

Weighting Procedure of the Austrian Microcensus using Administrative Data

Angelika Meraner
Statistics Austria

Daniela Gumprecht
Statistics Austria

Alexander Kowarik
Statistics Austria

Abstract

The Austrian microcensus is the biggest sample survey of the Austrian population, it is a regionally stratified cluster sample with a rotational pattern. The sampling fractions differ significantly between the regions, therefore the sample size of the regions is quite homogeneous. The primary sampling unit is the household, within each household all persons are surveyed. The design weights are the input for the calibration on population counts and household forecasts. It is performed by iterative proportional fitting. Until the third quarter of 2014 only demographic, regional and household information were used in the weighting procedure. From the fourth quarter 2014 onwards the weighting process was improved by adding an additional dimension to the calibration, namely a labour status generated from administrative data and available for the whole population. Apart from that some further minor changes were introduced. This paper describes the methodological and practical issues of the microcensus weighting process and the variance estimation applied from 2015 onwards. The new procedure was used for the first time for the fourth quarter of 2014, published at the end of March 2015. At the same time, all previous microcensus surveys back to 2004 were reweighted according to the new approach.

Keywords: microcensus, weighting, iterative proportional fitting, variance, bootstrap.

1. Introduction

The Austrian microcensus (MC) is regulated by law¹ and carried out by Statistics Austria (STAT) since the 1970s. Over the years, the survey was modified on several occasions to better estimate reality. A principal reformation took place in 2004 when the MC was completely reorganised. Since that time the Austrian Labour Force Survey (LFS) is part of the Austrian microcensus, see [Kytir and Stadler 2004](#). The LFS is an important data source for main economic and social indicators focusing on the labour market. It is based on definitions on employment and unemployment stated by the International Labour Organisation (ILO). Results are comparable with other countries.

For ten years the practices remained unchanged, until the utilisation of anonymised individual-related administrative data for official statistics was well developed and represented a great opportunity for quality improvement of the MC. The availability of administrative data and

¹Erwerbs- und Wohnungsstatistikverordnung, BGBl. II Nr. 111/2010.

the possibility of micro data linkage, i.e. to connect them with other administrative data as well as survey data anonymously, so without any knowledge of names or social insurance numbers etc., and the opportunity for use in official statistics lead to a wide range of innovative methods and proceedings.

The preparation and implementation of the latest census 2011 as register based census also lead to a number of new possibilities and applications for survey data like the microcensus. In this context, the new weighting of the MC was developed. Linking MC survey data with administrative data on the labour status showed a small but non-negligible bias in the MC data concerning the labour status. Employed persons were slightly over-covered and by contrast unemployed and persons out of labour force were under-covered. The calibration (used until reference period q3_2014, see [Haslinger and Kytir 2006](#)) reduced this bias, but not sufficiently, i.e. employed persons were still overestimated and non-employed were underestimated. As a consequence, administrative data on the labour status were included in the weighting process to correct this bias.

Beside the changes of the calibration procedure there was another reason for a revision. With the results of the latest census, which took place in 2011, new population numbers and household numbers were available to which the weighting procedure had to be adapted retroactively. To avoid breaks in the time series simply based on methodological changes and to allow for time series analysis, the whole MC back to the first quarter of 2004 was reweighted with the new procedure and the results were revised. The reweighting lead to a decrease of the estimated number of persons employed and to an increase of the estimated number of persons unemployed and out of labour force, e.g. the estimated yearly unemployment rate of 2013 (persons aged between 15 to 74) increased from 4.9% to 5.4% due to the change of the weighting, whereas the estimated employment rate of persons aged between 15 and 64 decreased from 72.3% to 71.4%. These changes are along the same lines for men and women, although the differences are usually bigger for men. Regarding the whole time series from 2004 onwards the reweighting lead to level shifts but to no change of the trends and seasonal patterns.

The latest renewals and changes of the Austrian microcensus weighting procedure are also described in a working paper (German-language only) published at the website of Statistics Austria, see [Meraner, Gumprecht, and Kowarik 2015](#).

In the following, this paper presents the sample of the Austrian microcensus, the weighting and the error estimation. Section 2 describes design and sampling frame of the Austrian microcensus as well as the problems with non-response which is on a low level but nevertheless lead to biased results which is shown in a non-response analysis using administrative data (see 2.3.1). Section 3 describes the new weighting procedure including among other things also calibration specifications and the steps of the iterative proportional fitting process. Section 4 deals with the calculation of the standard errors and confidence intervals via a bootstrap procedure as well as an approximation of the errors. As usual, the last section 5 gives a summary of the paper as well as an outlook to further topics of research.

2. Sample design and data collection

The MC is a stratified random sample of private households or rather addresses by NUTS-2 region, therefore results on NUTS-2 level are of (controlled) high quality. Due to organisational reasons the sample selection itself is done at the level of the political districts and within districts with a small population density the sampling fraction is doubled. All other districts have the same sampling fraction as its NUTS-2 region. The MC sample contains approximately 20,000 households per quarter, these are about 44,000 persons or 0.5% of the whole Austrian population.

2.1. Sampling frame

The sample frame is the Austrian central population register. All non-private households, households which were part of the MC within the last ten years (a household can participate only once in ten years), and households being no main residence of at least one person are excluded. About four to six months before the first interview takes place, the households are selected from a current copy of the central population register. Due to this time lag between sampling and interview it may happen that situations change in the meanwhile, e.g., persons move out and others move in, some die and others are born etc. Regardless of the situation in the central population register at the time of the sampling, also called “register reality”, the persons who are actually living at the selected household, i.e. at this address, at the time of the interview are subjects of the MC. If e.g., the address does not exist anymore it drops out of the MC sample.

2.2. Sample design

Once a household is selected, it stays in the sample for five consecutive quarters and is questioned in each of these five quarter, starting as wave 1 in the very first quarter, becoming wave 2 in the next quarter and so on. The MC is a rotating sample, each quarter one fifth of the total sample rotates in and out. Each quarter all persons who are currently living in the household are interviewed. The first interview is a computer assisted personal interview (CAPI), the following four interviews are usually computer assisted telephone interviews (CATI). Each household is assigned to a reference week and most questions refer to that week. Moreover, the households are evenly distributed across all the reference weeks of a quarter. The interviews should be done shortly after the reference week and not later than five (in summer six) weeks after the reference week. There is a legal obligation to participate in the microcensus, but if a person is not able or willing to do the interview by herself, any other adult person living in the same household is allowed to answer by proxy.

This and more information about the MC sample design and sampling frame can be found in [Haslinger and Kytir 2006](#).

2.3. Non-response

In general there is a very low level of non-response, mainly due to the legal obligation to participate, nevertheless there is a certain degree of non-response due to different reasons, like denial, non-reachability, language problems, wrong addresses, etc. The degree of non-response is usually given by the non-response rate or - its counterpart - the response rate, which both can be calculated in different ways. In Austria it is the share of non-neutral non-response in the gross-sample, which does not include neutral non-response cases (the whole sample is called gross-gross-sample; the sample without neutral non-response and non-neutral non-response is called net-sample). Thus, the non-response share depends on the classification of cases as neutral and non-neutral non-response. Neutral non-response cases are e.g., if a household is absent during the quarter, if household-members are incapable of being interviewed, if it is no private household, and the like; whereas non-neutral non-response is e.g., refusal. As the MC is a sample of households (not persons) non-response statistics also refer to households. In the first quarter 2013 the gross-gross-sample was 22,499 households, 1,188 or 5.3% of them were classified as neutral non-response, this lead to a gross-sample of 21,311 households and 19,931 or 93.5% of these were response households. The non-response rate is calculated by the non-response cases of the sample divided by the whole sample cases. The size depends on the definition of non-response cases as well as whole sample cases. Yet, even if the share of non-response is quite small, it might lead to problems if non-response is not at random, i.e. if there are structural differences between response and non-response cases, a bias can be introduced.

2.3.1. Non-response analysis using administrative data

In labour market statistics, the persons are the main subjects of interest, therefore the non-response analysis has to be done on the level of persons instead of households. Structural differences in the labour status between non-response and response persons can lead to a biased estimation of certain labour market groups. On the level of persons a non-response analysis is more complex than for households, for example, there might be persons living in “response” households who are non-respondents after all, because their existence is not disclosed by the interviewed household member(s). If a whole household does not answer, it is not easily possible to know the number of persons actually living there, therefore not even the number of non-response persons is easy to determine. Beside the pure number of non-response persons also their labour status is of interest, so that the assumption of a bias concerning the labour status in the MC net sample can be checked. A special non-response analysis was done for reference period 2012 using data from the central population register, the central social security register and the labour market service (see [Gumprecht and Oismüller 2013](#)), this information from administrative data was used as a proxy for the missing MC information. Correlation of ILO- and administrative labour status is very high for persons employed, more than 90% of persons employed according to ILO definition are also employed according to administrative data. However, correlation between ILO- and administrative definitions of unemployment is considerably lower with about 70% of persons unemployed according to ILO definition also being unemployed according to administrative data and about 42% of registered unemployed persons also being unemployed according to ILO definitions. Nevertheless, correlation between the labour status according to these two definitions is quite high and administrative data are good proxies for missing values in MC data. Non-response analysis showed that persons employed according to administrative information tend to be overrepresented in the MC sample whereas unemployed and persons out of labour force are rather underrepresented, and the calibration procedure used at that time (until the third quarter 2014) could not entirely correct this bias. The findings of this non-response analysis were the motivation to change the weighting of the Austrian microcensus. Administrative employment status is a good candidate for an additional weighting specification, because it is known in the sample as well as in the whole population and it is highly correlated with the main variable of interest, the ILO labour status.

3. Weighting

Drawing a random sample means reducing the target population to a subset that should represent this population in an unbiased way, therefore, computing population statistics requires a reversal of the reduction which is done by projection. The data collected from the sample is used to estimate the unknown population parameters of interest (totals, means, medians, ratios, ...) with the respective weighted estimator. The inverse of the selection probabilities is the so-called design weight, it is the most basic weighting scheme. The Austrian microcensus has different selection probabilities for each stratum (NUTS-2 region), so each region has to be projected separately. In addition to that, known population totals, such as age and gender, retrieved from sources like the statistical population register, are used to calibrate the weights. The calibration is performed with the iterative proportional fitting procedure, which basically computes multipliers known as calibration factors that adjust the sampling weights to make the population estimates agree with the known totals. Theoretically, every characteristic known for the units in the sample and for the whole population can be used as a calibration variable, but only variables correlated with target variables enhance precision. Practically of course there are some restrictions, e.g., the sample size.

The following description of the calibration specifications, computation of base weights and the steps within the iterative proportional fitting procedure is strongly based on [Haslinger and Kytir 2006](#) with the main difference being the usage of an additional calibration specification,

the administrative employment status.

3.1. Calibration specifications

The calibration specifications for the iterative proportional fitting procedure are the following:

- N_{rga} ... Total number of persons in private households in NUTS-2 region r ($=1, \dots, 9$), of gender g ($=1, 2$) and in age class a (1=0-2 years, 2=3-5 years, 3=6-9, ... 5-year classes ... , 18=80-84, 19=85+).
- N_{rn} ... Total number of persons in private households in NUTS-2 region r ($=1, \dots, 9$) with nationality n ($=1, \dots, 6$) comprising the groups “Austria”, “EU-15 (w.o. Austria)”, “EU 2004+ (joined the EU between 2004 and 2014)”, “European non-EU countries”, “Turkey”, “Other”.
- N_{rge} ... Total number of persons in private households in NUTS-2 region r ($=1, \dots, 9$), of gender g ($=1, 2$) and with administrative employment status e ($=1, \dots, 5$) which consists of the groups “Employee standard”, “Employee non-standard”, “Self-Employed”, “Unemployed” and “Inactive”.
- M_{rh} ... Total number of private households in NUTS-2 region r ($=1, \dots, 9$) of household size h ($=1, \dots, 5$) with values 1 to 5+.

Information about variables gender, age, and nationality come from the statistical population register “POPREG”, it includes all persons living in Austria at the beginning of a quarter. Information about the labour status from administrative data is available from the Austrian central social security register and the labour market service, see Section 3.1.3. The number of private households with h members stem from the household projection of STAT, see Section 3.1.2 which is based on the register based census for the census years (2011, 2021, etc.) and the register-based labour market statistics for the years in between.

3.1.1. Private and institutional households

To compute the number of persons in private households it is essential to segregate institutional households like prisons, care homes, residential schools etc. that are also included in the POPREG. Persons in institutional households are removed via so-called institutional rates (share of persons in institutional households) for all relevant combinations of the characteristics used as weighting specifications. These rates stem from a special survey done by STAT, called “institutional survey”, which is available for each year for reference date October 31st (from 2011 onwards), but with a time lag of up to two years. Institutional rates are held constant until new rates are available.

3.1.2. Household projection of STAT

In addition to the weighting specifications concerning the number of persons living in private households, the number of private households (by size and NUTS-2 region) is used for calibration. Since the real numbers are not known for all quarters, results from the household projection of STAT are used instead. For the fourth quarter of each year, real data is available, either from the census or the register-based labour market statistics, however with a time lag of about two years. For the quarters in between and thereafter, a two-step projection following Ediev 2007 is used.

3.1.3. Labour status from administrative data

Similar to the use in the non-response analysis, data from the Austrian central social security register and the labour market service as well as the statistical population register POPREG

from STAT is used to generate an administrative labour status without any survey information. It can be calculated on individual level for approximately 94% of all MC respondents, as well as for the whole population, i.e., all persons living in private households in Austria at the beginning of the MC reference quarter, for reference period end of a month. For about 6% of the MC-persons, administrative data cannot be linked to survey data because no area specific personal identifier (bPK) is available. For them, the administrative employment status is imputed using random hot-deck within a domain, which is the default procedure for most of the MC variables. In the imputation process ILO labour status, region and gender are used as explanatory variables, i.e. the correlation between administrative employment status and ILO labour status (see Section 2.3.1) can be used again.

The variable “administrative employment status” has the following five values:

- Employee standard: Persons with dependent employment concerning to the social security register, e.g., white- and blue-collar workers, civil servants.
- Employee non-standard: Persons with non-standard dependent employment concerning to the social security register, e.g., holder of non-standard contract, marginal part-timers, persons in parental leave, etc.
- Self-Employed: Self-employed persons concerning to the social security register, e.g., freelancers, self-employed and family workers in agriculture.
- Unemployed: Persons unemployed concerning to labour market service, e.g., job seeking persons, persons in training measures.
- Inactive: Persons living in Austria and being neither employed nor unemployed.

Data from the social security register and the labour market service can contain several cases per person, therefore, the dominant case has to be selected for every person, e.g., employment always is prevailing compared to unemployment and inactivity. To guarantee a complete coverage of the reference quarter, this is performed for the three end-of-month administrative data deliveries pertaining to the reference quarter as well as the end-of-month delivery of the month preceding the reference quarter, but only for persons corresponding to the population in private households at the beginning of the respective reference quarter. A weighted mean of these four final monthly results m_{t-1}, m_t, m_{t+1} and m_{t+2} with the first month t of a quarter gives the quarterly results N_{rge} as defined above which are used as weighting specifications:

$$N_{rge} = \frac{1}{3} \left(\frac{m_{t-1} + m_t}{2} + \frac{m_t + m_{t+1}}{2} + \frac{m_{t+1} + m_{t+2}}{2} \right). \quad (1)$$

3.2. Base weights

The final sampling weights are computed in an iterative process starting with the base weights which are then calibrated against the population totals defined in Section 3.1.

The base weights are basically the inverse of the selection probabilities and are determined for every person and household as

$$\frac{M_r}{m_r} \quad (2)$$

where M_r is the total number of inhabited addresses in NUTS-2 region r and m_r is the number of addresses in the net sample of NUTS-2 region r . As mentioned in Section 2, the sample selection itself is done at the level of the political districts and the sampling fraction is doubled within districts with small population density. For NUTS-2 regions containing districts like these, the base weights for persons and households corresponding to districts with “normal” sampling fraction are computed as

$$\frac{M_r}{m_{r1} + m_{r2}/2} \quad (3)$$

where m_{r1} is the number of sampled addresses belonging to districts with “normal” sampling fraction while m_{r2} denotes the number of those with double sampling fraction. The weights for persons and households with double sampling fraction are calculated as half the quotient (3).

3.3. Iterative proportional fitting

Projection with the base weights computed as the quotients (2)-(3) yields population statistics that differ from the known population figures according to age, gender, nationality and administrative labour status. Many microcensus variables used for analyses are highly correlated with these characteristics. As mentioned above, we can use the known population totals to adjust the base weights accordingly. The iterative proportional fitting procedure applied for this purpose consists of 5 iteration steps and is explained in the following:

For every person i in the sample the base weights which shall be referred to as w_i^0 and the values of the variables region r ($=1, \dots, 9$), gender g ($=1, 2$), age group a ($=1, \dots, 19$), nationality n ($=1, \dots, 6$) and administrative employment status e ($=1, \dots, 5$) are needed. The running index k is set to $k = 0$ at the outset. If the constraint after step 5 is satisfied, k is raised by 1 and the procedure starts again with step 1.

3.3.1. Steps 1-3

The base weight w_i^0 of each person is modified by multiplication with a factor so that the projected distribution of the population matches the respective calibration specification $N_v \in \{N_{rga}, N_{rn}, N_{rge}\}$ (see Section 3.1) in every iteration step. That is, the calibration specification N_v with $v \in \{rga, rn, rge\}$ varies depending on the iteration step t :

- if $t = 1$, $N_v = N_{rga}$,
- if $t = 2$, $N_v = N_{rn}$,
- if $t = 3$, $N_v = N_{rge}$.

In every step t , the weight calibrated against N_v is computed as

$$w_i^{5k+t} = w_i^{5k+t-1} \frac{N_{v_i}}{\sum_l w_l^{5k}} \quad \forall i. \quad (4)$$

For $v_i = r_i g_i a_i$ the summation in the equation expands over all observations l with the same values of the characteristics r, g and a as observation i . This applies to $v_i = r_i n_i$ and $v_i = r_i g_i e_i$ in analogue form.

Weights w_i^{5k+t} outside of $\frac{w_i^0}{4} \leq w_i^{5k+t} \leq 4w_i^0$ are recoded to the nearest of these two boundaries. The constraints are based on expert opinion and should in general restrict the variance which has a positive effect on the sampling error. Restricting the change of the base weights is also general practice in other countries, i.e. it is a common form of weight trimming where very large/small weights are trimmed back to an upper/lower boundary (see e.g., Potter 1990, 1993) in order to reduce the variance but with the possible side effect of introducing a bias into the estimates.

3.3.2. Step 4

Due to modifications of the sampling weights in iteration steps 1-3, the weights of the persons in a household are no longer identical unless the household members all show the same values for the characteristics age group, gender, nationality and administrative labour status. Such heterogeneous weights lead to inconsistencies between results projected with household and person weights if e.g., the weight of the household reference person is used as the household weight. To avoid this, every person in a household is assigned the mean of the person weights

corresponding to the household. For every person i who is a member of household j with N_j household members l , it holds that:

$$w_i^{5k+4} = \frac{\sum_{l=1}^{N_j} w_l^{5k+3}}{N_j} \quad \forall i. \quad (5)$$

As a result, the adaptation to the population structure performed in steps 1-3 may be lost again.

3.3.3. Step 5

The weights from step 4 are modified to make the projected distribution of household size by region approximate the calibration specification M_{rh} . We do not aim at an exact match since the specifications themselves are subject to uncertainty. The uncertainty p_h for the number of households of size h per region is assumed to be 0.005 for $h = 1, \dots, 4$ and 0.2 for $h = 5$. The adjusted weights are computed as

$$w_i^{5k+5} = \begin{cases} w_i^{5k+4} \frac{M_{r_i h_i}}{\sum_l w_l^{5k+4}} & \text{if } \sum_l w_l^{5k+4} \notin ((1 - p_h)M_{r_i h_i}, (1 + p_h)M_{r_i h_i}) \\ w_i^{5k+4} & \text{otherwise} \end{cases} \quad (6)$$

where the summation expands over all households l with the same values of the characteristics r and h as observation i . Again, the new weights w_i^{5k+5} should not exceed 4 times the base weights w_i^0 nor should they fall below one quarter of w_i^0 , i.e. if $\frac{w_i^{5k+5}}{w_i^0} > 4$, we set $w_i^{5k+5} = 4w_i^0$ and if $\frac{w_i^{5k+5}}{w_i^0} < \frac{1}{4}$ we set $w_i^{5k+5} = \frac{w_i^0}{4}$.

3.3.4. Check

Finally, we check whether the deviation of the projected results from the known totals is greater than 0.01% in any cell:

$$\max_v \left| \frac{\sum_l w_l^{5k+5} - N_v}{N_v} \right| > 0.0001 \quad \text{for } v \in \{rga, rn, rge\}. \quad (7)$$

For $v = rga$ the summation in the equation expands over all observations l sharing the same values of the characteristics r, g and a . This applies to $v = rn$ and $v = rge$ in analogue form. If the maximum deviation is exceeding this limit, k is raised by 1 and the procedure continues at step 1 with the weights computed in step 5 as initial weights w_i^{5k} . It should be noted, that the constraints restricting the variance of the weights in steps 1-3 and 5 always refer to the "original" base weights w_i^0 .

Convergence is reached if the maximum deviation falls below the threshold, in that case, w_i^{5k+5} are the final calibrated weights w_i . Usually, the method converges after approximately 130 iteration steps.

3.4. Monthly weights and yearly results

Monthly weights are computed almost the same way as quarterly weights with the difference of using the total number N_n of persons in private households with nationality n instead of the total number N_{rn} of persons in private households with nationality n in NUTS-2 region r as a calibration specification (see Section 3.1). Also, the specification N_{rge} is not computed as the mean of four months (see Section 3.1.3) but as the mean of two months, i.e. the end-of-month values of the reference month as well as the end-of-month values of the previous month.

For yearly data, the quarterly data sets corresponding to a year are aggregated and the quarterly weights are divided by four.

4. Error estimation

4.1. Bootstrapping

4.1.1. The naïve approach

Standard errors and confidence intervals are estimated with the help of a bootstrap procedure (see Efron 1979).

In addition to the calibrated sampling weight, each person and household in the sample is assigned a certain number c of bootstrap weights. Currently $c = 500$ bootstrap draws seem sufficient to estimate the sampling error with high quality, however, an increase to $c = 1000$ is under consideration.

Bootstrap weights are computed based on bootstrap samples. A bootstrap sample is a sample with replacement of size m taken from the original sample of the same size. This means that every original sample unit can appear 0- to m times in the bootstrap sample. In our case, the sampling units are households. The frequency of occurrence f_i^j of observation i in bootstrap sample j , where $j = 1, \dots, c$, multiplied with the calibrated sampling weights w_i of the original sample (see Section 3.3.3) renders the uncalibrated bootstrap weights

$$b_i^{0,j} = f_i^j w_i \quad (8)$$

which are identical for every person in a household as the bootstrap samples are drawn at household level. These weights are then calibrated using the iterative proportional fitting procedure from Section 3.3 but with initial weights $b_i^{0,j}$ instead of the base weights w_i^0 , returning the adapted bootstrap weights b_i^j .

4.1.2. The rescaling bootstrap

Since the microcensus is a sample without replacement drawn from a finite population, the “naïve” bootstrap procedure described above can not be applied in exactly this form. Instead, the “rescaled” bootstrap procedure introduced by Rao and Wu (1988) with the adjustment of using rescaled weights instead of rescaled survey data values (see Rao, Wu, and Yue 1992) is used with the additional modification of selecting bootstrap samples without replacement (see Chipperfield and Preston 2007; Preston 2009), also incorporating the stratification by region r (see Section 3.2). To be more specific, instead of drawing c bootstrap samples with replacement of the same size m_r as the original sample, subsamples without replacement of size $m_r^j = \lfloor m_r/2 \rfloor$ are drawn.

The uncalibrated bootstrap weights for every observation i belonging to region r are then computed for $j = 1, \dots, c$ as

$$b_i^{0,j} = w_i \left(1 - \lambda_r + \lambda_r \frac{m_r}{m_r^j} \delta_{r_i} \right) = w_i f_i^j \quad \forall i \in m_r \quad (9)$$

with

$$\lambda_r = \sqrt{\frac{m_r^j \left(1 - \frac{m_r}{M_r} \right)}{m_r - m_r^j}} \quad (10)$$

where w_i are the calibrated sampling weights of the original sample and $\delta_{r_i} = 1$ if observation i is selected in the subsample m_r^j and 0 otherwise. The $b_i^{0,j}$ are then calibrated to render the adapted bootstrap weights b_i^j as mentioned above in Section 4.1.1.

To determine the standard error and the approximate 95% confidence interval of the population estimate $\hat{\theta}$ of some population parameter θ we use the c bootstrap weights b_i^j to compute c population estimates $\hat{\theta}^j$. The estimated standard error of the population estimate $\hat{\theta}$ is then the standard deviation of these c estimates $\hat{\theta}^j$ with mean $\hat{\theta}$:

$$\text{SE}(\hat{\theta}) = \sqrt{\frac{\sum_j (\hat{\theta}^j - \tilde{\theta})^2}{c-1}}. \quad (11)$$

The 2.5% and 97.5% quantiles of the c bootstrap replicates $\hat{\theta}^j$ correspond to the lower and upper boundaries of the approximate 95% confidence interval (see [Efron 1981](#)).

4.1.3. Rotation

The Austrian microcensus is a rotating quarterly sample survey (see [Section 2.2](#)) where one fifth of the sample is replaced by a new random sample every quarter. The bootstrap procedure takes this into account by drawing bootstrap samples and computing the occurrence frequencies f_i^j for this new sample only while retaining the f_i^j from the previous quarter for the remaining four fifths of the sample.

This way, the overlap of two microcensus samples is considered in cases where the objective is to estimate standard error and confidence interval of change.

4.2. Error approximation

In cases where no bootstrap results are provided by STAT, users can resort to tables containing rough estimates of the relative sampling errors.

The relative sampling error, i.e. the sampling error divided by the estimate, can be expressed for frequency counts of persons and households by means of the binomial approach (see e.g., [Cochran 1977](#)) where the variance of $x = Np$ with sample proportion p and population size N is estimated as

$$\text{Var}(x) = \frac{N(N-n)}{n} p(1-p) \quad (12)$$

Strictly speaking, the frequency counts resulting from a sample without replacement are hypergeometrically distributed, however due to the small sampling fraction, the use of the binomial distribution as an approximation is justified.

Calibration (see [Section 3.3](#)) is ignored by this approximation, only the sampling fractions corresponding to the NUTS-2 regions are considered. For a population frequency count x of persons or households the relative sampling error at confidence level 95%, i.e. with the factor 2 as an approximation of the respective quantile of the normal distribution, is computed as

$$2 * 100 \sqrt{\frac{(N_r - n_r)(N_r - x)}{N_r n_r x}} \quad (13)$$

at NUTS-2 level r and as

$$2 * 100 \sqrt{\sum_r \frac{(N_r - n_r)(N - x)N_r}{n_r N^2 x}} \quad (14)$$

for frequency counts x corresponding to the whole population where the contributions of the NUTS-2 regions to the value x are assumed to be proportional to the size of the respective regions. N is the total population size while N_r and n_r are population and sample size of NUTS-2 region r . If x refers to households, (13) and (14) are computed with M , M_r and m_r instead (see [Section 3.2](#)).

This simple calculation of the sample error is useful for a rough estimation. It gives an idea of the magnitude of the error, however it does not give the exact values. Neither calibration which leads to a reduction of the error nor sample clusters which lead to an increase of the error are taken into account. For both effects no general rule can be given, e.g. sample clusters

lead to bigger errors but the magnitude depends very much on the variable, e.g. errors of variables with a very low cluster correlation are nearly unaffected by cluster sampling. In this way, differences between the bootstrap-errors and the simple approximation vary from estimate to estimate. As an example of the differences between the error approximation and the bootstrap errors, the estimated number of persons unemployed at the age of 15 to 74 in year 2014 are used. This estimated number is 244,883 and the estimated relative standard error calculated on the bootstrap weights is 1.63% whereas the error based on the simple approximation is estimated to be 5.2%. This shows that for the number of unemployed the positive effect of calibration exceeds the negative effect of the cluster sampling.

5. Summary

One decade after the last substantial renewal of the Austrian microcensus the weighting procedure was again revised. Although the method itself is unchanged, i.e. calibration of weights via iterative proportional fitting, the external information used for calibration changed significantly. Some were updated and a completely new specification was launched, the employment status from administrative data. This led to a big improvement of the quality of the MC due to a reduction of the non-response bias as well as the standard error of variables of great interest.

For the near future there are two important tasks to implement: Firstly, the aim is to enable users to compute the exact standard errors and confidence intervals on their own, i.e. to provide the required tools to make use of the bootstrap weights and eliminate user's dependency on the error approximation.

Secondly, the focus on the longitudinal dimension of the MC, to provide information on labour market dynamics and flows, is of increasing importance. Therefore, the weighting of longitudinal data, with its related issues such as panel attrition, consistency between cross sectional and longitudinal results, plausibility along the time line, is an area of constant improvement.

Acknowledgement

The authors would like to thank the editor and the anonymous reviewer for their careful work and their helpful and constructive comments.

References

- Chipperfield J, Preston J (2007). "Efficient Bootstrap for Business Surveys." *Survey Methodology*, **33**(2), 167–172.
- Cochran W (1977). *Sampling Techniques, 3rd Edition*. John Wiley.
- Ediev D (2007). "On Projecting the Distribution of Private Households by Size." *Vienna Institute of Demography of Austrian Academy of Sciences. Working Paper*, **4**.
- Efron B (1979). "Bootstrap Methods: Another Look at the Jackknife." *The Annals of Statistics*, pp. 1–26.
- Efron B (1981). "Nonparametric Estimates of Standard Error: The Jackknife, the Bootstrap and Other Methods." *Biometrika*, **68**(3), 589–599.
- Gumprecht D, Oismüller A (2013). "Non-Response im Mikrozensus." *Statistische Nachrichten*, **11**, 1046–1061 [in German].

- Haslinger A, Kytir J (2006). “Stichprobendesign, Stichprobenziehung und Hochrechnung des Mikrozensus ab 2004.” *Statistische Nachrichten*, **6**, 510–519 [in German].
- Kytir J, Stadler B (2004). “Die kontinuierliche Arbeitskräfteerhebung im Rahmen des neuen Mikrozensus.” *Statistische Nachrichten*, **6**, 511–520 [in German].
- Meraner A, Gumprecht D, Kowarik A (2015). “Die neue Hochrechnung des Mikrozensus - Methodenbeschreibung.” *Technical report*, Statistik Austria [in German]. URL http://www.statistik.at/web_de/statistiken/menschen_und_gesellschaft/arbeitsmarkt/index.html.
- Potter FJ (1990). “A Study of Procedures to Identify and Trim Extreme Sampling Weights.” In *Proceedings of the American Statistical Association, Section on Survey Research Methods*, volume 225230.
- Potter FJ (1993). “The Effect of Weight Trimming on Nonlinear Survey Estimates.” In *Proceedings of the American Statistical Association, Section on Survey Research Methods*, volume 758763.
- Preston J (2009). “Rescaled Bootstrap for Stratified Multistage Sampling.” *Survey Methodology*, **35**(2), 227–234.
- Rao J, Wu C, Yue K (1992). “Some Recent Work on Resampling Methods for Complex Surveys.” *Survey methodology*, **18**(2), 209–217.
- Rao JN, Wu C (1988). “Resampling Inference with Complex Survey Data.” *Journal of the american statistical association*, **83**(401), 231–241.

Affiliation:

Angelika Meraner, Alexander Kowarik

Methods Unit

Statistics Austria

Guglgasse 13, 1110 Vienna, Austria

E-mail: angelika.meraner@statistik.gv.at, alexander.kowarik@statistik.gv.at

URL: <http://www.statistik.at>

Daniela Gumprecht

Department of Social Statistics

Statistics Austria

Guglgasse 13, 1110 Vienna, Austria

E-mail: daniela.gumprecht@statistik.gv.at

URL: <http://www.statistik.at>