

Sampling and Weighting the Sample of
the Adult Cohort of the National
Educational Panel Study (Wave 2 to 5)
Technical Report on SUF SC6
Version 5.0.0

*Sabine Zinn, Christian Aßmann, and
Ariane Würbach*

Copyrighted Material
Leibniz Institute for Educational Trajectories (LifBi)
Wilhelmsplatz 3, 96047 Bamberg
Director: Prof. Dr. Hans-Günther Roßbach
Executive Director of Research: Dr. Jutta von Maurice
Executive Director of Administration: Dr. Robert Polgar
Bamberg, 2015

**Sampling and Weighting the Sample of the Adult
Cohort of the National Educational Panel Study
(Wave 2 to 5)**

Technical Report on SUF SC6 Version 5-0-0

Sabine Zinn and Christian Aßmann and Ariane Würbach

Corresponding address: methoden@lifbi.de

version: February 10, 2016

1 Prequel

This report documents the target population, the sampling, the sample sizes, and the weighting procedures of the panel Waves¹ 2 to 6 of the NEPS Starting Cohort 6 (Adult Education and Lifelong Learning).² The remainder of this report is structured as follows: Section 2 introduces the target population of the Starting Cohort and the sampling design applied. Furthermore, here the composition of the gross and the net samples of the different waves are described. In Section 3, the derivation of the sampling weights are elaborated in very detail. This includes, the computation of design weights, non-response adjustments, and post-stratification of weights. Section 4 gives a summary of the design variables and sampling weights provided. Section 5 concludes with some comments regarding the usage of sampling weights for analysis.

2 Population, Sampling Design, and Sample Sizes

The target population of the Starting Cohort 6 comprises people living in private households in Germany and are born in the years between 1944 and 1986. Access to this population is gained via three subsamples. The first subsample consists of the sample of the survey “Working and Learning in a Changing World (ALWA)” conducted in 2009 by the Institute for Employment Research (IAB); for further details see Antoni et al. (2010). The ALWA sample covers birth cohorts from 1956 to 1986. All participants of ALWA were asked to participate in NEPS. Those who agreed to participate form the first subsample of the initial sample of the Starting Cohort 6. In addition to this subsample, two further subsamples have been established: a refreshment sample that also covers the birth cohorts from 1956 to 1986 and an enhancement sample covering individuals born between 1944 and 1954. The refreshment sample was drawn from the same target population as the ALWA sample, that is, within the same communities. Likewise, these communities served as the basic population to draw the enhancement sample of elderly people from. In other words, all individuals who are born between 1944 and 1986 and who lived at the date of drawing (January 2005) in one of the municipalities which were sampled in the context of ALWA form the target population of the Starting Cohort 6. The sampling of the refreshment and the enhancement sample of Starting Cohort 6 was conducted on the basis of a stratified two stage sampling approach. First, all German communities were subject to an *implicit* stratification according to Federal States, administrative districts, and classification of urbanization (BIK categorization). Then, within each stratum municipalities are sampled³ proportional to the resident population of the target population of ALWA corresponding to the respective stratum. The measure of size was the number of individuals born between 1956 and 1986. The sampling frame used for this purpose was built on the basis of the German resident population data provided by the German Federal Statistical Office and the statistical offices of the German Länder. To finally sample municipalities, 281

¹The sample of the ALWA survey is counted as Wave 1. It served as a basis to establish the initial sample of Starting Cohort 6. For further information see Section 2.

²The five waves correspond to the studies B72 (Wave 2), B67 (Wave 3), B68 (Wave 4), and B69 (Wave 5).

³Actually, these communities had already been sampled in the context of ALWA.

Table 1: The subsamples of the initial sample of Starting Cohort 6, the corresponding gross sample sizes, as well as net sample sizes realized in Wave 2 (i.e., in study B72).

Subsamples	Gross sample size	Net sample size	Response rate
ALWA sample	8,997	6,572	73.05%
Refreshment sample	6,547	1,971	30.11%
Enhancement sample	11,465	3,106	27.09%
Total	27,009	11,649	43.13%

sampling points⁴ corresponding to 250 communities have been selected. Sampling points have been allocated according to the size of the resident population of a municipality.⁵ Sampled municipalities which dropped out are replaced by municipalities from the same stratum which are structurally similar concerning size of resident population. Thus, in the end only 271 sampling points had been allocated.⁶ From the registries of the registration offices of the corresponding municipalities addresses are drawn by means of systematic random sampling. Thus, municipalities form the primary sampling units and addresses the secondary sampling units. In the sampling process, all individuals who were part of the resident population of the sampled municipalities at the date of sampling (2008) and who were born between 1944 and 1986 have been considered. In the refreshment sample, 24 addresses had been drawn per sampling point and in the enhancement sample 45 addresses per sampling point. That way, 6,547 addresses with telephone number could be determined for the refreshment sample and 11,465 addresses with telephone number for the enhancement sample. In sum, 8,997 individuals who participated in ALWA agreed to take part in NEPS. Table 1 depicts the resulting gross sample and the number of individuals who gave an evaluable interview in Wave 1 (i.e., in study B72), that is the net sample size.

The gross sample of Wave 3 (i.e., of study B67) comprises all individuals who were asked for an interview in Wave 2 minus those individuals who refused to further take part in the panel until September 2010. In sum, 287 individuals refused to further participate. Additionally, 833 individuals who participated in ALWA and could not be reach before (i.e., when establishing the gross sample of Wave 2) gave their permission to be surveyed in the context of NEPS. Table 2 gives the related gross and net sample sizes. Generally, all gross sample sizes given in this report are cleansed for individuals who refused to further participate and for retroactive data deletion. The net samples presented always exclude unfinished interviews. In total, 805 individuals of the panel cohort of Wave 3 are final drop-outs. In Wave 4 (i.e., in study B68), the sample of Starting Cohort 6 was enriched by a further refreshment sample covering the birth cohorts from 1944 to 1988. For this

⁴Commonly, for administrative reasons within municipalities only multiples of a fixed quantum can be sampled. Therefore, the overall goal to sample addresses of individuals is achieved via sampling artificial units called sample points.

⁵Note that such processing allows for multiple sampling points per municipality. In the considered case, there is one municipality with either assigned four, five, six, or twelve sampling points and eight municipalities were assigned two sampling points.

⁶The reason is that the NEPS sample was sampled from exactly the same municipality as the ALWA sample, and of that sample ten municipalities decided not to participate any longer. Note that these ten municipalities of the 250 selected ones could not be replaced that way.

Table 2: The gross and net sample sizes of Wave 3 (i.e., of study B67).

Subsamples	Gross sample size	Net sample size	Response rate
ALWA sample	7,402	5,637	76.16%
Refreshment sample	1,871	1,395	74.56%
Enhancement sample	2,922	2,288	78.30%
Total	12,195	9,320	76.42%

Table 3: The gross and net sample sizes of Wave 4 (i.e., of study B68).

Subsamples	Gross sample size	Net sample size	Response rate
Refreshment sample Wave 4	17,111	5,204	30.41%
ALWA sample	6,714	5,379	80.12%
Refreshment sample Wave 2	1,835	1,324	72.15%
Enhancement sample	2,841	2,197	77.33%
Total	28,501	14,104	49.49%

purpose, the same sampling procedure as for the refreshment sample of the initial sample of Starting Cohort 6 was applied. That is, the refreshment sample of Wave 4 was drawn within the 250 municipalities of the ALWA sample. At the end, 242 municipalities (with 273 sampling points allocated) provided information about their resident population. Per sampling point, from each register of a municipality, 63 addresses were drawn—resulting in a total of 17,111 addresses. Finally, 5,208 individuals gave their consent for participating in NEPS. Apart from this, all individuals who had already given their consent to attend in the studies of Starting Cohort 6 and who did not withdraw it or refuse further participation up to September 2011 were asked for an interview, that is, individuals who are part of the ALWA sample, the refreshment sample of Wave 2, or the enhancement sample. The gross sample of Wave 5 (i.e., of study B69) is composed by all individuals who gave their panel consent for taking part in NEPS and who did not refused before the onset of study B69, that is before September 2012. Between Wave 4 and 5, 1,349 individuals dropped out because of participation refusal or due to other reasons (e.g., moving abroad and dying). The Tables 3 and 4 give the gross and net sample sizes of Wave 4 and 5. Note that the sampling of the ALWA study, of Wave 2, and of Wave 4 had been conducted by the infas *Institut für angewandte Sozialwissenschaft GmbH*, see Aust et. al. (2011) and Aust et. al. (2013).

3 Derivation of Sampling Weights

Alike the sampling, the computation of sampling weights corresponding to Wave 2 to 5 inclusively nonresponse adjustments had been conducted by infas, cp. Aust et al. (2011), Aust et al. (2012), Aust et al. (2013) and Bech et al. (2014). In addition, infas calibrated the sampling weights of Wave 2 and 3 to external benchmark values taken from the Microcensus 2009 and 2010. The sampling weights of Wave 4 and 5 were calibrated to values of the Microcensus 2011 and 2012 by the NEPS method group.

Table 4: The gross and net sample sizes of Wave 5 (i.e., of study B69).

Subsamples	Gross sample size	Net sample size	Response rate
Refreshment sample Wave 4	4,964	3,716	74.86%
ALWA sample	6,196	4,880	78.76%
Refreshment sample Wave 2	1,603	1,166	72.74%
Enhancement sample	2,486	1,934	77.80%
Total	15,249	11,696	76.70%

3.1 Design Weights

For all considered subsamples, design weights are calculated as inverse sampling probabilities allowing to adjust the sampling design for disproportional stratification. That is, when assuming for an individual an inclusion probability π , its corresponding design weight is $1/\pi$. For all subsamples a stratified two stage sampling approach has been adopted. First, the target population had been stratified according to Federal States, administrative districts, and classification of urbanization (BIK scale), yielding a total of L strata. Then, sampling points had been allocated and thus municipalities had been selected. Finally, from the selected municipalities addresses are sampled on the basis of the number of sampling points allocated.

For the initial sample of Starting Cohort 6 and the refreshment sample of Wave 4, 250 municipalities (281 sampling points) had been sampled from a total of 12,429 German municipalities.⁷ For this purpose, s_l municipalities had been sampled proportional to size within each stratum l , $l = 1, \dots, L$. The measure of size (MOS) applied for this purpose is N_{m_l}/N_l , with N_{m_l} denoting the number of available addresses within municipality m in stratum l and N_l denotes the total number of addresses available in stratum l . Subsequently, s_{mlk} denotes the number of sampling points allocated to municipality m in stratum l in subsample k , and c_k the number of addresses drawn per sampling point in the subsample k . Thus, the sampling probability of an individual address i in stratum l in municipality m in subsample k is given as

$$\pi_{ilmk} = \frac{s_l N_{m_l}}{N_l} \times \frac{c_k s_{mlk}}{N_{m_l}} = \frac{c_k s_{mlk} s_l}{N_l} \approx \frac{c_k s_l}{N_l},$$

since s_{ml} is in general equal to one. By design, the sampling procedure of Starting Cohort 6 resembles a simple random sampling approach. In detail, the number s_l of municipalities sampled at the first stage is chosen such that $s_l \propto N_l/N$, where $N = 39,235,797$ is the total of the German resident population born between 1944 and 1986 at survey start. Thus, the sampling probability π_{ilmk} is (approximately) equal to $\pi = (\sum_{l,k} c_k s_l)/N = n/N$ with n denoting the number of all addresses that have overall been sampled.⁸

⁷For the sake of convenience, we consider the drop out among the 250 sampled municipalities—resulting in either a sample of 240 municipalities (refreshment and enhancement sample of Wave 2) or a sample of 242 municipalities (refreshment same of Wave 4)—as being completely at random.

⁸Due to the applied sampling procedure, the ALWA subsample and the refreshment sample of Wave 2 might overlap. This issue has been tackled by computing for all individuals who can be part of more

3.2 Cross-sectional and Longitudinal Weights

Up to now, for all individuals who have been selected to be part of Wave 2, design weights are computed. To account for nonresponse among these individuals, the design weights had to be adjusted accordingly.

3.2.1 Wave 2

In order to compute nonresponse adjusted sampling weights for individuals i who are part of the ALWA subsample, first the probability $W_{\pi_{i1}}$ of panel willingness and then the probability $P_{\pi_{i1}}$ of participation has to be derived. Thereafter, the nonresponse adjusted sampling weights w_{ilm1} can be computed as:

$$w_{ilm1} = w_{ilm}^{ALWA} \cdot (W_{\pi_{i1}} \cdot P_{\pi_{i1}})^{-1}.$$

Here, w_{ilm1}^{ALWA} denotes the original design weight of an individual being part of the ALWA subsample (i.e., $k = 1$). In other words, the weight w_{ilm1} is the cross-sectional weight of an individual of the ALWA subsample to participate in Wave 2. Logit regressions are used to estimate the probabilities $W_{\pi_{i1}}$ and $P_{\pi_{i1}}$. The set of covariates incorporated within the regression and resulting odds ratios are given in the Tables 7 and 8 in the Appendix. Overall, the regressions only point to modest selectivity concerning *educational attainment* and *income*. Individuals with a high level of education show a slightly higher probability to attend in the survey than individuals with a low educational level. Likewise, individuals with higher income are more willing to attend in the survey than individuals with lower income.

To derive sampling weights for all individuals i being part of the refreshment and the enhancement subsample, the probabilities $P_{\pi_{ik}}$ of the current participation have to be derived ($k = 2, 3$). The corresponding adjusted weights are

$$w_{ilmk} = (\pi_{ilmk} \cdot P_{\pi_{ik}})^{-1}.$$

with $k = 2, 3$. The weight w_{ilmk} corresponds to the cross-sectional weight of an individual attending Wave 2. Again, logit regressions are used to estimate the probabilities $P_{\pi_{i2}}$ and $P_{\pi_{i3}}$. The estimation results are given in Table 9 in the Appendix. Small selection effects can be observed related to *country of birth*. Furthermore, people born in the years from 1944 to 1955 have a slightly lower probability to attend in the survey than people born later.

Besides nonresponse adjustments, the weights of Wave 2 are calibrated to make the distribution of sample data concordant with known totals. Adjusting data to external population totals reduces the bias in the sampled data, but at the same time it tends to increase the variance in the data (i.e., the sampling error). This trade-off has to be regarded in the calibration process. To avoid any substantial enhancement of the sampling error, we adjust only few relevant marginal distributions of the sample of Starting Cohort

than one subsample design weights for each of the subsample of which they can be part. The individual design weights are computed as a linear combination minimizing the variance of an estimator for the total population number serving as a benchmark.

6. Calibration factors are determined using the so-called linear GREG estimation method, see Särndal & Lundström (2005) and Särndal (2007). This method allows specifying adjusted design weights as products of design weights and calibration factors. That is, for a sample unit i with adjusted weight w_{ilmk} and calibration factor g_i the calibrated weight is given as $w_{ilmk}^{\text{cal}} = g_i w_{ilmk}$. External benchmark distributions are taken from the German Microcensus 2009. Calibration factors are computed using marginal distributions for the following variable combinations:

- gender and educational attainment (according to ISCED97 categories) and
- birth year and educational attainment (according to ISCED97 categories).

The Tables 10 and 11 in the Appendix provide a comparison between sample distribution and reference distribution for the above mentioned benchmark variables. The observed differences can be gauged on the basis of the efficiency measure $E = \tilde{n}/n$ with n denoting the sample size and \tilde{n} the effective number of cases. The latter indicates the number of respondents that would have produced the same sampling error under a simple random sampling design (given the variance of the attributes accounted for in the calibration process). It can be computed as follows.⁹

$$\tilde{n} = \frac{(\sum_{i=1}^n g_i)^2}{\sum_{i=1}^n (g_i)^2}$$

In the considered setting, the efficiency measure is approximately 60 percent. Minding the multilevel weighting concept applied, with design weighting, nonresponse adjustment, and calibration, it can be considered as being good.

3.2.2 Wave 3

The longitudinal and cross-sectional weights for the attendance in Wave 3 are computed starting from the calibrated (cross-sectional) weights of attending Wave 2. For this purpose, two groups of participants need to be differentiated. The first group consists of all individuals who had already participated in the Wave 2, denoted as “repeaters”. The second group is made up by those individuals who attended the ALWA study, agreed to participate in NEPS, but finally refused in Wave 2, i.e. they did not dropped out ultimately. These individuals are called “temporary drop-outs”. The longitudinal weights ${}^R w_i^L$ of repeaters i are computed by means of their cross-sectional weights w_i of Wave 2 and their probability ${}^R \rho_i$ of participating in Wave 3:

$${}^R w_i^L = w_i \cdot {}^R \rho_i^{-1}.$$

A logistic regression model had been used to estimate the participation probabilities ${}^R \rho_i$ for all repeaters. In accordance therewith, all cases that had already participated in Wave 2 formed the basis of computation (in total, 11,362 cases). The parameters and results of the logistic regression analysis are shown in Table 12 in the Appendix. The longitudinal weights ${}^{TA} w_i^L$ of the temporary drop-outs i have been computed by means of the sampling weights w_i^{ALWA} of these cases attending the ALWA study, their probabilities ${}^W \pi_{i1}$ of panel

⁹For reasons of clarity, subsequently all indices related to stratum, municipality, and subsample are omitted.

willingness, their participation probabilities ${}^P\pi_{i1}$ of taking part in Wave 2, as well as their participation probabilities ${}^{TA}\rho_i$ of taking part in Wave 3:

$${}^{TA}w_i^L = w_i^{ALWA} \cdot ({}^W\pi_{i1} \cdot (1 - {}^P\pi_{i1}) \cdot {}^{TA}\rho_i)^{-1}.$$

Again, a logistic regression had been used to estimate the probabilities of temporary drop-outs to participate in Wave 3. In sum, the participation probabilities of 833 temporary drop-out cases had been modeled. The parameters and the results of this regression analysis are given in Table 13 in the Appendix. (The derivation of ${}^W\pi_{i1}$ and ${}^P\pi_{i1}$ is described in Section 3.2.1.) Now, the cross-sectional weights for participants in Wave 3 can be computed as

$$\begin{aligned} {}^Rw_i^C &= {}^Rw_i^L \cdot n_R / (n_R + n_{TA}) && \text{for repeaters and as} \\ {}^{TA}w_i^C &= {}^{TA}w_i^L \cdot n_{TA} / (n_R + n_{TA}) && \text{for temporary drop-outs,} \end{aligned}$$

where n_R is the number of repeaters and n_{TA} the number of temporary drop-out cases. Here, the panel attrition due to individuals who refuse to further participate is assumed to occur completely at random.

To make the distribution of sample data concordant with known totals, the cross-sectional weights of Wave 3 are calibrated to benchmark distributions taken from the German Microcensus 2010. Before, the weights have been trimmed at the 5th and 95th percentile in order to limit extreme outliers and thus also the variance of the weights. Calibration has then been conducted applying GREG estimation on the basis of the marginal distributions for the following variable combinations:

- gender and educational attainment (according to ISCED97 categories),
- birth year and educational attainment (according to ISCED97 categories),
- place of living (Federal State categories),
- BIK categories of municipality size,
- birth year and country of birth.

A comparison of the Microcensus distribution 2010 and the unweighted realized sample does not indicate any major differences; cp. Tables 14 to 19 given in the Appendix. Nevertheless, there are differences between the realized cases and the basic population, particularly pertaining to attributes of *country of birth* and *education*. These differences were equalized through the calibration procedure. The related efficiency measure E is close to 56 percent—which is acceptable in view of the complex sampling design.

3.2.3 Wave 4

The sample of Wave 4 comprises—besides the individuals who had already agreed to participate in the studies of Starting Cohort 6 in Wave 2 and who did not withdraw their panel consent up to September 2011—a refreshment sample of individuals who were born between 1944 and 1988. The sampling procedure applied to establish this refreshment

sample is identical to the one applied to establish the sample of Wave 2; see Section 3. In accordance therewith, the derivation of design weights of the refreshment sample corresponds to the derivation of design weights of Wave 2. To this end, Section 3.1 gives a comprehensive description. In sum, design weights are computed for the 17,111 individuals who are part of the gross sample of the refreshment sample. Note that an individual who is part of the refreshment sample of Wave 4 has a nonzero probability to be also part of the sample of Wave 2. To counteract this incoherence, design weights have been computed for both settings (i.e., for being part of the sample of Wave 2 and of the sample of Wave 4) and then linearly combined such that the variance of an estimator for the total population number becomes minimal; see also footnote 6. Clearly, not all individuals who had initially been sampled participated in the studies of Wave 4. This was accounted for by adjusting the design weights accordingly. For this purpose, participation probabilities had been estimated using logistic regression models. Table 20 (in the Appendix) shows the respective parameters and estimation results. On the basis of the estimated participation probabilities adjustment factors had been computed and being multiplied to the design weights.

Sampling weights of the panel sample of Wave 4 have been derived in a similar way as the sampling weights for the panel of sample Wave 3. First, two groups of participants are differentiated: repeaters and temporary drop-outs. Repeaters constitute those individuals who took part in Wave 3 and did not refuse up to September 2011. Likewise, the group of temporary drop-outs is made up by those individuals who did neither participate in Wave 3 nor refuse further participation. For repeaters, first the probability to not refuse has been modeled and then the probability to actually participate in the study. The results are given in the Tables 21 and 22 in the Appendix. The product of both probabilities gives the propensity of an individual to participate in Wave 3 and 4, and its inverse constitutes the accordant adjustment factor. That is, multiplied with the cross-sectional weight of Wave 3 it yields the cross-sectional weight of repeaters of Wave 4. The parameters and results of the logit regression analysis of temporary drop-outs are shown in Table 23 in the Appendix. The related inverse participation probabilities form the adjustment factors of temporary drop-out cases to temporarily drop-out in Wave 3 and to participate in Wave 4. By means of these adjustment factors, the temporary drop-outs' cross-sectional weights of Wave 2, and their non-participation probability of Wave 3 corresponding longitudinal weights can be derived. Combining the longitudinal weights of repeaters and temporary drop-outs as described for Wave 3 (cp. Section 3.2.2) allows deriving cross-sectional sampling weights for Wave 4.

To improve the representativeness of the sample, the cross-sectional weights have been calibrated to benchmark distributions taken the Microcensus 2011. To this end, the following marginal distributions have been considered:

- gender and educational attainment (according to ISCED97 categories),
- birth year and educational attainment (according to ISCED97 categories),
- place of living (Federal State categories),
- BIK categories of municipality size, as well as

- birth year and country of birth.

The Tables 24 to 29 in the Appendix contrast the corresponding distributions derived from the Microcensus 2011 data with the accordant distributions taken from the realized unweighted sample of Wave 4. The differences between the studied distributions are small. Nevertheless, calibration seems to be reasonable, in particular, with respect to *country of birth* and educational attainment. Ultimately, calibration resulted in an efficiency measure of approximately 60 percent, indicating reasonable design effects.

3.2.4 Wave 5

The procedure to compute longitudinal and cross-sectional weights for Wave 5 is equivalent to the one applied for the panel sample of Wave 4 as well as for the panel sample of Wave 3. That is, to specify the propensity of individuals to take part in Wave 5, repeaters and temporary drop-outs are distinguished, and related models describing the participation probabilities are estimated. These models allow deriving adjustment factors which then are used to calculate longitudinal and cross-sectional weights. (See Sections 3.2.1 and 3.2.2 for a detailed description of the related computation.) The parameters and results of the models estimated are given in the Tables 30, 31 and 32 in the Appendix.

Similarly to the Waves 2 to 4, the cross-sectional weights of Wave 5 were calibrated such that the weighted sample data matches with external benchmark distributions. The variables considered in this context are the same as in Wave 3 and 4 (cp. Section 3.2.1 and Section 3.2.2). For calibration the data of the Microcensus 2012 has been used. The Tables 33 to 38 in the Appendix show the comparison of the related distributions. Differences concerning the distribution of the *educational attainment* and the *country of birth* are revealed. Hence, calibration yields an efficiency measure of circa 36 percent which points to considerable design effects.

Table 5: Types of weights provided.

Type of weight	Label
Weights of individuals participating in Wave 2 (study B72)	w_t2
Weights of individuals participating in Wave 3 (study B67)	w_t3
Weights of individuals participating in Wave 4 (study B68)	w_t4
Weights of individuals participating in Wave 5 (study B69)	w_t5
Weights of individuals participating in Wave 2 and 3	w_t23
Weights of individuals participating in Wave 2, 3, and 4	w_t234
Weights of individuals participating in Wave 2, 3, 4, and 5	w_t2345
Weights of individuals participating in Wave 4 and 5	w_t45

4 Summary of Design Variables and Weights

The weights are provided ‘purely’ and—to ease statistical analysis—in a trimmed and standardized form. Trimming was conducted before the 5th and the after 95th percentile to

Table 6: Summary statistics for (calibrated and standardized) weights.

Label of weight	Number of students	Min.	Lower Quart.	Median	Mean	Upper Quart.	Max.
w_t2_std	11,649	0.309	0.576	0.907	1.000	1.252	2.453
w_t2_cal	11,649	0.116	0.483	0.769	1.000	1.185	6.869
w_t3_cal	9,320	0.064	0.416	0.720	1.000	1.233	11.813
w_t4_cal	14,104	0.000	0.414	0.842	1.000	1.260	4.023
w_t5_cal	11,696	0.000	0.217	0.462	1.000	1.074	5.280
w_t23_std	9,037	0.079	0.495	0.789	1.000	1.207	7.561
w_t234_std	7,901	0.088	0.515	0.810	1.000	1.223	8.219
w_t2345_std	6,820	0.064	0.524	0.823	1.000	1.235	8.812
w_t45_std	11,196	0.054	0.478	0.852	1.000	1.179	9.244

remove outliers. Standardization was performed to have weights with mean one. The standardized form of the weights are marked by the suffix `_std`. Weights which are additionally calibrated are labeled with the suffix `_cal`. Table 5 lists the types of weights provided for the SUF SC6 release version 5-0-0 and Table 7 gives some summary statistics of the (standardized) weights provided. Along with sampling weights, variables highlighting the sampling design are published. They are summarized in Table 6.

Table 7: Design variables provided.

Type of design information	Label
Primary Sampling Unit (Sampling Point Number)	psu
Identifier of stratum (implicit stratification)	stratum
Initial sample (ALWA, NEPS)	sample
Initial sample detailed (ALWA, NEPS enhancement, NEPS refreshment)	subsample
Federal state	tx80101
BIK classification	tx80102

5 Comments regarding the Usage of Weights

No general recommendation for the usage of sampling weights can be given. Whether and how weights have to be used depends on the problem to be studied. Commonly, it is recommended to apply sampling weights when conducting descriptive statistics. For analytical analysis, models have to be tested for their dependence on the sampling design. Concretely, this means that the user has to ensure that the way of sampling has no or only a negligible effect on the model results or/and that the sampling design is considered in the model definition adequately. A general description of how to test and account for the sampling design is given in, for example, Snijder and Bosker (2012). As a guideline, we recommend to include all variables employed for constructing the (used set of) weights as explanatory variables into the model under consideration.

References

- [1] Antoni, M., Drasch, K., Kleinert, C. Matthes, B., Ruland, M. & Trahms, A. (2010). *Arbeiten und Lernen im Wandel, Teil I: Überblick über die Studie.* (FDZ-Methodenreport 5/2010), Nürnberg: Forschungsdatenzentrum (FDZ) der Bundesagentur für Arbeit im Institut für Arbeitsmarkt- und Berufsforschung.
- [2] Aust, F., Gilberg, R., Hess, D., Kleudgen, M. & Steinwede, A. (2011), *Methodenbericht NEPS Etappe 8: Befragung von Erwachsenen - Haupterhebung 1. Welle 2009/2010*, infas Institut für angewandte Sozialwissenschaft GmbH.
- [3] Aust, F., Gilberg, R., Hess, D., Kersting, A., Kleudgen, M. & Steinwede, A. (2012), *Methodenbericht NEPS Etappe 8: Befragung von Erwachsenen - Haupterhebung 2. Welle (B67)*, infas Institut für angewandte Sozialwissenschaft GmbH.
- [4] Aust, F., Hess, D., Kleudgen, M., Malina, A. & Steinwede, A. (2013), *Methodenbericht NEPS Etappe 8: Befragung von Erwachsenen - Haupterhebung 3. Welle 2011/2012 (B68)*, infas Institut für angewandte Sozialwissenschaft GmbH.
- [5] Bech, K., Hess, D., Kleudgen, & Steinwede, A. (2014), *Methodenbericht NEPS Etappe 8: Befragung von Erwachsenen - Haupterhebung 4. Welle 2012/2013 (B69)*, infas Institut für angewandte Sozialwissenschaft GmbH.
- [6] Särndal, C.E., & Lundström, S. (2005). *Estimation in surveys with nonresponse.* New York: Wiley.
- [7] Särndal, C.E. (2007). *The calibration approach in survey theory and practice.* *Survey Methodology*, Vol. 33 (2), pp. 99-119.
- [8] Snijder, T., and Bosker, R. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling.* In (2nd ed., p. 216-246). Sage Publications.

A Results of Nonresponse Modeling

Table 7: Results of the logit regression model measuring the panel willingness of participants of the ALWA survey.

Variable	Reference Category	Odds Ratio	P-Value
<i>Birth year</i>	1980 – 1986		
1956 – 1969		1.05	0.73
1970 – 1979		1.02	0.86
<i>Gender</i>	female		
male		0.99	0.93
<i>Country of birth</i>	born in Germany		
born abroad		0.72	0.06
<i>Mother tongue</i>	Non-German		
German		1.22	0.28
<i>Marital status</i>	unmarried		
married		1.03	0.84
separated		1.89	0.00
widowed		2.34	0.16
<i>Household size</i>	three and more persons		
one person		1.30	0.08
two persons		1.08	0.47
<i>School qualification</i>	'Realschule'		
'Hauptschule'		0.92	0.41
upper secondary education		1.03	0.75
other		0.61	0.01
<i>School qualification parents</i>	'Realschule'		
'Hauptschule'		0.91	0.35
upper secondary education		1.23	0.09
other		0.51	0.00
<i>Income</i>	1,501 – 3,500 Euro		
up to 1,500 Euro		0.80	0.08
more than 3,500 Euro		1.88	0.00
<i>Federal state</i>	Nordrhein-Westfalen		
Schleswig-Holstein		1.14	0.61
Hamburg		0.99	0.99
Niedersachsen		0.96	0.76
Bremen		0.95	0.92
Hessen		1.04	0.79
Rheinland-Pfalz		1.21	0.35
Baden-Württemberg		1.02	0.86
Bayern		0.81	0.09
Saarland		0.90	0.75
Berlin		0.94	0.79
Brandenburg		1.32	0.30
Mecklenburg-Vorpommern		0.91	0.77
Sachsen		1.08	0.70
Sachsen-Anhalt		1.38	0.25
Thüringen		1.49	0.18
Pseudo R^2	0.03		
Number of cases	10,404		

Table 8: Results of logit regression model measuring the participation probability of individuals of the ALWA subsample.

Variable	Reference Category	Odds Ratio	P-Value
<i>Birth year</i>	1980 – 1986		
1956 – 1969		1.38	0.00
1970 – 1979		1.34	0.00
<i>Gender</i>	female		
male		1.08	0.12
<i>Country of birth</i>	born in Germany		
born abroad		0.76	0.03
<i>Mother tongue</i>	Non-German		
German		1.46	0.01
<i>Marital status</i>	unmarried		
married		1.20	0.03
separated		1.09	0.42
widowed		1.09	0.77
<i>Household size</i>	three persons and more		
one person		0.87	0.11
two persons		0.89	0.07
<i>School qualification</i>	'Realschule'		
'Hauptschule'		0.87	0.06
upper secondary education		1.43	0.00
other		0.93	0.62
<i>School qualification parents</i>	'Realschule'		
'Hauptschule'		1.12	0.09
upper secondary education		1.12	0.12
other		0.83	0.11
<i>Income</i>	1,501 – 3,500 Euro		
up to 1,500 Euro		0.82	0.03
more than 3,500 Euro		1.01	0.85
<i>Federal State</i>	Nordrhein-Westfalen		
Schleswig-Holstein		0.87	0.35
Hamburg		1.35	0.15
Niedersachsen		0.92	0.38
Bremen		0.85	0.60
Hessen		0.94	0.59
Rheinland-Pfalz		0.95	0.66
Baden-Württemberg		0.92	0.37
Bayern		1.02	0.78
Saarland		1.08	0.73
Berlin		0.96	0.80
Brandenburg		0.82	0.20
Mecklenburg-Vorpommern		1.16	0.52
Sachsen		0.97	0.79
Sachsen-Anhalt		0.75	0.06
Thüringen		1.26	0.17
<i>BIK categories</i>	500,000 and more inhab. (styp 1)		
less than 2000 inhab.		1.24	0.28
2000 – 5000 inhab.		1.08	0.64
5000 – 20,000 inhab.		1.02	0.88
20,000 – 50,000 inhab.		1.10	0.34

50,000 – 100,000 inhab. (styp 2/3/4)		1.24	0.06
50,000 – 100,000 inhab. (styp 1)		0.97	0.89
100,000 – 500,000 inhab. (styp 2/3/4)		0.97	0.76
100,000 – 500,000 inhab. (styp 1)		0.86	0.08
500,000 and more inhab. (styp 2/3/4)		0.97	0.77
<i>Attempts to contact target</i>	1 to 3 attempts		
4 to 6 attempts		1.04	0.63
7 to 10 attempts		0.97	0.69
more than 10 attempts		0.35	0.00
Pseudo R^2	0.07		
Number of cases	8,997		

Table 9: Results of logit regression model measuring the participation probability of the refreshment sample and of the additional sample.

Variable	Reference Category	Odds Ratio	P-Value
<i>Birth year</i>	1980 – 1988		
1944 – 1955		0.83	0.00
1956 – 1969		0.98	0.78
1970 – 1979		0.96	0.66
<i>Gender</i>	female		
male		0.95	0.15
<i>Country of birth</i>	born in Germany		
born abroad		0.52	0.00
<i>Federal state</i>	Nordrhein-Westfalen		
Schleswig-Holstein		0.88	0.24
Hamburg		0.95	0.67
Niedersachsen		1.04	0.58
Bremen		0.90	0.62
Hessen		1.02	0.77
Rheinland-Pfalz		0.89	0.19
Baden-Württemberg		0.93	0.24
Bayern		0.98	0.79
Saarland		1.11	0.48
Berlin		0.97	0.72
Brandenburg		0.93	0.47
Mecklenburg-Vorpommern		0.80	0.12
Sachsen		1.19	0.04
Sachsen-Anhalt		0.94	0.56
Thüringen		0.92	0.50
<i>BIK categories</i>	500,000 and more inhab. (styp 1)		
less than 2000 inhab.		1.38	0.03
2000 – 5000 inhab.		0.81	0.08
5000 – 20,000 inhab.		1.09	0.24
20,000 – 50,000 inhab.		1.13	0.05
50,000 – 100,000 inhab. (styp 2/3/4)		1.15	0.06
50,000 – 100,000 inhab. (styp 1)		1.10	0.44
100,000 – 500,000 inhab. (styp 2/3/4)		0.99	0.89
100,000 – 500,000 inhab. (styp 1)		0.91	0.13
500,000 and more inhab. (styp 2/3/4)		1.20	0.01

<i>Attempts to contact target</i>	1 to 3 attempts		
5 to 6 attempts		1.46	0.00
7 to 10 attempts		1.25	0.00
more than 10 attempts		0.72	0.00
Pseudo R^2	0.02		
Number of cases	18,012		

Table 10: Sample and reference distribution according to gender and educational attainment.

	actual distribution net sample				target distribution population (Microcensus 2009)	
	re-freshment	addi-tional sample	panel sample	total	%	total
Gender and education	%	%	%	%	%	total
male						
ISCED 1	1.32	0.97	0.33	0.67	1.50	712,401
ISCED 2	3.70	3.03	1.42	2.23	4.63	2,194,902
ISCED 3ca	3.40	2.16	3.15	2.93	2.54	1,203,307
ISCED 3b	16.44	21.12	17.16	18.10	23.92	11,343,006
ISCED 4ab	4.46	2.19	4.85	4.08	3.32	1,573,744
ISCED 5b	5.58	8.18	6.33	6.70	5.16	2,446,774
ISCED 5a	10.81	12.01	14.09	12.98	8.29	3,932,478
ISCED 6	1.07	1.16	1.57	1.37	0.84	396,103
female						
ISCED 1	1.47	1.48	0.30	0.82	1.8	853,680
ISCED 2	7.56	9.05	2.51	5.11	6.81	3,231,635
ISCED 3ca	4.57	2.41	2.30	2.71	2.12	1,007,536
ISCED 3b	22.83	23.34	22.47	22.77	23.77	11,270,789
ISCED 4ab	6.24	1.87	8.00	6.07	4.18	1,982,235
ISCED 5b	0.81	1.8	1.16	1.27	3.88	1,841,603
ISCED 5a	8.93	8.73	13.54	11.48	6.84	3,246,127
ISCED 6	0.81	0.52	0.81	0.73	0.4	187,680
Total	100.00	100.00	100.00	100.00	100.00	47,424,000

Table 11: Sample and reference distribution according to birth year and educational attainment.

	actual distribution net sample				target distribution population (Microcensus 2009)	
	re-freshment	addi-tional sample	panel sample	total	%	total
Birth year and education	%	%	%	%	%	total
1975 – 1986						
ISCED 1	1.12	-	0.23	0.32	0.76	360,672
ISCED 2	4.67	0.03	1.13	1.43	2.87	1,362,317
ISCED 3ca	4.52	-	3.83	2.93	2.96	1,405,517
ISCED 3b	10.1	0.1	6.76	5.55	9.65	4,578,228

ISCED 4ab	4.11	0.03	4.26	3.11	2.97	1,407,526
ISCED 5b	0.76	-	0.96	0.67	1.49	706,275
ISCED 5a	6.14	-	6.35	4.62	3.7	1,756,143
ISCED 6	0.05	-	0.35	0.21	0.15	69,322
1965 – 1974						
ISCED 1	0.66	-	0.2	0.22	0.89	421,422
ISCED 2	3.45	-	0.96	1.12	2.51	1,188,010
ISCED 3ca	1.67	-	0.68	0.67	0.67	316,067
ISCED 3b	14.71	0.13	13.72	10.27	12.17	5,773,486
ISCED 4ab	3.7	-	3.91	2.83	2.24	1,064,593
ISCED 5b	2.69	-	2.91	2.09	2.48	1,176,972
ISCED 5a	6.9	0.06	8.98	6.25	4.27	2,024,834
ISCED 6	0.66	-	0.93	0.64	0.39	182,616
1956 – 1964						
ISCED 1	0.91	-	0.21	0.27	0.81	382,079
ISCED 2	3.09	0.06	1.84	1.58	2.65	1,257,552
ISCED 3ca	1.73	-	0.93	0.82	0.57	271,768
ISCED 3b	14.1	0.13	19.13	13.21	12.58	5,965,853
ISCED 4ab	2.84	0.06	4.69	3.14	1.53	726,051
ISCED 5b	2.94	0.16	3.61	2.58	2.59	1,229,473
ISCED 5a	6.49	0.1	12.31	8.07	3.54	1,680,748
ISCED 6	1.12	-	1.1	0.81	0.36	168,476
1944 – 1955						
ISCED 1	0.1	2.45	-	0.67	0.85	401,908
ISCED 2	0.05	11.98	-	3.2	3.41	1,618,658
ISCED 3ca	0.05	4.57	-	1.23	0.46	217,491
ISCED 3b	0.36	44.11	0.03	11.84	13.28	6,296,228
ISCED 4ab	0.05	3.96	-	1.06	0.75	357,809
ISCED 5b	-	9.82	0.02	2.63	2.48	1,175,657
ISCED 5a	0.2	20.57	-	5.52	3.62	1,716,880
ISCED 6	0.05	1.67	-	0.45	0.34	163,369
Total	100.00	100.00	100.00	100.00	100.00	47,424,000

Table 12: Results of the logit regression model measuring the participation propensity of repeaters in Wave 3.

Variable	Reference	Odds Ratio	P-Value
<i>Birth year</i>	1980 – 1986		
1970 – 1979		1.20	0.06
1956 – 1969		1.38	0.00
1944 – 1955		1.04	0.74
<i>Gender</i>	female		
male		1.04	0.41
<i>Country of birth</i>	born in Germany		
born abroad		0.87	0.30
<i>Mother tongue</i>	Non- German		
German		1.39	0.02
<i>Marital status</i>	unmarried		
married		1.12	0.16
separated		1.21	0.07

widowed		1.20	0.30
<i>Household size</i>	three and more		
one person		0.88	0.15
two persons		0.89	0.06
<i>School qualification</i>	'Realschule'		
'Hauptschule'		0.80	0.00
upper secondary education		1.36	0.00
other		1.17	0.13
<i>Secondary school qualification of parents</i>	'Realschule'		
'Hauptschule'		1.19	0.01
upper secondary education		1.10	0.18
other		1.11	0.68
<i>Income</i>	1501 – 3500 Euro		
up to 1500 Euro		0.92	0.28
more than 3500 Euro		1.05	0.40
<i>Federal State</i>	Nordrhein-Westfalen		
Schleswig-Holstein		1.25	0.17
Hamburg		1.19	0.37
Niedersachsen		1.01	0.91
Bremen		1.29	0.41
Hessen		1.03	0.74
Rheinland-Pfalz		1.08	0.54
Baden-Württemberg		1.12	0.22
Bayern		1.20	0.03
Saarland		1.12	0.60
Berlin		0.90	0.44
Brandenburg		1.16	0.35
Mecklenburg-Vorpommern		0.81	0.29
Sachsen		1.29	0.05
Sachsen-Anhalt		1.61	0.01
Thüringen		1.26	0.16
<i>BIK categories</i>	500,000 and more inh. (styp 1)		
less than 2000 inhab.		1.38	0.14
2000 – 5000 inhab.		1.10	0.59
5000 – 20,000 inhab.		1.10	0.39
20,000 – 50,000 inhab.		1.06	0.55
50,000 – 100,000 inhab. (styp 2/3/4)		1.15	0.21
50,000 – 100,000 inhab. (styp 1)		1.19	0.40
100,000 – 500,000 inhab. (styp 2/3/4)		1.00	0.99
100,000 – 500,000 inhab. (styp 1)		0.99	0.94
more than 500,000 inhab. (styp 2/3/4)		0.86	0.13
<i>Attempts to contact target</i>	1 to 3 attempts		
4 to 6 attempts		0.79	0.00
7 to 10 attempts		0.39	0.00
more than 10 attempts		0.15	0.00
Pseudo R2	0.10		
Number of cases	11,362		

Table 13: Results of the logit regression model measuring the participation propensity of individuals who participated in Wave 3 but not in Wave 2.

Variable	Reference	Odds Ratio	P-Value
<i>Birth year</i>	1980 – 1986		
1970 – 1979		1.18	0.47
1944 – 1969		1.13	0.53
<i>Gender</i>	female		
male		1.04	0.79
<i>Country of birth</i>	born in Germany		
born abroad		0.80	0.43
<i>Federal State</i>	Nordrhein-Westfalen		
Schleswig-Holstein		0.57	0.22
Hamburg		0.31	0.15
Niedersachsen		1.40	0.28
Bremen		5.00	0.24
Hessen		0.88	0.71
Rheinland-Pfalz		0.61	0.24
Baden-Württemberg		0.70	0.21
Bayern		0.80	0.38
Saarland		1.33	0.67
Berlin		0.74	0.53
Brandenburg		0.45	0.13
Mecklenburg-Vorpommern		1.27	0.78
Sachsen		0.94	0.89
Sachsen-Anhalt		0.38	0.05
Thüringen		0.77	0.65
<i>BIK categories</i>	500,000 and more inh. (styp 1)		
less than 2000 inhab.		2.16	0.23
2000 to 5000 inhab.		1.37	0.51
5000 to 20,000 inhab.		1.03	0.93
20,000 to 50,000 inhab.		1.75	0.07
50,000 to 100,000 inhab. (styp 2/3/4)		3.04	0.00
50,000 to 100,000 inhab. (styp 1)		1.88	0.36
100,000 to 500,000 inhab. (styp 2/3/4)		1.55	0.12
100,000 to 500,000 inhab. (styp 1)		1.22	0.46
more than 500,000 inhab. (styp 2/3/4)		1.38	0.33
<i>Attempts to contact target</i>	1 to 3 attempts		
4 to 6 attempts		0.86	0.45
7 to 10 attempts		0.54	0.01
more than 10 attempts		0.15	0.00
Pseudo R2	0.11		
Number of cases	833		

Table 14: Comparison of the distribution of the Wave 3 sample data and the target distribution (Mikrocensus 2010) according to gender and educational attainment.

Gender and education	actual distribution	target distribution	
	net sample	population (Mikrocensus 2010)	
	%	%	total
male			
ISCED 1	0.48	1.58	744,484
ISCED 2	1.92	4.44	2,095,599
ISCED 3ca	2.74	1.94	918,490
ISCED 3b	17.58	24.06	1,364,786
ISCED 4ab	4.22	3.39	1,601,706
ISCED 5b	6.95	5.48	2,590,162
ISCED 5a	13.91	8.36	3,948,233
ISCED 6	1.45	0.88	415,862
female			
ISCED 1	0.63	1.89	892,575
ISCED 2	4.66	6.57	3,102,092
ISCED 3ca	2.40	1.61	762,387
ISCED 3b	22.42	24.10	11,382,921
ISCED 4ab	6.28	4.24	2,002,132
ISCED 5b	1.26	4.03	1,901,064
ISCED 5a	12.26	6.97	3,291,538
ISCED 6	0.84	0.45	211,969
Total	100.00	100.00	47,266,000

Table 15: Comparison of the distribution of the Wave 3 sample data and the target distribution (Microcensus 2010) according to birth year and educational attainment.

Birth year and education	actual distribution	target distribution	
	net sample	population (Microcensus 2010)	
	%	%	total
1975 – 1986			
ISCED 1	0.28	0.80	375,452
ISCED 2	1.15	2.65	1,250,839
ISCED 3ca	2.37	2.38	1,123,346
ISCED 3b	5.25	9.78	4,618,870
ISCED 4ab	3.00	3.05	1,441,577
ISCED 5b	0.76	1.71	807,122
ISCED 5a	5.17	4.07	1,921,433
ISCED 6	0.25	0.20	93,361
1965 – 1974			
ISCED 1	0.10	0,94	441,947
ISCED 2	1.08	2.46	1,161,747
ISCED 3ca	0.65	0.52	246,645
ISCED 3b	10.42	12.31	5,815,781
ISCED 4ab	2.86	2.25	1,064,096
ISCED 5b	2.22	2.60	1,227,183
ISCED 5a	6.77	4.22	1,994,299
ISCED 6	0.69	0.41	1,95,302

1956 – 1964			
ISCED 1	0.25	0,85	399,636
ISCED 2	1.53	2.53	1,194,871
ISCED 3ca	0.84	0.40	1,190,735
ISCED 3b	13.85	12.74	6,014,722
ISCED 4ab	3.54	1.56	735,693
ISCED 5b	2.65	2.71	1,277,624
ISCED 5a	8.86	3.53	1,669,186
ISCED 6	0.90	0.36	168,533
1944 – 1955			
ISCED 1	0.49	0.89	420,024
ISCED 2	2.81	3.37	1,590,234
ISCED 3ca	1.28	0.25	120,151
ISCED 3b	10.49	13.34	6,298,334
ISCED 4ab	1.08	0.77	362,472
ISCED 5b	2.57	2.50	1,179,297
ISCED 5a	5.37	3.50	1,654,853
ISCED 6	0.45	0.36	170,635
Total	100.00	100.00	47,226,000

Table 16: Comparison of the distribution of the Wave 3 sample data and the target distribution (Microcensus 2010) according to Federal State.

Federal State	actual distribution	target distribution	
	net sample	population (Microcensus 2010)	
	%	%	total
Schleswig-Holstein	2.99	3.37	1,593,000
Hamburg	2.04	2.30	1,085,000
Niedersachsen	10.28	9.50	4,487,000
Bremen	0.62	0.82	388,000
Nordrhein-Westfalen	22.38	21.62	10,211,000
Hessen	7.82	7.46	3,522,000
Rheinland-Pfalz	4.86	4.84	2,284,000
Baden-Württemberg	12.29	12.95	6,118,000
Bayern	15.48	15.40	727,000
Saarland	1.51	1.25	588,000
Berlin	3.51	4.46	2,108,000
Brandenburg	3.25	3.20	1,509,000
Mecklenburg-Vorpommern	1.51	2.07	979,000
Sachsen	5.54	5.07	2,394,000
Sachsen-Anhalt	3.05	2.88	1,358,000
Thüringen	2.86	2.82	1,330,000
Total	100.00	100.00	47,226,000

Table 17: Comparison of the distribution of the Wave 3 sample data and the target distribution (Microcensus 2010) according to BIK categories of municipal size.

BIK categories	actual distribution	target distribution	
	net sample	population (Microcensus 2010)	
	%	%	total
less than 2000 inhab.	2.17	1.92	909,000
2000 to 5000 inhab.	2.68	2.76	1,304,000
5000 to 20,000 inhab.	8.05	7.81	3,686,000
20,000 to 50,000 inhab.	12.39	11.43	5,399,000
50,000 to 100,000 inhab. styp 2/3/4	9.10	7.82	3,692,000
50,000 to 100,000 inhab. styp 1	2.02	2.23	1,055,000
100,000 to 500,000 inhab. styp 2/3/4	15.72	14.84	7,007,000
100,000 to 500,000 inhab. styp 1	15.48	16.16	7,630,000
500,000 and more inhab. styp 2/3/4	8.41	9.08	4,288,000
500,000 and more inh. styp 1	23.99	25.95	12,256,000
Total	100.00	100.00	47,226,000

Table 18: Comparison of the distribution of the Wave 3 sample data and the target distribution (Microcensus 2010) according to birth year.

Year of birth	actual distribution	target distribution	
	net sample	population (Microcensus 2010)	
	%	%	total
1944	1.72	1.95	919,000
1945	1.43	1.42	671,000
1946	1.64	1.69	797,000
1947	1.83	1.89	892,000
1948	1.75	2.03	957,000
1949	2.31	2.17	1,023,000
1950	2.10	2.25	1,062,000
1951	2.29	2.26	1,065,000
1952	2.44	2.28	1,075,000
1953	2.11	2.30	1,087,000
1954	2.64	2.38	1,125,000
1955	2.31	2.38	1,123,000
1956	3.30	2.48	1,170,000
1957	3.11	2.56	1,210,000
1958	3.27	2.57	1,215,000
1959	4.14	2.69	1,272,000
1960	3.80	2.80	1,323,000
1961	3.48	2.82	1,332,000
1962	3.80	2.80	1,323,000
1963	3.68	2.94	1,389,000
1964	3.84	3.00	1,417,000
1965	3.89	3.02	1,428,000
1966	3.45	3.11	1,470,000
1967	2.97	2.94	1,388,000
1968	2.84	2.83	1,336,000
1969	2.48	2.71	1,278,000
1970	2.40	2.59	1,221,000

1971	1.97	2.41	1,139,000
1972	1.91	2.18	1,031,000
1973	1.49	1.98	933,000
1974	1.37	1.95	923,000
1975	1.32	1.97	931,000
1976	1.23	1.99	940,000
1977	1.46	2.01	950,000
1978	1.35	2.04	962,000
1979	1.47	2.03	957,000
1980	1.38	2.18	1,031,000
1981	1.37	2.12	1,003,000
1982	1.46	2.15	1,013,000
1983	1.71	2.10	991,000
1984	1.46	2.02	953,000
1985	1.65	1.98	935,000
1986	2.36	2.05	966,000
Total	100.00	100.00	XXX

Table 19: Comparison of the distribution of the Wave 3 sample data and the target distribution (Microcensus 2010) according to country of birth.

Country of birth	actual distribution	target distribution	
	net sample	population (Microcensus 2010)	
	%	%	total
born abroad	8.30	17.48	8,257,000
born in Germany	91.70	82.52	38,969,000
Total	100.00	100.00	47,226,000

Table 20: Results of the logit regression model measuring the participation propensity of individuals of the refreshment sample of Wave 4.

Variable	Reference	Odds Ratio	P-Value
<i>Birth Year</i>	1980 – 1988		
1970 – 1979		1.02	0.75
1956 – 1969		1.12	0.04
1944 – 1955		1.14	0.02
<i>Gender</i>	female		
male		0.89	0.00
<i>Country of birth</i>	born in Germany		
born abroad		0.49	0.00
not specified		0.76	0.00
<i>Federal State</i>	Nordrhein-Westfalen		
Schleswig-Holstein		0.99	0.95
Hamburg		0.81	0.11
Niedersachsen		1.07	0.29

Bremen		0.89	0.59
Hessen		0.92	0.24
Rheinland-Pfalz		0.89	0.19
Baden- Württemberg		0.89	0.06
Bayern		1.01	0.83
Saarland		0.79	0.12
Berlin		0.81	0.03
Brandenburg		0.86	0.14
Mecklenburg-Vorpommern		0.81	0.10
Sachsen		0.93	0.44
Sachsen-Anhalt		0.90	0.33
Thüringen		1.13	0.26
<i>BIK categories</i>	500,000 and more inh. (styp 1)		
less than 2000 inhab.		1.47	0.01
2000 – 5000 inhab.		0.95	0.66
5000 – 20,000 inhab.		1.26	0.00
20,000 – 50,000 inhab.		1.16	0.03
50,000 – 100,000 inhab. (styp 2/3/4)		1.20	0.01
50,000 – 100,000 inhab. (styp 1)		0.98	0.89
100,000 – 500,000 inhab. (styp 2/3/4)		1.20	0.00
100,000 – 500,000 inhab. (styp 1)		1.08	0.21
more than 500,000 inhab. (styp 2/3/4)		1.16	0.04
<i>Attempts to contact target</i>	1 to 3 attempts		
4 to 6 attempts		1.14	0.00
7 to 10 attempts		1.15	0.01
more than 10 attempts		0.86	0.00
Pseudo R2	0.01		
Number of cases	17,111		

Table 21: Results of the logit regression model measuring the participation willingness of repeaters in Wave 4.

Variable	Reference	Odds Ratio	P-Value
<i>Birth year</i>	1980 – 1986		
1970 – 1979		1.13	0.31
1956 – 1969		1.18	0.16
1944 – 1955		2.84	0.00
<i>Gender</i>	female		
male		0.97	0.72
<i>Country of birth</i>	born in Germany		
born abroad		0.97	0.86
<i>Mother tongue</i>	Non-German		
German		1.01	0.98
<i>Marital status</i>	unmarried		
married		3.38	0.00
separated		1.72	0.00
widowed		1.82	0.05
<i>Household size</i>	three and more		
one person		1.03	0.82

two persons		1.00	0.98
<i>School qualification</i>	'Realschule'		
'Hauptschule'		0.98	0.87
upper secondary education		1.38	0.00
other		1.56	0.16
<i>School qualification of parents</i>	'Realschule'		
'Hauptschule'		0.85	0.09
upper secondary education		0.97	0.76
other		0.17	0.00
<i>Income</i>	1.501 – 3500 Euro		
up to 1500 Euro		2.08	0.00
more than 3500 Euro		1.23	0.04
<i>Federal State</i>	Nordrhein-Westfalen		
Schleswig-Holstein		0.78	0.26
Hamburg		1.08	0.79
Niedersachsen		1.03	0.87
Bremen		0.91	0.84
Hessen		0.79	0.12
Rheinland-Pfalz		0.92	0.64
Baden-Württemberg		0.85	0.22
Bayern		0.88	0.31
Saarland		1.14	0.71
Berlin		0.86	0.48
Brandenburg		0.98	0.94
Mecklenburg-Vorpommern		0.98	0.94
Sachsen		0.93	0.70
Sachsen-Anhalt		0.72	0.14
Thüringen		1.33	0.30
Pseudo R2	0.07		
Number of cases	12,195		

Note: At the end of the studies of Wave 3, the panel sample of Starting Cohort 6 comprised 12,195 cases who were willing to further participate in NEPS. In the beginning of the studies of Wave 4, this number reduced to 11,390.

Table 22: Results of the logit regression model measuring the participation propensity of repeaters in Wave 4.

Variable	Reference	Odds Ratio	P-Value
<i>Birth year</i>	1980 – 1986		
1970 – 1979		1.12	0.35
1956 – 1969		1.47	0.00
1944 – 1955		1.08	0.55
<i>Gender</i>	female		
male		1.02	0.80
<i>Country of birth</i>	born in Germany		
born abroad		1.16	0.38
<i>Mother tongue</i>	Non-German		
German		1.53	0.02
<i>Marital status</i>	unmarried		
married		1.22	0.09
separated		1.10	0.50
widowed		1.52	0.14

<i>Household size</i>	three and more		
one person		1.13	0.34
two persons		1.01	0.90
<i>School qualification</i>	'Realschule'		
'Hauptschule'		0.75	0.00
upper secondary education		1.25	0.01
other		1.27	0.09
<i>School qualification of parents</i>	'Realschule'		
'Hauptschule'		0.95	0.53
upper secondary education		0.99	0.89
other		0.59	0.07
<i>Income</i>	1.501 – 3500 Euro		
up to 1500 Euro		0.86	0.14
more than 3500 Euro		1.24	0.01
<i>Federal State</i>	Nordrhein-Westfalen		
Schleswig-Holstein		0.93	0.69
Hamburg		0.82	0.40
Niedersachsen		0.89	0.36
Bremen		1.31	0.55
Hessen		1.18	0.27
Rheinland-Pfalz		0.97	0.88
Baden-Württemberg		0.81	0.06
Bayern		0.85	0.15
Saarland		0.93	0.78
Berlin		1.28	0.24
Brandenburg		0.87	0.50
Mecklenburg-Vorpommern		1.50	0.20
Sachsen		0.93	0.65
Sachsen-Anhalt		1.09	0.68
Thüringen		1.42	0.14
<i>BIK categories</i>	500,000 and more inh. (styp 1)		
less than 2000 inhab.		0.66	0.07
2000 to 5000 inhab.		1.29	0.25
5000 to 20,000 inhab.		1.25	0.14
20,000 to 50,000 inhab.		0.84	0.14
50,000 to 100,000 inhab. (styp 2/3/4)		1.10	0.51
50,000 to 100,000 inhab. (styp 1)		0.93	0.76
100,000 to 500,000 inhab. (styp 2/3/4)		1.06	0.60
100,000 to 500,000 inhab. (styp 1)		1.00	0.99
more than 500,000 inhab. (styp 2/3/4)		1.11	0.45
<i>Attempts to contact target</i>	1 to 3 attempts		
4 to 6 attempts		1.11	0.36
7 to 10 attempts		0.74	0.01
more than 10 attempts		0.18	0.00
Pseudo R2	0.12		
Number of cases	9,321		

Table 23: Results of the logit regression model measuring the participation propensity of individuals who participated in Wave 4 but not in Wave 3.

Variable	Reference	Odds Ratio	P-Value
<i>Birth year</i>	1980 – 1986		
1970 – 1979		1.25	0.15
1956 – 1969		1.13	0.38
1944 – 1955		0.96	0.78
<i>Gender</i>	female		
male		1.07	0.43
<i>Country of birth</i>	born in Germany		
born abroad		0.67	0.01
<i>Federal State</i>	Nordrhein-Westfalen		
Schleswig-Holstein		1.05	0.87
Hamburg		1.07	0.86
Niedersachsen		1.56	0.01
Bremen		0.79	0.72
Hessen		1.29	0.17
Rheinland-Pfalz		1.21	0.40
Baden-Württemberg		0.93	0.66
Bayern		0.93	0.64
Saarland		0.64	0.30
Berlin		1.75	0.02
Brandenburg		0.74	0.32
Mecklenburg-Vorpommern		1.30	0.49
Sachsen		0.85	0.51
Sachsen-Anhalt		1.33	0.43
Thüringen		1.19	0.56
<i>BIK categories</i>	500,000 and more inhab. Styp 1		
less than 2000 inhab.		1.37	0.48
2000 to 5000 inhab.		1.26	0.45
50000 to 20.000 inhab.		0.93	0.75
20,000 to 50,000 inhab.		1.08	0.66
50,000 to 100,000 inhab. (styp 2/3/4)		0.70	0.10
50,000 to 100,000 inhab. (styp 1)		0.66	0.30
100,000 to 500,000 inhab. (styp 2/3/4)		0.96	0.82
100,000 to 500,000 inhab. (styp 1)		1.08	0.62
more than 500,000 inhab. (styp 2/3/4)		0.91	0.64
<i>Attempts to contact target</i>	1 to 3 attempts		
4 to 6 attempts		1.11	0.44
7 to 10 attempts		1.42	0.02
more than 10 attempts		0.66	0.00
Pseudo R2	0.03		
Number of cases	2069		

Table 24: Comparison of the distribution of the Wave 4 sample data and the target distribution (Microcensus 2011) according to gender and educational attainment.

Gender and education	actual distribution	target distribution	
	net sample	population (Microcensus 2011)	
	%	%	total
male			
ISCED 1	0.48	1.49	703,000
ISCED 2	2.00	4.28	2,012,000
ISCED 3a	1.84	1.65	776,000
ISCED 3b	16.98	23.81	11,203,000
ISCED 3c	0.40	0.45	201,000
ISCED 4ab	2.96	3.40	1,602,000
ISCED 5a	12.54	8.16	3,841,000
ISCED 5b	11.18	6.05	2,845,000
ISCED 6	1.06	0.85	398,000
female			
ISCED 1	0.60	1.79	843,000
ISCED 2	3.47	6.43	3,024,000
ISCED 3a	1.57	1.36	639,000
ISCED 3b	18.38	23.65	11,131,000
ISCED 3c	0.19	0.31	144,000
ISCED 4ab	3.62	4.16	1,956,000
ISCED 5a	10.06	6.69	3,150,000
ISCED 5b	12.10	5.07	2,384,000
ISCED 6	0.57	0.42	199,000
Total	100.00	100.00	47,060,000

Table 25: Comparison of the distribution of the Wave 4 sample data and the target distribution (Microcensus 2011) according to birth year and educational attainment.

Birth year and education	actual distribution	target distribution	
	net sample	population (Microcensus 2011)	
	%	%	total
1975 – 1986			
ISCED 1	0.23	0.74	348,000
ISCED 2	1.02	2.57	1,207,000
ISCED 3ca	1.10	1.88	883,000
ISCED 3b	4.02	9.72	4,570,000
ISCED 4ab	0.03	0.03	13,000
ISCED 5b	1.55	3.00	1,412,000
ISCED 5a	3.67	4.25	1,999,000
ISCED 6	2.78	2.33	1,097,000
ISCED 6	0.23	0.21	97,000
1965 – 1974			
ISCED 1	0.33	0.90	425,000
ISCED 2	1.98	2.41	1,132,000
ISCED 3a	1.08	0.52	244,000
ISCED 3b	15.74	12.16	5,717,000
ISCED 3c	0.22	0.17	82,000

ISCED 4ab	2.94	2.25	1,060,000
ISCED 5a	9.83	4.96	1,860,000
ISCED 5b	10.47	3.00	1,411,000
ISCED 6	0.71	0.38	178,000
1956 – 1964			
ISCED 1	0.18	0.82	385,000
ISCED 2	1.08	2.46	1,155,000
ISCED 3a	0.54	0.37	174,000
ISCED 3b	8.47	12.53	5,895,000
ISCED 3c	0.11	0.24	114,000
ISCED 4ab	1.43	1.55	728,000
ISCED 5a	5.13	3.38	1,588,000
ISCED 5b	5.50	3.06	1,437,000
ISCED 6	0.32	0.34	160,000
1944 – 1955			
ISCED 1	0.26	0.83	389,000
ISCED 2	1.24	3.28	1,543,000
ISCED 3a	0.45	0.23	109,000
ISCED 3b	7.60	13.09	6,154,000
ISCED 3c	0.22	0.25	116,000
ISCED 4ab	0.55	0.77	360,1000
ISCED 5a	3.99	3.28	1,544,000
ISCED 5b	4.70	2.74	1,287,000
ISCED 6	0.32	0.32	156,000
Total	100.00	100.00	47,029,000

Table 26: Comparison of the distribution of the Wave 4 sample data and the target distribution (Microcensus 2011) according to Federal State.

Federal State	actual distribution	target distribution	
	net sample	population (Microcensus 2011)	
	%	%	total
Schleswig-Holstein	2.94	3.39	1,598,000
Hamburg	1.92	2.29	1,079,000
Niedersachsen	10.61	9.50	4,475,000
Bremen	0.63	0.83	392,000
Nordrhein-Westfalen	22.40	21.66	10,207,000
Hessen	7.64	7.50	3,533,000
Rheinland-Pfalz	4.87	4.82	2,272,000
Baden-Württemberg	12.24	12.93	6,094,000
Bayern	15.61	15.40	7,258,000
Saarland	1.42	1.24	582,000
Berlin	3.76	4.47	2,106,000
Brandenburg	3.23	3.16	1,491,000
Mecklenburg-Vorpommern	1.74	2.07	977,000
Sachsen	5.01	5.05	2,378,000
Sachsen-Anhalt	2.94	2.88	1,355,000
Thüringen	3.01	2.82	1,328,000
Total	100.00	100.00	47,125,000

Table 27: Comparison of the distribution of the Wave 4 sample data and the target distribution (Microcensus 2011) according to BIK categories of municipal size.

BIK categories	actual distribution	target distribution	
	net sample	population (Microcensus 2011)	
	%	%	total
less than 2000 inhab.	1.99	1.81	852,000
2000 to 5000 inhab.	2.55	2.75	1,298,000
5000 to 20,000 inhab.	8.03	8.10	3,819,000
20,000 to 50,000 inhab.	11.85	11.54	5,438,000
50,000 to 100,000 inhab. styp 2/3/4	9.05	7.84	3,695,000
50,000 to 100,000 inhab. styp 1	1.98	2.32	1,094,000
100,000 to 500,000 inhab. styp 2/3/4	16.40	14.41	6,795,000
100,000 to 500,000 inhab. styp 1	15.71	15.61	7,358,000
500,000 and more inhab. styp 2/3/4	8.85	9.37	4,418,000
500,000 and more inh. styp 1	23.58	26.25	12,374,000
Total	100.00	100.00	47,141,000

Table 28: Comparison of the distribution of the Wave 4 sample data and the target distribution (Microcensus 2011) according to birth year.

Year of birth	actual distribution	target distribution	
	net sample	population (Microcensus 2011)	
	%	%	total
1944	1.82	1.89	892,000
1945	1.38	1.43	674,000
1946	1.68	1.66	781,000
1947	1.85	1.86	878,000
1948	1.91	2.03	955,000
1949	2.39	2.18	1,025,000
1950	2.37	2.20	1,038,000
1951	2.42	2.23	1,053,000
1952	2.57	2.28	1,073,000
1953	2.20	2.31	1,088,000
1954	2.69	2.32	1,095,000
1955	2.35	2.38	1,122,000
1956	3.08	2.49	1,175,000
1957	2.96	2.54	1,195,000
1958	3.14	2.57	1,210,000
1959	3.79	2.68	1,261,000
1960	3.37	2.78	1,312,000
1961	3.47	2.85	1,344,000
1962	3.27	2.84	1,339,000
1963	3.52	2.95	1,392,000
1964	3.52	3.01	1,416,000
1965	3.58	2.97	1,399,000
1966	3.46	3.05	1,435,000
1967	2.95	2.96	1,397,000
1968	2.80	2.83	1,333,000
1969	2.39	2.68	1,264,000
1970	2.57	2.58	1,215,000

1971	1.97	2.47	1,165,000
1972	1.98	2.19	1,032,000
1973	1.70	2.01	946,000
1974	1.47	1.98	933,000
1975	1.56	1.93	910,000
1976	1.40	2.01	948,000
1977	1.54	1.99	939,000
1978	1.62	2.04	963,000
1979	1.55	2.04	960,000
1980	1.59	2.19	1,032,000
1981	1.47	2.14	1,007,000
1982	1.59	2.15	1,014,000
1983	1.70	2.11	994,000
1984	1.60	2.04	962,000
1985	1.75	2.04	961,000
1986	2.01	2.10	991,000
Total	100.00	100.00	47,118,000

Table 29: Comparison of the distribution of the Wave 4 sample data and the target distribution (Microcensus 2011) according to country of birth.

Country of birth	actual distribution	target distribution	
	net sample	population (Microcensus 2011)	
	%	%	total
born abroad	9.63	17.69	8,335,000
born in Germany	90.37	82.31	38,783,000
Total	100.00	100.00	47,118,000

Table 30: Results of the logit regression model measuring the participation willingness of repeaters in Wave 5.

Variable	Reference	Odds Ratio	P-Value
<i>Birth year</i>	1980 – 1986		
1970 – 1979		1.08	0.52
1956 – 1969		0.99	0.92
1944 – 1955		0.60	0.00
<i>Gender</i>	female		
male		1.11	0.13
<i>Country of birth</i>	born in Germany		
born abroad		0.76	0.06
<i>Mother tongue</i>	Non-German		
German		1.21	0.22
<i>Marital status</i>	unmarried		
married		2.15	0.00
separated		1.99	0.00
widowed		2.63	0.00

<i>Household size</i>	two persons		
one person		1.34	0.01
three persons		0.87	0.11
four persons		0.87	0.17
five or more persons		0.94	0.66
<i>School qualification</i>	ISCED 3b		
ISCED 1/2		0.91	0.38
ISCED 3ca/4ab		1.29	0.01
ISCED 5b		0.87	0.25
ISCED 5a/6		1.25	0.01
<i>School qualification parents</i>	'Realschule'		
'Hauptschule'		1.22	0.01
upper secondary education		1.46	0.00
other		2.38	0.02
<i>Income</i>	1,501 – 3,500 Euro		
up to 1,500 Euro		1.01	0.89
more than 3,500 Euro		1.29	0.00
<i>Federal state</i>	Nordrhein-Westfalen		
Schleswig-Holstein		1.15	0.48
Hamburg		0.90	0.63
Niedersachsen		1.37	0.01
Bremen		1.12	0.78
Hessen		1.32	0.05
Rheinland-Pfalz		1.26	0.15
Baden-Württemberg		1.03	0.81
Bayern		1.00	0.97
Saarland		0.92	0.74
Berlin		1.55	0.03
Brandenburg		0.93	0.68
Mecklenburg-Vorpommern		1.50	0.14
Sachsen		1.22	0.20
Sachsen-Anhalt		2.08	0.00
Thüringen		1.45	0.08
Pseudo R^2	0.0271		
Number of cases	16,356		

Note: At the end of the studies of Wave 4 the panel sample of Starting Cohort 6 comprised 16,356 cases who were willing to further participate in NEPS. In the beginning of the studies of Wave 5, this number reduced to 15,249.

Table 31: Results of the logit regression model measuring the participation propensity of repeaters in Wave 5.

Variable	Reference	Odds Ratio	P-Value
<i>Birth year</i>	1980 – 1986		
1970 – 1979		1.14	0.13
1956 – 1969		1.35	0.00
1944 – 1955		1.24	0.02
<i>Gender</i>	female		
male		0.98	0.72
<i>Country of birth</i>	born in Germany		
born abroad		0.99	0.94
<i>Mother tongue</i>	Non-German		
German		1.82	0.00
<i>Marital status</i>	unmarried		
married		1.40	0.00
separated		1.24	0.03
widowed		1.39	0.06
<i>Household size</i>	two persons		
one person		1.15	0.09
three persons		1.03	0.68
four persons		1.08	0.28
five persons and more		0.99	0.94
<i>School qualification</i>	ISCED3B		
ISCED1/2		0.84	0.04
ISCED3CA/4AB		1.35	0.00
ISCED5B		1.19	0.06
ISCED5A/b		1.45	0.00
<i>School qualification of parents</i>	'Realschule'		
'Hauptschule'		1.03	0.64
upper secondary education		0.96	0.53
other		0.67	0.03
<i>Income</i>	1.501 – 3500 Euro		
up to 1500 Euro		1.00	0.97
more than 3500 Euro		0.95	0.35
<i>Federal State</i>	Nordrhein-Westfalen		
Schleswig-Holstein		0.96	0.79
Hamburg		1.12	0.52
Niedersachsen		1.27	0.01
Bremen		1.49	0.22
Hessen		1.35	0.00
Rheinland-Pfalz		0.99	0.96
Baden-Württemberg		1.20	0.03
Bayern		1.30	0.00
Saarland		0.93	0.73
Berlin		1.17	0.24
Brandenburg		1.09	0.53
Mecklenburg-Vorpommern		0.69	0.03
Sachsen		1.40	0.01
Sachsen-Anhalt		1.30	0.09
Thüringen		1.62	0.00
<i>BIK categories</i>	500,000 and more inh. (styp 1)		
less than 2000 inhab.		1.61	0.02

2000 to 5000 inhab.		0.77	0.09
5000 to 20,000 inhab.		0.91	0.35
20,000 to 50,000 inhab.		0.93	0.44
50,000 to 100,000 inhab. (styp 2/3/4)		0.89	0.23
50,000 to 100,000 inhab. (styp 1)		1.19	0.35
100,000 to 500,000 inhab. (styp 2/3/4)		1.05	0.57
100,000 to 500,000 inhab. (styp 1)		1.14	0.12
more than 500,000 inhab. (styp 2/3/4)		0.97	0.78
<i>Attempts to contact target</i>	1 to 3 attempts		
4 to 6 attempts		0.96	0.48
7 to 10 attempts		0.61	0.00
more than 10 attempts		0.18	0.00
Pseudo R2	0.10		
Number of cases	13,860		

Table 32: Results of the logit regression model measuring the participation propensity of individuals who participated in Wave 5 but not in Wave 4.

Variable	Reference	Odds Ratio	P-Value
<i>Birth year</i>	1980 – 1986		
1970 – 1979		1.08	0.69
1956 – 1969		1.06	0.72
1944 – 1955		0.58	0.01
<i>Gender</i>	female		
male		1.27	0.04
<i>Country of birth</i>	born in Germany		
born abroad		0.55	0.01
<i>Household size</i>	two persons		
one person		1.25	0.20
three persons		1.53	0.01
four persons		0.76	0.14
five persons and more		0.85	0.54
<i>Federal State</i>	Nordrhein-Westfalen		
Schleswig-Holstein		1.10	0.81
Hamburg		1.26	0.05
Niedersachsen		1.96	0.00
Bremen		2.91	0.19
Hessen		2.20	0.00
Rheinland-Pfalz		1.04	0.89
Baden-Württemberg		1.36	0.15
Bayern		1.35	0.15
Saarland		1.06	0.91
Berlin		1.65	0.15
Brandenburg		1.43	0.37
Mecklenburg-Vorpommern		1.42	0.51
Sachsen		3.12	0.00
Sachsen-Anhalt		0.68	0.35
Thüringen		1.26	0.59
<i>BIK categories</i>	500,000 and more inh. Styp 1		

less than 2000 inhab.		2.94	0.02
2000 to 5000 inhab.		1.31	0.55
50000 to 20,000 inhab.		0.75	0.28
20,000 to 50,000 inhab.		1.09	0.68
50,000 to 100,000 inhab. (styp 2/3/4)		1.09	0.74
50,000 to 100,000 inhab. (styp 1)		1.33	0.54
100,000 to 500,000 inhab. (styp 2/3/4)		1.06	0.80
100,000 to 500,000 inhab. (styp 1)		1.17	0.44
more than 500,000 inhab. (styp 2/3/4)		0.89	0.65
<i>Attempts to contact target</i>	1 to 3 attempts		
4 to 6 attempts		1.25	0.12
7 to 10 attempts		1.18	0.41
more than 10 attempts		0.35	0.00
Pseudo R2	0.09		
Number of cases	1389		

Table 33: Comparison of the distribution of the Wave 5 sample data and the target distribution (Microcensus 2012) according to gender and educational attainment.

Gender and education	actual distribution	target distribution	
	net sample	population (Microcensus 2012)	
	%	%	total
male			
ISCED 1	0.32	1.48	696,000
ISCED 2	0.31	4.26	2,006,000
ISCED 3a	0.38	1.44	680,000
ISCED 3b	9.44	23.80	11,208,000
ISCED 3c	0.25	0.36	170,000
ISCED 4ab	0.97	3.35	1,579,000
ISCED 5a	21.63	8.49	4,000,000
ISCED 5b	14.92	6.02	2,837,000
ISCED 6	1.07	0.87	408,000
female			
ISCED 1	0.84	1.78	893,000
ISCED 2	0.67	6.47	3,047,000
ISCED 3a	1.63	1.18	558,000
ISCED 3b	8.01	23.56	11,093,000
ISCED 3c	0.07	0.27	127,000
ISCED 4ab	0.45	4.09	1,924,000
ISCED 5a	21.46	7.07	3,329,000
ISCED 5b	16.79	5.05	2,378,000
ISCED 6	0.61	0.45	214,000
Total	100.00	100.00	47,093,000

Table 34: Comparison of the distribution of the Wave 5 sample data and the target distribution (Microcensus 2012) according to birth year and educational attainment.

Birth year and education	actual distribution	target distribution	
	net sample	population (Microcensus 2012)	
	%	%	total
1975 – 1986			
ISCED 1	0.26	0.73	338,000
ISCED 2	0.09	2.34	1,076,000
ISCED 3a	0.06	1.19	548,000
ISCED 3b	0.76	9.13	4,208,000
ISCED 3c	0.01	0.01	600,346
ISCED 4ab	0.03	2.72	1,255,000
ISCED 5a	6.00	4.62	2,126,000
ISCED 5b	6.47	2.31	1,062,000
ISCED 6	0.26	0.26	121,000
1965 – 1974			
ISCED 1	0.28	0.89	408,000
ISCED 2	0.46	2.47	1,137,000
ISCED 3a	0.96	0.54	247,000
ISCED 3b	6.71	12.35	5,690,000
ISCED 3c	0.05	0.15	68,000
ISCED 4ab	0.54	2.26	1,040,000
ISCED 5a	20.29	4.14	1,909,000
ISCED 5b	13.67	2.99	1,379,000
ISCED 6	0.71	0.40	183,000
1956 – 1964			
ISCED 1	0.13	0.80	369,000
ISCED 2	0.24	2.55	1,174,000
ISCED 3a	0.48	0.38	177,000
ISCED 3b	3.81	12.89	5,940,000
ISCED 3c	0.01	0.20	90,000
ISCED 4ab	0.30	1.57	721,000
ISCED 5b	10.91	3.50	1,610,000
ISCED 5a	6.96	3.11	1,433,000
ISCED 6	0.31	0.35	159,000
1944 – 1955			
ISCED 1	0.54	0.86	397,000
ISCED 2	0.28	3.36	1,548,000
ISCED 3a	0.53	0.25	6,102,000
ISCED 3b	6.11	13.25	104,000
ISCED 3c	0.18	0.23	1,190,735
ISCED 4ab	0.53	0.75	346,000
ISCED 5b	6.72	3.40	1,564,000
ISCED 5a	4.15	2.74	1,261,000
ISCED 6	0.32	0.33	152,000
Total	100.00	100.00	46,065,000

Table 35: Comparison of the distribution of the Wave 5 sample data and the target distribution (Microcensus 2012) according to Federal State.

Federal State	actual distribution	target distribution	
	net sample	population (Microcensus 2012)	
	%	%	total
Schleswig-Holstein	2.80	3.41	1,606,000
Hamburg	1.89	2.29	1,082,000
Niedersachsen	10.92	9.53	4,492,000
Bremen	0.68	0.81	384,000
Nordrhein-Westfalen	21.39	21.64	10,205,000
Hessen	7.85	7.52	3,546,000
Rheinland-Pfalz	4.69	4.83	2,279,000
Baden-Württemberg	12.27	12.89	6,080,000
Bayern	15.77	15.44	7,281,000
Saarland	1.39	1.23	578,000
Berlin	3.83	4.51	2,125,000
Brandenburg	3.27	3.17	1,497,000
Mecklenburg-Vorpommern	1.65	2.04	964,000
Sachsen	5.42	5.02	2,365,000
Sachsen-Anhalt	3.06	2.86	1,350,000
Thüringen	3.10	2.80	1,319,000
Total	100.00	100.00	47,153,000

Table 36: Comparison of the distribution of the Wave 5 sample data and the target distribution (Microcensus 2012) according to BIK categories of municipal size.

BIK categories	actual distribution	target distribution	
	net sample	population (Microcensus 2012)	
	%	%	total
less than 2000 inhab.	2.18	1.65	780,000
2000 to 5000 inhab.	2.46	2.59	1,220,000
5000 to 20,000 inhab.	7.92	8.45	3,985,000
20,000 to 50,000 inhab.	11.91	10.47	4,935,000
50,000 to 100,000 inhab. styp 2/3/4	8.96	8.20	3,866,000
50,000 to 100,000 inhab. styp 1	2.00	2.28	1,074,000
100,000 to 500,000 inhab. styp 2/3/4	16.58	14.65	6,908,000
100,000 to 500,000 inhab. styp 1	15.77	15.12	7,127,000
500,000 and more inhab. styp 2/3/4	8.88	9.59	4,523,000
500,000 and more inh. styp 1	23.34	27.00	12,731,000
Total	100.00	100.00	47,149,000

Table 37: Comparison of the distribution of the Wave 5 sample data and the target distribution (Microcensus 2012) according to birth year.

Year of birth	actual distribution	target distribution	
	net sample	population (Microcensus 2012)	
	%	%	total
1944	1.75	1.86	875,000
1945	1.45	1.44	678,000
1946	1.68	1.64	772,000
1947	1.91	1.82	859,000
1948	1.90	2.02	951,000
1949	2.34	2.19	1,032,000
1950	2.38	2.20	1,037,000
1951	2.46	2.22	1,047,000
1952	2.61	2.27	1,070,000
1953	2.26	2.27	1,072,000
1954	2.67	2.28	1,077,000
1955	2.40	2.41	1,135,000
1956	3.23	2.50	1,179,000
1957	2.94	2.53	1,194,000
1958	3.29	2.61	1,230,000
1959	3.83	2.67	1,261,000
1960	3.51	2.80	1,320,000
1961	3.40	2.85	1,343,000
1962	3.47	2.85	1,343,000
1963	3.57	2.97	1,399,000
1964	3.60	3.01	1,421,000
1965	3.69	2.94	1,388,000
1966	3.49	2.95	1,391,000
1967	2.89	2.88	1,358,000
1968	2.99	2.89	1,364,000
1969	2.45	2.71	1,277,000
1970	2.49	2.59	1,222,000
1971	1.89	2.44	1,152,000
1972	1.92	2.20	1,038,000
1973	1.73	1.99	940,000
1974	1.41	2.00	944,000
1975	1.44	1.94	917,000
1976	1.42	2.03	955,000
1977	1.68	1.99	939,000
1978	1.51	2.09	984,000
1979	1.52	2.07	975,000
1980	1.39	2.16	1,019,000
1981	1.37	2.14	1,009,000
1982	1.48	2.19	1,032,000
1983	1.55	2.09	987,000
1984	1.44	2.08	983,000
1985	1.57	2.09	985,000
1986	2.06	2.11	996,000
Total	100.00	100.00	4,715,000

Table 38: Comparison of the distribution of the Wave 5 sample data and the target distribution (Microcensus 2012) according to country of birth.

Country of birth	actual distribution	target distribution	
	net sample	population (Microcensus 2012)	
	%	%	total
born abroad	8.44	17.96	8,466,000
born in Germany	91.56	82.04	38,684,000
Total	100.00	100.00	47,150,000