Using Auxiliary Information to Improve Agricultural Statistics – Advantages of the Calibration Approach over Poststratification Weights

Lucian Stanca
Technische Universität München, Germany

Daniel Hoop
Agroscope, Switzerland

Johannes Sauer
Technische Universität München, Germany

Abstract

Official statistics are often based on samples representing a certain population. Because participation in a sample is usually voluntary, bias might result from so-called non-sampling errors such as nonresponse. Weighting procedures are intended to correct these errors by assigning a certain weight to each observation in the sample. In many official agricultural statistics, such as the Bavarian Agricultural Report, poststratification is used. In this process, the population is stratified according to different dimensions (e.g. farm type, farm location and farm size) and weights are assigned to all farms in a stratum so that the sum of the weights in that stratum corresponds to the number of observations in that stratum in the population. However, when estimating the population average, important characteristics (such as the farm size) may still be biased. Using a Bavarian farm sample, the present study shows how the so-called calibration approach, utilising auxiliary variables to adjust weights, outperforms the poststratification procedure in terms of estimating important population characteristics.

Keywords

calibration; unit nonresponse bias; auxiliary information; design-based estimation; weighting adjustment.

1 Introduction

The application of statistical inference can be an easy and straightforward endeavour in presence of ideal statistical conditions. These conditions are given when non-sampling errors are absent, such as nonresponse, measurement error or frame imperfection. However, in practice, this is seldom the case.

In absence of ideal statistical conditions, when carrying out parameter estimations, the error is formed by two components: sampling or random error and non-sampling or systematic errors. The first-mentioned is impossible to eliminate completely owing to the nature of sampling, where parameters are estimated only by means of a sample. The sampling error is characterised by the variance or standard error of the estimates, given that other errors are absent. The non-sampling errors, such as coverage or nonresponse errors, might be reduced at the sampling stage. However, it is not always possible to intervene substantially at the sampling stage in order to reduce non-sampling errors. Once the sampling is concluded, different techniques can be employed at the estimation stage to adjust the weights according to auxiliary information.

An important source of non-sampling error is nonresponse. In the literature, nonresponse is regularly differentiated into two types: item and unit nonresponse (SÄRNDAL and LUNDSTRÖM, 2005). Item nonresponse occurs when, for various reasons, only part of a questionnaire is answered, whereas unit nonresponse is given when the survey response from selected elements is not obtainable (KALTON and KASPRZYK, 1986; BETHLEHEM, 2009). Thus, with unit nonresponse, the required sample data determined by the sample allocation is not fully obtained. This is a major concern because it is an important source of bias for estimates of population features. Moreover, in case of a reduced number of respondents, the variance might augment (SÄRNDAL and LUNDSTRÖM, 2005). However, this is considered a minor problem when facing bias.

Statistical inference and, more specifically, estimation is applied in many fields. Agriculture is no exception: The Bavarian Agricultural Report (BAR),

DOI: 10.30430/gjae.2022.0294

estimating farm income, has its origin in a resolution of the Bavarian State Parliament from 1971. From its first publication, a core issue of the report was the description of the economic situation of the Bavarian agriculture. The survey that serves for the estimation is based on the information contained in the financial statements provided by the Bavarian test farms network. Prior to 1978, the information was presented as unweighted averages. Thereafter, a weighted estimation was introduced.

The aim of the survey is to obtain as much information as possible about different population parameters, such as means and ratios, by using statistical inference, so that it can be used to describe the whole population. However, information is also needed at different domain levels (such as regions or farm types). The disaggregated information allows, for instance, to better formulate and monitor regional policies such as the allocation of funds.

The BAR sampling design is based on a stratified sampling procedure (i.e. the target population is subdivided according to different dimensions such as farm size). Because the participation in the survey is not mandatory, there is both overrepresentation and underrepresentation in most strata derived from unit nonresponse. Nonresponse is an intrinsic aspect of surveys nowadays, and it is also the case for the Bavarian test farms network; it is undesirable because it can substantially affect the accuracy of estimates. To reduce the adverse impacts of nonresponse, the current estimation method employs a poststratified estimator. Nevertheless, the quality of the estimation can still be improved by adjusting the estimation weights using additional information through quantitative auxiliary variables. This is what the so-called calibration approach pursues.

The objective of this paper is to provide a step-by-step procedure for practitioners in agricultural statistics – a cookbook so to speak – to improve the quality of estimates by applying the calibration approach, because as far as we know, this method is not widely employed in Germany for income estimation in the agricultural sector. For this purpose, we used the BAR sample and auxiliary information from the Integrated Administration and Control System (IACS) dataset, in order to illustrate the improvements that can be achieved by means of the calibration approach, compared with poststratification, but, as well, to show that the IACS dataset could represent a reliable source of information for such a purpose.

With this objective, the next section presents the employed datasets that served for the weights adjustment and for the estimation. The third section describes the methodological framework used for the current estimation (as conducted by the Bavarian State Ministry of Food, Agriculture and Forestry) and for the calibration approach. Next, in the fourth section, the different estimation methods are put side by side to determine empirically the advantages of the adjusted weights obtained by means of the calibration approach over the poststratified estimator. Last, we present the main conclusions derived from the study.

2 Data

For the present analysis, two data sources were used from the IACS and from the Bavarian test farms network. Data refer to the fiscal year 2013/2014.

The IACS dataset consists of a series of interconnected databases used to monitor the direct payments of the Common Agricultural Policy. In this analysis, the dataset is used to determine the target population and to provide the auxiliary variables for calibrating the design weights, such as the hectares of the utilised agricultural area (UAA) and the number of livestock units (LSU). The target population consists of farms with a standard output (SO) greater than 25,000 €. More precisely, the small and part-time farms are represented by those with a SO between 25,000 € and 50,000 €, while full-time farms are represented by those with a SO of at least 50,000 €. The resulting target population consists of 56,569 farms. Of these, 35,972 are full-time farms and 20,597 are small and part-time farms.

The Bavarian test farms network provides the sample that serves for the estimation of population characteristics (e.g. farm income) and consists of 2,696 farms for the analysed year. It contains 1,975 full-time farms and 721 small and part-time farms. The average nonresponse rate is about 62% and especially by farms under 30 ha which explains the slight overrepresentation of full-time farms in the sample. By using a unique identifier, auxiliary information from the IACS dataset can be merged to the Bavarian test farms.

3 Methodology

The sampling design is based on a stratified sample with H strata and $h = \{1, \dots, H\}$ in order to achieve two key objectives. In the first place, it aims to improve the quality of estimation by creating homogeneous strata from a heterogeneous sample. By this means, it is intended to obtain sample units with a variance within every stratum, smaller than the variance between different strata (HANSEN et al., 1953; COCHRAN, 1977; THOMPSON, 2002). Secondly, there is also a need for information at different levels, as it is the domain level.

Predefined by the Bavarian State Ministry of Food, Agriculture and Forestry, the following criteria were taken into account to determine the strata:

1. Type of engagement: full-time farms, and small and part-time farms.
2. Farm location: northern and southern Bavaria.
3. Farm type: field crops, grazing livestock, granivores, mixed crops–livestock, viticulture, dairying, horticulture.
4. Farm size: For the part-time farms, only size classes from >7.5 to ≤10 ha and from >10 to ≤30 ha are relevant, whereas for the full-time farms, only those from >10 to ≤30 ha, from >30 to ≤60 ha and from >60 to ≤200 ha are taken into account.

As described in the introduction, estimators derived from samples may suffer from errors such as nonresponse and frame imperfection. These errors can be mitigated by introducing weights to estimate sample means. Currently, a poststratified estimator is used for the BAR, based on the four already described criteria. However, this estimator may still provide biased estimates, e.g. in the presence of correlation between target variables and response mechanism. This study determines to which extent a more advanced methodological approach would be more suitable, i.e. the calibration approach that adjusts given weights (e.g. design weights) according to known population totals.

3.1 Poststratified Estimator

Under ideal statistical conditions, sample data, $s = \{1, \dots, i, \dots, n\}$, is expanded to the population level, $U = \{1, \dots, i, \dots, N\}$ with $s \subset U$, by means of design weights, which correspond to the inverse of the inclusion probability, π_i. The design weights reflect the number of units in the population that are represented by one sampled unit. Thus, the weights are always greater than or equal to one, because each element is representing at least itself.

However, in the presence of nonresponse, $r \subset s$ with $r = \{1, \dots, i, \dots, m\}$, the design weights are considered to be on average too small to yield adequate estimates, so that they lead to an underestimation of the target variable, Y. Thus, a poststratified estimator is used for the current estimation, as described by THOMSEN (1973) or COCHRAN (1977), where weights are adjusted in correspondence with the response rate

$$\hat{\bar{Y}}_{PS} = \frac{1}{N} \sum_{h=1}^{H} \frac{N_h}{m_h} \sum_{i=1}^{m_h} y_{ih}, \qquad (1)$$

where H represents the number of strata,
N represents the number of units in the total population,
n_h represents the number of observations to be sampled in the hth stratum (not considering nonresponse), while
N_h and m_h represent the number of units and respondents in the hth stratum, respectively.

By this means, the estimation weights, $w_{PS_h} = N_h/m_h$, are corrected accounting for nonresponse, (w_{PS_h} stands for the poststratified weight a farm has in the hth stratum). In fact, the estimation can also be considered, as suggested by VANDERHOEFT (2001), as a two-step correction procedure
$w_{PS_h} = \frac{N_h}{n_h}\frac{n_h}{m_h} = \frac{d_h}{rr_h} = \frac{N_h}{m_h}$ (where $d_h = \frac{N_h}{n_h}$ represents the design weight and $rr_h = \frac{m_h}{n_h}$ represents the response rate for the hth stratum, respectively). In the first stage, weights are adjusted for nonresponse (response rate), from r (respondents) to s (sample), and in the second stage, the information is expanded from the sample, s, to population level, U, correcting thus for sampling errors. However, when response behaviour across strata is not homogeneous (e.g. some farm types have a lower response rate; LITTLE, 1986) and/or there is an underlying correlation between the target variables and the response mechanism (e.g. farms with higher income tend to have a higher response rate; BETHLEHEM, 2009), this method fails to produce unbiased estimates. This attribute can be improved by a more thorough adjustment or, literally, by the calibration approach (KALTON and FLORES-CERVANTES, 2003). Moreover, as described by SÄRNDAL and LUNDSTRÖM (2005), we will also be able to obtain consistent weights for a multipurpose survey.

3.2 Calibration Estimator

As previously explained, when facing nonresponse the design weights are too small, on average, to represent the population. For this reason, the design weights need to be adjusted to be in line with known external information. The calibration estimator, widely employed in the last decades, exploits the information obtained from auxiliary variables, adjusts the design weights, reduces bias (i.e. reduces differences in the sample and the population total) and increases efficiency (i.e. reduces variance of the estimator). As it turns out, the poststratified estimator is a special case of the calibration estimator, because it only corrects for the number of observations in different strata but does not use further auxiliary variables. The poststratified estimator and the calibration estimator are similar when facing limited nonresponse rate or complete response (SÄRNDAL and LUNDSTRÖM, 2005).

The basic idea of the calibration approach is to determine new weights, w_i, which satisfy the condition of being as close as possible to the design weights, d_i, according to the objective function (distance function), and fulfilling the constraints of the calibration function (DEVILLE and SÄRNDAL, 1992). This can be translated mathematically into minimising the distance function:

$$\min_{w_i} D(w,d) = \sum_{i=1}^{n} d_i G(w_i, d_i), \tag{2}$$

where d_i represents the design weights and $G(w_i, d_i)$ is the distance function, satisfying the calibration equations that serve as benchmark:

$$\sum_{i=1}^{n} w_i x_{ij} = X_j, \tag{3}$$

where x_{ij} stands for the jth auxiliary variable for the ith unit, X_j represents the calibration totals of the jth auxiliary variable, and w_i is the calibration weight of the ith unit, subjected to the following boundary constraints:

$$L \leq \frac{w_i}{d_i} \leq U, \text{ with } 0 \leq L \leq 1 \leq U. \tag{4}$$

When $w_i = d_i$, then there is no need for an adjustment. According to calibration equations, by using the new determined weights, it should be possible to estimate the auxiliary variables' totals with zero variance.

When using the calibration approach, a certain response propensity is implicitly assumed, which can be inverse linear, exponential, logistic, or quadratic. Therefore, it is highly important to ensure that bias is reduced, independently of the response mechanism. For this objective, in a first phase, we have determined the relationship between the auxiliary variables and the parameters of interest. In a second phase, conditioned by the first, the distance function that best meets our objectives has been chosen (HAZIZA and LESAGE, 2016). To ensure that calibration weights are neither negative nor extreme, boundaries for the design weight adjustment factor are introduced.

However, regardless of the chosen calibration method and the consistency of the auxiliary information, we must be aware that, regarding the variable of interest, bias could still be present when there is nonresponse. The key question is then, how to reduce bias to an acceptable minimum. HAZIZA and LESAGE (2016) showed that improved auxiliary information, i.e. more 'informative' variables, is the fundamental answer to this issue. Therefore, the quality of the adjusted weights depends on the availability of good and robust auxiliary information. For this reason, building consistent auxiliary vectors is crucial to achieve our target.

When non-random nonresponse is not neutralised through adjustments, it can be an important cause of bias, especially when variables determining nonresponse are related to target variables. Even when a small variance is obtained, the quality of estimates could be strongly affected. Therefore, the focus should be on reducing bias as far as possible. For this, the selection of auxiliary variables is the central concept within this approach. Even though, as mentioned by SÄRNDAL and LUNDSTRÖM (2005), the selection process of the relevant variables is to a certain extent based on a heuristic approach, the coefficient of correlation between target variables and auxiliary variables could form a decisive criterion.

For our analysis, considering the available information from the IACS dataset, we have employed the information at population level (U), where information for respondents (r) is available as well as the vector of population totals ($X = \sum_U x$).

4 Estimates Comparison

As highlighted previously, it is important to detect in a first stage the relationship between the auxiliary variable(s) and the target variable(s), in order to be able to choose in a second stage the adequate distance function. Using correlation analysis, we have chosen LSU for dairying, grazing livestock, granivores, and mixed crops–livestock. UAA has been employed for field crops, viticulture, and horticulture. The calibration of specific variables for specific farm types was achieved by setting these variables to zero for other (irrelevant) farm types.[1] To illustrate the quality of the results, we have estimated the following target variables: total income, profit, and standard gross margin (SGM). The relationship between auxiliary variables and target

[1] This means that the sample and the target population have an average value of zero for these variables. Therefore, the target is already met, and these variables are not calibrated in strata where they are of no relevance.

variables proved to be roughly linear with an average correlation coefficient of 0.53.

For our purpose, according to HAZIZA and LESAGE (2016), the most suitable distance function, G, would be the truncated linear model:

$$G(w_i, d_i) = (w_i - d_i)^2/2d_i. \quad (5)$$

To allow meaningful interpretation of the dataset employed in the present work, the distance function should yield strictly positive weights.

Additionally, this function allows to control the range of the correction weights, also called g-$weights$, g_{ih}, i.e. to assure that weights are neither negative nor too large. The algorithm used for the calibration has been implemented using the open-source software R (R CORE TEAM, 2017) extended by the 'sampling' package (TILLÉ and MATEI, 2016).

To describe the Bavarian agriculture as accurately and precisely as possible, the weights have been calibrated on different levels that were chosen according to a heuristic process:

- Region: land area and LSU
- Strata: number of farms

The calibration matrix, X_j, has been built using the information according to the previously mentioned levels to assure consistency not only at an aggregated level but also at a disaggregated level, i.e. population and domain level.

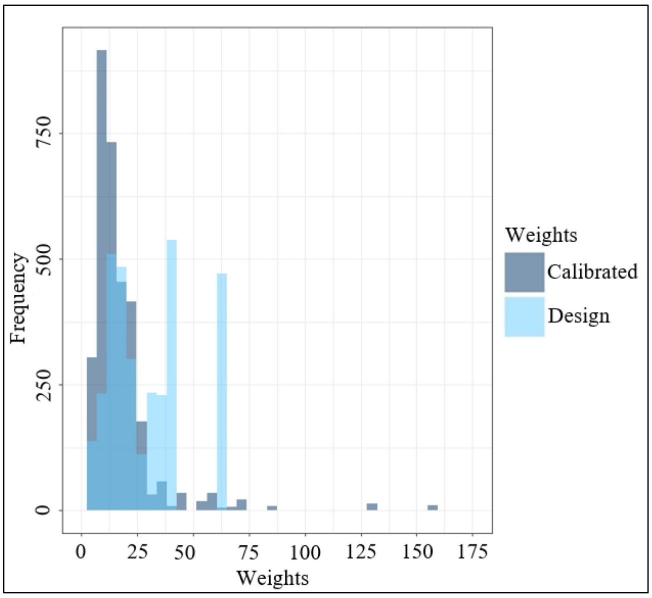
The final weights, i.e. the calibrated weights, have been calculated by multiplying the design weights, d_{ih}, with the resulting factor of correction (g-$weights$) obtained from the optimisation program.

4.1 Distribution of Weights

Figure 1 shows the distribution of initial or design weights (light blue) and the corrected or calibrated weights (grey) regarding the frequency of their occurrence. Overlaps are represented in dark blue. From this chart, we can assure that the calibrated weights are neither negative nor less than one. Noteworthy is that the calibration procedure has determined a reduction in the distribution of the weights, clustering them towards the lower bottom, prompting by this means a reduction in frequency of calibrated weights above 25. On the other hand, the calibration has produced a few relatively high weights, which may suggest that the design weights due to nonresponse would have been relatively too low for some observations.

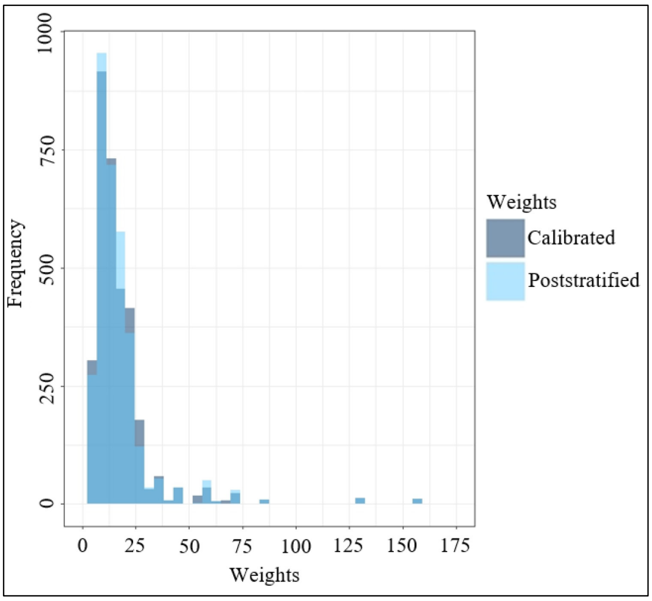
When confronting the calibrated weights (grey) with the poststratification weights (light blue) in Figure 2, it can be observed that the distributions of both weights have similar patterns (overlaps are represented in dark blue). This result is to some extent expected, because poststratification is also considered to belong to the greater family of weighting adjustment methods. Nevertheless, as will be shown below, the relatively small differences are key for the quality of the estimates. The large calibration weights are nearly identical to those originating from poststratification.

Figure 1. Distribution of calibrated and design weights for 2013/2014



Source: own calculations

Figure 2. Distribution of calibrated and poststratified weights for 2013/2014



Source: own calculations

Table 1. Ad hoc bias comparison for livestock units (LSU) and utilised agricultural area (UAA) between population and estimated means (calibration and poststratification) at population and at domain level

Characteristic		IACS	Estimation using calibrated weights		Estimation using poststratified weights	
			Average	*Discrepancy*	*Average*	*Discrepancy*
Bavaria	*LSU*	51.96	51.93	−0.06%	35.74	−31.23%
	UAA	45.45	45.69	0.52%	29.44	−35.22%
North	*LSU*	52.17	52.17	0.00%	31.04	−40.50%
	UAA	50.64	50.83	0.39%	27.24	−46.20%
South	*LSU*	51.83	51.78	−0.10%	39.91	−22.99%
	UAA	41.95	42.21	0.62%	31.40	−25.16%

IACS: Integrated Administration and Control System
Source: own calculations

4.2 Accuracy

Bias measures the accuracy of an estimated parameter, and it can be calculated only for variables that are available for the entire population and in the sample, such as LSU, UAA and SGM. For these variables, we calculated the population mean and compared it with a) the weighted mean calculated with the calibrated weights and b) the weighted mean calculated with poststratified weights.

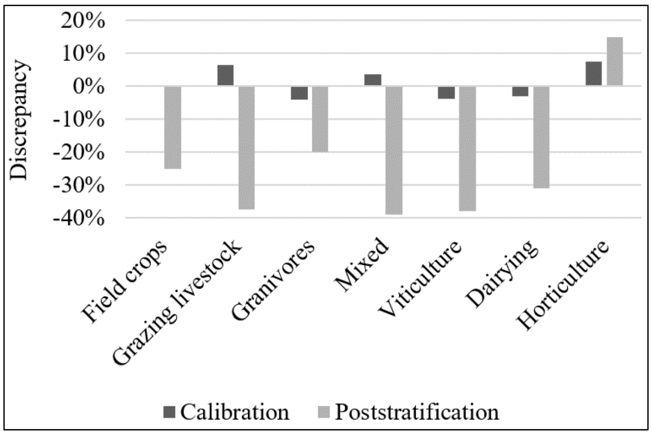
Table 1 displays the true mean of the Bavarian population and the two weighted means (derived from the sample) computed at population (Bavaria) and at regional level (North and South Bavaria), a more aggregated domain level. Because we did not calibrate LSU and UAA for all farm types, it is noteworthy that the calibration approach does not estimate exactly the population means. The estimated means, however, match relatively closely the population means. This stands in stark contrast to the means estimated using poststratified weights. For LSU at population level, the estimated mean amounts to 69% of the true value, and for the UAA it is even more underestimated, with 65% of the actual value. At the regional level, poststratified weights still considerably underestimate LSU and UAA, whereas the calibrated weights maintain the good estimation performance as it was the case at the population level.

However, as already stated, we were also interested in obtaining accurate estimates at an even more disaggregated domain level. Therefore, the same comparison has been established but, in this case, according to the farm type. Likewise, at farm type level (Figure 3), the estimates for LSU and UAA using weights corrected by means of the calibration approach provide again a better adjustment with respect to the population mean as compared with poststratification. In the case of field crops, the calibration estimation is identical to the population mean, whereas the poststratification underestimates the true value by about 25%. At this domain level, the calibration estimates do not perform as good as at a more aggregated level. Nevertheless, compared with poststratification, calibration provides more satisfactory results because the relative difference is negligible. A possible explanation for the constant underestimation by poststratification weights is that for the presented levels, the farms delivering data have on average a smaller UAA and less LSU than the actual farms in the population, independently of the presented characteristic. The only exception is horticulture, where both weights overestimated the population mean. However, the calibration approach overestimates only half as much as the poststratification weights.

By means of the ad hoc procedure, we could assure that the calibration approach assures, to a certain extent, the unbiasedness in comparison with the poststratified estimates. Thus, we can expand the analysis to additional variables.

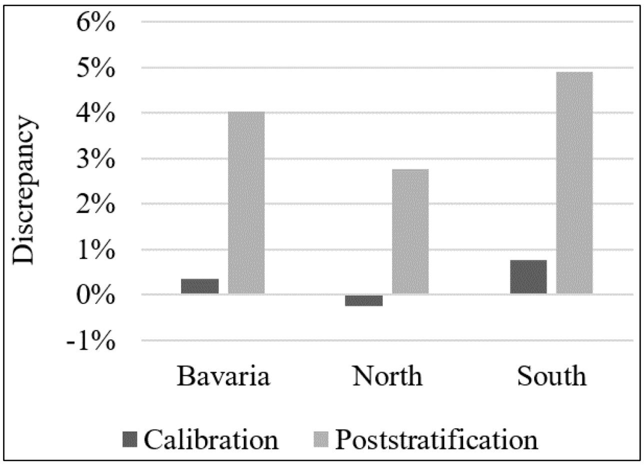
Because LSU and UAA were used as benchmark variables for the calibration, the previous results were as expected. Therefore, to assess further the quality of

Figure 3. Ad hoc bias comparison for livestock units and utilised agricultural area between population and estimated means (calibration and poststratification weights) at farm type level



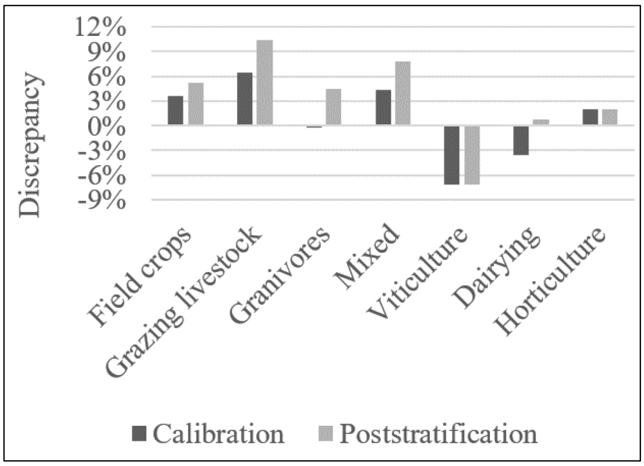
Source: own calculations

Figure 4. Ad hoc bias comparison for standard gross margin between population and estimated means (calibration and poststratification weights)



Source: own calculations

Figure 5. Ad hoc bias comparison for standard gross margin between population and estimated means (calibration and poststratification weights) at farm type level



Source: own calculations

the calibrated estimation, we performed an estimation for a variable, the SGM, that we did not calibrate for but for which information exists at population and at sample level. At a more aggregated level, as can be noted in Figure 4, the estimation using calibrated weights produces more accurate results than poststratification at all three levels. The results suggest that farms having a higher SGM are more likely to participate in the sample. Weight calibration manages to compensate for this response behaviour, whereas poststratification fails to do so.

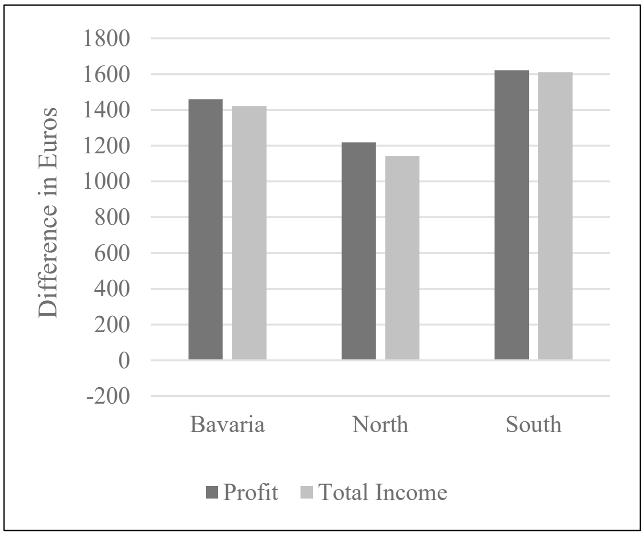
Even at a more disaggregated level, as shown in Figure 5, the calibrated weights prove to yield results that are more accurate than the poststratified results. However, like in the case of dairy farms, calibration underestimates the true value, whereas poststratification provides a more accurate result. Regarding viticulture and horticulture, both approaches underestimate or overestimate the actual value to the same extent, respectively. However, in both cases the differences are negligible.

Figure 6 shows the differences between poststratification and calibration estimates (i.e. poststratification estimate minus calibration estimate) at population and regional level for two target variables, namely total income and profit. Given that poststratified weights are based on those farms that actually provide data, we might assume that, on average, there is a higher predisposition for more profitable farms to deliver accountancy information (see also Figure 4). Therefore, as stressed in Section 3.1, poststratified weights fail to produce unbiased estimates when there is an underlying correlation between the target variables and the response mechanism.

Figure 7 presents the difference between poststratification and calibration estimates at a more disaggregated level, i.e. at farm type level. Viticulture and horticulture, shown in Figure 7, are farm types where the poststratified weights seem to yield similar results as the calibrated ones. This would mean that the data-delivering farms for these farm types could represent the average profit or total income. This observation seems to corroborate the findings from Figure 5. Regarding the rest of farm types, assuming that the calibrated estimates are almost unbiased, the responding farms seem to have higher profits or total incomes on average, i.e. responding farms are more profitable on average, as compared with the population.

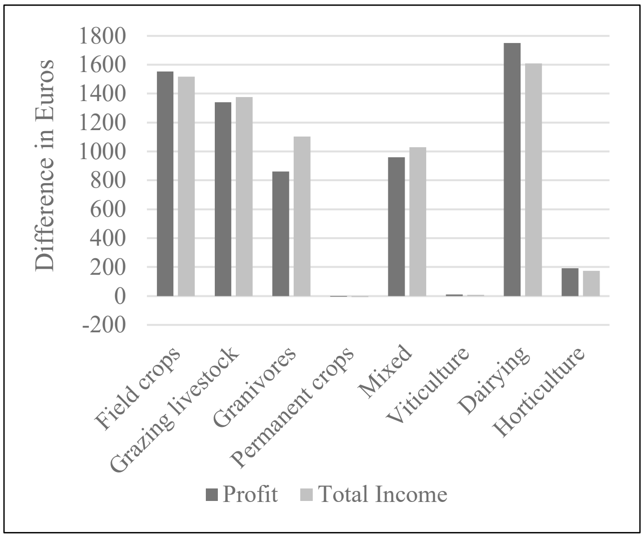
In theory, with increased sample size, the probability to yield biased estimators decreases. Because dairy farms are overrepresented in the sample, one could expect the bias of the poststratified estimator to be low. Given systematic nonresponse, however, this is not the case: despite the large sample size, poststratification weights seem to underestimate target variables of dairy farms.

Figure 6. Difference between poststratification and calibration estimates at population and regional level



Source: own calculations

Figure 7. Difference between poststratification and calibration estimates at farm type level



Source: own calculations

5 Conclusion

Sampling is an intrinsic part of official statistics because it substantially reduces the costs but also the tedious process of a census. However, sampling has its drawbacks because it is connected with both sampling and non-sampling errors. The impacts of these errors are important because they can significantly determine the illustration of a population that is intended to be recreated by means of estimation. The recreation is obtained by expanding the sampling information through weights. Because poststratification takes into account only the number of farms, calibrating weights by using auxiliary variables is a potential option to improve the accuracy of an estimator.

Poststratified weights are currently used for the BAR to adjust the estimation weights. However, under certain circumstances, such as correlation between target variables and response mechanism, the adjusting power of this approach is limited. The quality of weights can be further improved by including additional (auxiliary) variables. This paper shows that the one-step calibration approach is a useful method to further improve the current estimation method by reducing the systematic error.

Through the calibration approach, we have shown that the response mechanism could be disregarded by determining first the relationship between target and auxiliary variables. According to the determined relationship, the calibration function is chosen such that, independently of the response mechanism, a reduction in bias is assured.

Moreover, the IACS dataset proved to be a reliable source of robust and consistent auxiliary information that served to obtain a reduction in bias and to assure a consistent estimation. Obviously, when estimating a myriad of variables, the extent to which the calibration approach proves to have a positive effect on bias depends on the correlation between the variables of interest and the auxiliary variables used for the benchmark. For important variables, however, empirical results showed that calibration could reduce bias.

Consequently, it can be concluded that even obviating the modelling of the nonresponse mechanism, the estimates calculated using weights adjusted through the calibration approach perform better than poststratification estimates. Therefore, in the presence of 'good' auxiliary information, the calibration approach could provide accurate and efficient estimates that can serve as robust policy instruments.

References

BETHLEHEM, J. (2009): Applied survey methods: A statistical perspective. John Wiley & Sons, New York.

COCHRAN, W.G. (1977): Sampling techniques. 3rd edition. John Wiley & Sons, New York.

DEVILLE, J.-C. and C.-E. SÄRNDAL (1992): Calibration estimators in survey sampling. In: Journal of the American Statistical Association 87 (418): 376-382.

HANSEN, M.H., W.N. HURWITZ and W.G. MADOW (1953): Sample survey methods and theory. Vol. 1. Methods and applications. John Wiley & Sons, New York.

HAZIZA, D. and É. LESAGE (2016): A discussion of weighting procedures for unit nonresponse. In: Journal of Official Statistics 32 (1): 129-145.

KALTON, G. and I. FLORES CERVANTES (2003): Weighting methods. In: Journal of Official Statistics 19(2): 81-97.

KALTON, G. and D. KASPRZYK (1986): The treatment of missing survey data. In: Survey Methodology 12 (1): 1-16.

LITTLE, R.J.A. (1986): Survey nonresponse adjustments for estimates of means. In: International Statistical Review/Revue Internationale de Statistique 54 (2): 139-157.

R CORE TEAM (2017): R: A language and environment for statistical computing [Internet]. R Foundation for Statistical Computing, Vienna.

SÄRNDAL, C.-E. and S. LUNDSTRÖM (2005): Estimation in surveys with nonresponse. John Wiley & Sons, New York.

THOMPSON, S.K. (2002): Sampling. John Wiley & Sons, New York.

THOMSEN, I. (1973): A note on the efficiency of weighting subclass means to reduce the effects of nonresponse when analyzing survey data. In: Statistisk Tidskrift 11 (4): 278-283.

TILLÉ, Y. and A. MATEI (2016): R package 'sampling': Survey sampling. https://CRAN.R-project.org/package=sampling.

VANDERHOEFT, C. (2001): Generalised calibration at Statistics Belgium: SPSS module g-CALIB-S and current practices. Institut National de Statistique, Belgium.

Contact author:
DANIEL HOOP
Tänikon 1, 8356 Ettenhausen, Switzerland
e-mail: daniel.hoop@agroscope.admin.ch