFCS final weights on estimations by comparing selected weighted and unweighted mean values of HFCS variables. For example, we can see that households in Vienna were strongly downweighted from 24.7% to 23.1%. This means that despite their high nonresponse rate, overall households in Vienna were clearly overrepresented in the sample with respect to the target population. The comparison also shows that households with higher income or higher net wealth were underrepresented in the unweighted sample, which is probably caused by these households' higher nonresponse rate.

[11] *Table 14 shows a fairly strong increase in nonresponse weights from class IV to class V, which causes the above-mentioned bipolarity in the distribution of the final weights.*

Using the final HFCS weights is sufficient when calculating the weighted statistics shown in table 16. To calculate the appropriate correct variances or standard errors of these estimators, however, replicate weights, which are described in chapter 8, are necessary.

7.4 Concluding remarks

We computed a set of final weights to correct imperfections in the unweighted HFCS sample with respect to the HFCS target population. These imperfections are 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.[12] 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 16.7% *(UWE = 1 + coefficient of variation^2 = 1.167)* as a result of weighting. The adapted sample design made it possible to further improve this value compared to the first wave. Therefore, it is not necessary to apply weight trimming methods. Furthermore, a small increase in variance is acceptable in return for a significant reduction in the bias if it helps to avoid distorted results being classified as significant too often.

An explanation of how to correctly use the weights in Stata® is provided in chapter 9 (User guide).

Table 16

Comparison of weighted and unweighted means of selected HFCS variables (imputed)

	Mean	
	Unweighted	Weighted
Household size (number of persons)	2.07	2.14
	% of households	
Vienna	24.7	23.1
Lower Austria	16.6	18.9
Burgenland	3.2	3.2
Styria	13.8	13.7
Carinthia	6.3	6.5
Upper Austria	15.9	15.9
Salzburg	7.2	6.1
Tyrol	7.8	8.4
Vorarlberg	4.3	4.2
	EUR	
Estimated household monthly net income	2,388	2,450
Household net wealth	227,887	258,414

Source: HFCS Austria 2014, OeNB.

[12] *The poststratification step can restrict this increase in sample variance (see Levy and Lemeshow, 2008).*