Statistica & Applicazioni - Vol. XX, n. 2, 2022, p. 114-148
First published online may 2025

FINITE POPULATION INFERENCE FOR SKEWNESS MEASURES

Leo Pasquazzi *

SUMMARY

In this article we consider Bowley's skewness measure and the Groeneveld-Meeden b_3 index in the context of finite population sampling. We employ the functional delta method to obtain asymptotic variance formulae for plug-in estimators and propose corresponding variance estimators. We then consider plug-in estimators based on the Hájek cdf-estimator and on a Deville-Särndal type calibration estimator and test the efficiency of the corresponding variance estimators and normal confidence intervals in a simulation study.

Keywords: Skewness, Bowley's Index, Groeneveld-Meeden b_3 Index, Finite Pop- Ulation Inference
DOI: 10.26350/999999_000057
ISSN: 1824-6672 (print) 2283-6659 (digital)

1. INTRODUCTION

In order to allow for asymmetry and other deviations from normality, Pearson (1895) introduced his well-known distribution family and proposed

$$\zeta := \frac{\text{mode} - \text{mean}}{\text{standard deviation}}$$

as a measure for asymmetry. Curiously, only for Type III distributions (location scale transformations of gamma distributions) Pearson uses the difference between the mean and the mode in the numerator rather than the other way round. As Pearson considered the method of moments to fit his distributions ("generalized probability curves") to empirical data, he provided formulae to compute ζ from the standardized third central moment, i.e. from μ_3/σ^3. Despite its limitations (e.g. nonexistence of moments, large sample variability), μ_3/σ^3 became very popular in the aftermath and is still widely used today.

Some years after the publication of Pearson's paper, Bowley (1901) proposed

$$\frac{\text{upper interquartile range} - \text{lower interquartile range}}{\text{central interquartile range}}$$

as a measure for skewness. Actually, in the first edition of his 1901 book "Elements of Statistics", Bowley did not write down the explicit definition of this

* Dipartimento di Statistica e Metodi Quantitative - Università degli Studi di Milano-Bicocca - Piazza dell'Anteneo Nuovo, 1 - 20126 MILANO (e-mail: leo.pasquazzi@unimib.it).

measure, but the underlying idea can be inferred from some comments on an example about the distribution of wages among branches of the Amalgameted Society of Engineers (see pp. 134-136 in Bowley, 1901). For sure the explicit definition of Bowley's measure appears in the fourth edition of his book which was published in 1920 (see p. 116 in Bowley, 1920). Perhaps it appears also in the second and/or third edition which where published in 1902 and 1907, respectively, but I was not able to verify this since I had no access to these intermediate editions.

For estimation purposes in the Box-Cox model, Hinkley (1975) considers an immediate generalization of Bowley's measure which is given by

$$b_2(r) := \frac{F^{-1}(1-r) + F^{-1}(r) - 2F^{-1}(0.5)}{F^{-1}(1-r) - F^{-1}(r)}, \quad r \in (0,1),$$

where $F^{-1}(r) := \inf\{x \in \mathbb{R} : F(x) \geq r\}$ denotes the r-quantile of the distribution function $F(x)$. Of course, $b_2(r)$ coincides with Bowley's measure when $r = 0.25$. To save notation, we will just write b_2 when $r = 0.25$.

In close relation to $b_2(r)$, Groeneveld and Meeden (1984) propose

$$b_3 := \frac{\int_0^1 \{F^{-1}(1-r) + F^{-1}(r) - 2F^{-1}(0.5)\}dr}{\int_0^1 \{F^{-1}(1-r) - F^{-1}(r)\}dr} = \frac{\mu - \nu}{E(|X - \nu|)}$$

as a global measure for skewness. Here, we used X to denote a random variable with distribution function $F(x)$, μ to indicate $E(X)$ and ν to indicate the median $F^{-1}(0.5)$. As shown in Groeneveld and Meeden (1984), $b_2(r)$ and b_3 are both consistent with the partial skewness ordering relation of van Zwet (1964).

The present article is about estimation of $b_2(r)$ and b_3 in the context of finite population sampling. In Section 2 we introduce plug-in estimators based on the Hájek cdf-estimator and on a Deville-Särndal type calibration estimator. We also provide asymptotic variance formulae and corresponding variance estimators. In Section 3 we test the estimators in a simulation study. Normal confidence intervals are also tested. Conclusions and final remarks end this paper in Section 4.

2. ESTIMATORS, ASYMPTOTICE VARIANCE AND VARIANCE ESTIMATORS

Let $U_N := \{1, 2, \ldots, N\}$ be the set of labels which identify the units of a finite population, let $y_1, \ldots, y_N$ *be the values taken on by a study variable* Y and let $x_1, \ldots, x_N$ be the values taken on by an auxiliary variable X. The finite population cdf of Y is given by $F_N(t) := \frac{1}{N}\sum_{i=1}^{N} I(y_i \leq t)$ and henceforth we assume that $b_2(r)$ and b_3 refer to $F_N(t)$.

Now, let $s := \{i_1, i_2, \ldots, i_n\} \subset U_N$ be a sample taken from U_N (n denotes the sample size). In this article we assume the X-values to be known for every unit $i \in U_N$, but that the Y-values be known only for $i \in s$. For $i, j \in U_N$ we denote the first and second order sample inclusion probabilities by π_i and π_{ij}, respectively.

As already mentioned above, in this article we consider plug-in estimators based on

a) the Hájek estimator

$$\widehat{F}_{Ha}(t) := \frac{1}{\widehat{N}} \sum_{i \in s} \frac{1}{\pi_i} I(y_i \leq t),$$

where $\widehat{N} := \sum_{i \in s} \frac{1}{\pi_i}$ is the Horvitz-Thompson estimator of N;

b) and the calibration estimator (see Deville and Särndal, 1992)

$$\widehat{F}_{cal}(t) := \frac{1}{N} \sum_{i \in s} w_i I(y_i \leq t)$$

with sample weights w_i given by

$$w_i := \frac{\exp(\widehat{\beta}_0 + x_i \widehat{\beta}_1)}{\pi_i} \tag{1}$$

where $\widehat{\beta} := [\widehat{\beta}_0, \widehat{\beta}_1]$ solves the calibration equations

$$\sum_{i \in s} \frac{\exp(\widehat{\beta}_0 + x_i \widehat{\beta}_1)}{\pi_i} = N$$

$$\sum_{i \in s} \frac{\exp(\widehat{\beta}_0 + x_i \widehat{\beta}_1)}{\pi_i} x_i = \sum_{i=1}^{N} x_i. \tag{2}$$

which are always solvable.

Note that for some widely used sample designs (e.g. simple random sampling and stratified simple random sampling) $\widehat{F}_{Ha}(t)$ reduces to the Horvitz-Thompson estimator

$$\widehat{F}_{HT}(t) := \frac{1}{N} \sum_{i \in s} \frac{1}{\pi_i} I(y_i \leq t).$$

Also note that $\widehat{F}_{Ha}(t)$ and $\widehat{F}_{HT}(t)$ do not account for the auxiliary information unless the sample design does so through the first order inclusion probabilities π_i. When the latter is not the case, it seems more appealing to use $\widehat{F}_{cal}(t)$ instead of $\widehat{F}_{Ha}(t)$ or $\widehat{F}_{HT}(t)$.

Now, consider the plug-in estimators. The ones corresponding to $\widehat{F}_{Ha}(t)$ will be denoted by $\widehat{b}_{2,Ha}(r)$ and $\widehat{b}_{3,Ha}$, respectively, while those corresponding to $\widehat{F}_{cal}(t)$ will be denoted by $\widehat{b}_{2,cal}(r)$ and $\widehat{b}_{3,cal}$. Appendix A provides an argument by which

$$\frac{\widehat{b}_{\bullet,\diamond} - b_{\bullet}}{V_{N,\bullet,\diamond}} \xrightarrow{\mathcal{L}} N(0,1)$$

under broad conditions, where $\bullet$ is either ”2” or ”3”, $\diamond$ is either ”*Ha*” or ”*cal*” and

$V^2_{N,\bullet,\diamond}$ is the sequence of asymptotic variances to be defined below. Of course, when $\bullet = 2$, $\widehat{b}_{\bullet,\diamond}$ and $V^2_{N,\bullet,\diamond}$ depend also on r but this dependence does not show up in the notation. From formulae (A.3) and (A.4) in appendix A we know that

$$V^2_{N,\bullet,\diamond} = \frac{1}{N^2} var\left(\sum_{i\in s} d_i g(y_i)\right)$$

where

$$d_i = \begin{cases} N/(\widehat{N}\pi_i) & \text{if } \diamond = \text{"Ha"}, \\ w_i \text{ defined in (1) and (2)} & \text{if } \diamond = \text{"cal"}. \end{cases}$$

and where the function $g(t)$ is defined by either (A.2) or (A.5) in Appendix A according to whether $\bullet$ is "3" or "2".

Now, consider the case where $\diamond$ = "*Ha*". A straightforward application of the Delta-method yields

$$\sum_{i\in s} d_i g(y_i) = \sum_{i\in s} \frac{N}{\widehat{N}\pi_i} g(y_i) \simeq N\overline{g} + \sum_{i\in s} \frac{1}{\pi_i}[g(y_i) - \overline{g}]$$

where $\overline{g} := \frac{1}{N}\sum_{i=1}^N g(y_i)$. Hence we obtain

$$N^2 V^2_{N,\bullet,Ha} \simeq var\left(\sum_{i\in s} \frac{1}{\pi_i}[g(y_i) - \overline{g}]\right) = \sum_{i=1}^N \sum_{j=1}^N \frac{\pi_{ij} - \pi_i\pi_j}{\pi_i\pi_j}[g(y_i) - \overline{g}][g(y_j) - \overline{g}].$$

As estimator for $V^2_{N,\bullet,Ha}$ we may consider the standard Horvitz-Thompson variance estimator

$$\widehat{V}^2_{N,\bullet,Ha,HT} := \frac{1}{N^2} \sum_{i\in s} \sum_{j\in s} \frac{\pi_{ij} - \pi_i\pi_j}{\pi_{ij}\pi_i\pi_j}[\widehat{g}(y_i) - \overline{\widehat{g}}][\widehat{g}(y_j) - \overline{\widehat{g}}],$$

where $\widehat{g}(t)$ is an estimate for the unknown function $g(t)$ (below we propose such an estimate) and $\overline{\widehat{g}} := \sum_{i\in s} \frac{1}{\pi}\widehat{g}(y_i)/\widehat{N}$. However, it is well known that for fixed size designs the Sen-Yates-Grundy estimator

$$\widehat{V}^2_{N,\bullet,Ha,SYG} := -\frac{1}{2N^2} \sum_{i\in s} \sum_{j\in s} \frac{\pi_{ij} - \pi_i\pi_j}{\pi_{ij}} \left(\frac{\widehat{g}(y_i) - \overline{\widehat{g}}}{\pi_i} - \frac{\widehat{g}(y_j) - \overline{\widehat{g}}}{\pi_j}\right)^2$$

is usually preferable (see Vijayan, 1975). In the simulation study of this article, where only fixed size designs are considered, we used the latter.

Next consider the case where $\diamond$ = "*cal*". In this case we have (see Deville and Särndal, 1992)

$$V^2_{N,\bullet,cal} \simeq \frac{1}{N^2} var\left(\sum_{i\in s} w_i g(y_i)\right) = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N (\pi_{ij} - \pi_i\pi_j) w_i w_j E_i E_j$$

where the E_i's are the residuals of the least squares regression of the $g(y_i)$'s on the x_i's and a constant. A corresponding Horvitz-Thompson-type variance estimator is given by

$$\widehat{V}^2_{N,\bullet,cal,HT} := \frac{1}{N^2} \sum_{i \in s} \sum_{j \in s} \frac{\pi_{ij} - \pi_i \pi_j}{\pi_{ij}} w_i w_j \widehat{E}_i \widehat{E}_j,$$

where the $\widehat{E}_i$'s are the residuals of the w_i-weighted least squares regression of the sample $\widehat{g}(y_i)$'s on the sample x_i's and a constant. The corresponding Sen-Yates-Grundy estimator is given by

$$\widehat{V}^2_{N,\bullet,cal,SYG} := -\frac{1}{2N^2} \sum_{i \in s} \sum_{j \in s} \frac{\pi_{ij} - \pi_i \pi_j}{\pi_{ij}} \left(w_i \widehat{E}_i - w_j \widehat{E}_j \right)^2.$$

As pointed out in Deville and Särndal (1992), in the definition of the variance $\widehat{V}^2_{N,\bullet,cal,HT}$ and $\widehat{V}^2_{N,\bullet,cal,SYG}$ we may use the inverse inclusion probabilities $1/\pi_i$ in place of the sample weights w_i.

To complete the definitions of the variance estimators we still must define an estimator for the $g(t)$ functions in formulae (A.2) and (A.5) in Appendix A. These functions depend on ν_r, δ, $f(\nu_r)$, $F(\nu)$, b_3 and $b_2(r)$ which must be estimated from the sample. In the simulation study of this article we estimated the $g(t)$ functions by substituting plug-in estimates for ν_r, σ, F(ν), b_3 and $b_2(r)$. In order to estimate $f(\nu_r)$ we used a method which is already known from Francisco and Fuller (1986), Kovar et al. (1988) and Rao et al. (1990). The method consists in equating the length of the 95% normal confidence interval for ν_r to the length of the corresponding Woodruff confidence interval. The former is given by

$$2z_{0.025} \frac{\widehat{\sigma}_\diamond}{\widehat{f(\nu_r)}}$$

where $\widehat{\sigma}_\diamond$ is an estimate for the standard deviation of the cdf-estimator $\widehat{F}_\diamond(t)$ at the point $t = \widehat{\nu}_r$ and $\widehat{f(\nu_r)}$ is the estimate for $f(\nu_r)$ we are looking for. On the other side of the equation, the length of the Woodruff confidence interval is given by

$$\widehat{F}_\diamond^{-1}(r + z_{0.025}\widehat{\sigma}_\diamond) - \widehat{F}_\diamond^{-1}(r - z_{0.025}\widehat{\sigma}_\diamond).$$

Hence, the estimator for $f(\nu_r)$ used in the simulations is given by

$$\widehat{f(\nu_r)} = \frac{2z_{0.025}\widehat{\sigma}_\diamond}{\widehat{F}_\diamond^{-1}(r + z_{0.025}\widehat{\sigma}_\diamond) - \widehat{F}_\diamond^{-1}(r - z_{0.025}\widehat{\sigma}_\diamond)}. \tag{3}$$

Note that $\widehat{f(\nu_r)}$ is a density estimator and hence we expect some bias from it. A formal analysis of this bias lies outside of the scope of this article. However, our simulation results suggest that it goes to zero slowly since our estimates of the relative bias of the variance estimators $\widehat{V}^2_{N,\bullet,Ha,SYG}$ and $\widehat{V}^2_{N,\bullet,cal,SYG}$ (which depend $\widehat{f(\nu_r)}$) do not always decrease in absolute value as we pass from the smaller to the larger sample size.

As for $\widehat{\sigma}_{\diamond}$ which appears in the definition of $\widehat{f(\nu_r)}$, in the simulations we used the square root of the Sen-Yates-Grundy estimators

$$\widehat{\sigma}^2_{Ha} := -\frac{1}{2N^2}\sum_{i\in s}\sum_{j\in s}\frac{\pi_{ij}-\pi_i\pi_j}{\pi_{ij}}\times$$

$$\times\left(\frac{I(y_i \le \widehat{\nu}_r)-\widehat{F}_{Ha}(\widehat{\nu}_r)}{\pi_i}-\frac{I(y_j \le \widehat{\nu}_r)-\widehat{F}_{Ha}(\widehat{\nu}_r)}{\pi_j}\right)^2$$

and

$$\widehat{\sigma}^2_{cal} := -\frac{1}{2N^2}\sum_{i\in s}\sum_{j\in s}\frac{\pi_{ij}-\pi_i\pi_j}{\pi_{ij}}\left(w_iI(y_i \le \widehat{\nu}_r)-w_jI(y_j \le \widehat{\nu}_r)\right)^2.$$

One final remark with regard to variance estimation is in order. The estimators proposed in this section require knowledge of the second order inclusion probabilities π_{ij}. However, the latter are often not available for secondary analyses on large scale surveys. To obtain variance estimates in such cases, the second order inclusion probabilities are often computed from the first order inclusion probabilities according to some approximation like e.g. the Hájek approximation

$$\pi_{ij} \simeq \pi_i\pi_j\left(1-\frac{(1-\pi_i)(1-\pi_j)}{d_0}\right) \tag{4}$$

which, in case of conditional Poisson sampling, becomes more and more precise as $d_0 := \sum_{i=1}^{N}\pi_i(1-\pi_i)$ gets large (see Theorem 5.2 in Hajek 1964). In the simulation study of this paper we used this approximation for the simulated samples from the conditional Poisson designs. Tillé (2006, pages 84-86) contains an overview about recursive algorithms which provide better approximations.

3. SIMULATION STUDY

3.1 *Simulations with artificial populations*

In the first part of the simulation study we considered two artificial populations of size $N = 800$ where the X values are realizations of i.i.d. random variables with log-normal distribution with $E(\ln X) = 0$ and $var(\ln X) = 1$. Once we generated the X values, we obtained the Y values according to

$$y_i = x_i + v(x_i)\varepsilon_i, \quad i = 1, 2, \ldots, N$$

where the ε_i's are realizations of independent standard normal random variables. For $v(x_i)$ we used $v(x_i) := x_i^{\gamma}$ with γ equal to either 0 or 1. The two populations obtained in this way are shown in Figure 1. For the population obtained from the model with $\gamma = 0$ we have $b_2 = 0.040$ and $b_3 = 0.226$, while for the population from the model with $\gamma = 1$ we have $b_2 = 0.321$ and $b_3 = 0.455$.

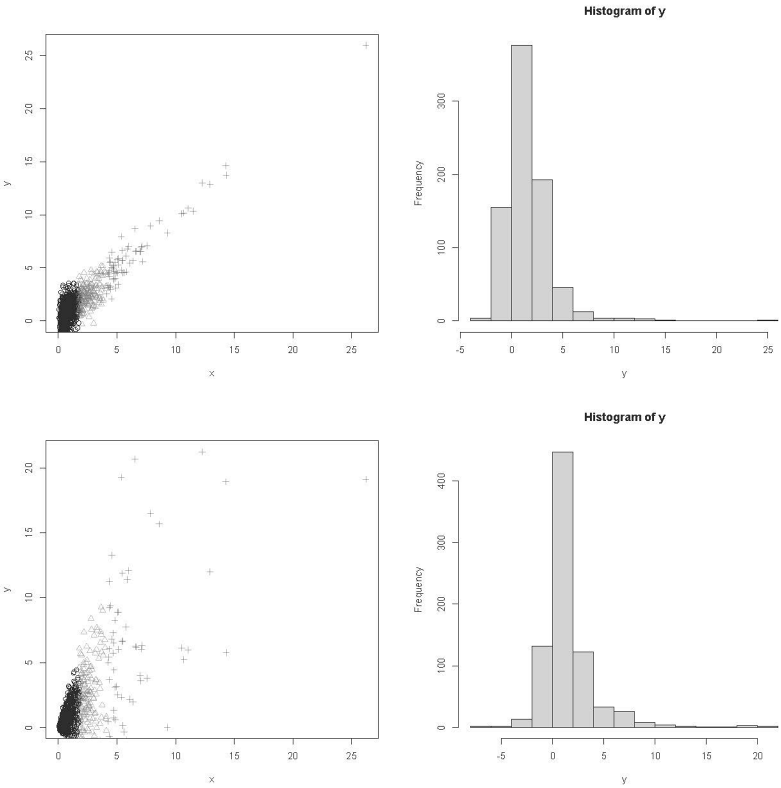

FIGURE 1. - *Population from the model with $\gamma = 0$ (upper part) and $\gamma = 1$ (lower part). Different symbols (circle, triangle, plus sign) are used to show strata membership in the stratified sampling designs*

As for the sample design, we considered three fixed size designs each with sample size $n = 40$ and $n = 80$ as to achieve sampling fractions of 5% and 10%, respectively. The three sample designs are:

i) simple random sampling without replacement;
ii) stratified simple random sampling with proportional allocation where the population is divided into three strata defined by intervals of X such that the aggregate value of X is the approximately the same in each stratum;
iii) conditional Poisson sampling with first order inclusion probabilities proportional to the x_i values.

For each considered population and sampling design we simulated a first set of

1000 independent samples in order to estimate the mse of the of the estimators $\widehat{b}_{2,Ha}$, $\widehat{b}_{2,cal}$, $\widehat{b}_{3,Ha}$ and $\widehat{b}_{3,cal}$, i.e. in order to obtain

$$\text{mse}_{\bullet,\diamond} = \frac{1}{1000}\sum_{h=1}^{1000}\left(\widehat{b}_{\bullet}^{(h)} - b_{\bullet,\diamond}\right)^2$$

where $\widehat{b}_{\bullet,\diamond}^{(h)}$ is the value taken on by $\widehat{b}_{\bullet,\diamond}$ on the hth simulated sample. The square root of this mse is reported in the "rmse" column of the tables in Appendix B. Moreover, it is also used as denominator for the computation of the bias contribution to the mse (the "bias^2/mse" column) and for the computation of the relative bias, the relative stability and the relative estimation errors of the variance estimators. The definitions of these quantities are given below. Note that for comparison the tables report also simulation results for the Hájek and calibration estimators for the mean, i.e. for the estimators

$$\bar{y}_{Ha} := \frac{1}{\widehat{N}}\sum_{i\in s}\frac{1}{\pi_i}y_i$$

and

$$\bar{y}_{cal} := \frac{1}{N}\sum_{i\in s} w_i y_i.$$

Apart from the mse, all other quantities needed for the construction of the tables in Appendix B are computed from an independent second set of 1000 simulated samples. In particular, given the realizations

$$\widehat{b}_{\bullet,\diamond}^{(h)}, \widehat{V}_{\bullet,\diamond}^{(h)}, \quad h = 1, 2, \ldots, 1000,$$

for this second set of samples, we computed, for $1-\alpha = 0.90, 0.95, 0.99$, the coverage rates

$$\text{coverage rate} := \frac{1}{1000}\sum_{h=1}^{1000} \mathrm{I}(\widehat{b}_{\bullet,\diamond} - z_{\alpha/2}\widehat{V}_{\bullet,\diamond} < b_{\bullet} < \widehat{b}_{\bullet,\diamond} + z_{\alpha/2}\widehat{V}_{\bullet,\diamond}),$$

the corresponding left and right tail errors

$$\text{left tail error LTE} := \frac{1}{1000}\sum_{h=1}^{1000} \mathrm{I}(b_{\bullet} < \widehat{b}_{\bullet,\diamond} - z_{\alpha/2}\widehat{V}_{\bullet,\diamond})$$

$$\text{right tail error RTE} := \frac{1}{1000}\sum_{h=1}^{1000} \mathrm{I}(\widehat{b}_{\bullet,\diamond} + z_{\alpha/2}\widehat{V}_{\bullet,\diamond} < b_{\bullet})$$

and

$$\text{bias} := \frac{1}{1000}\sum_{h=1}^{1000}\widehat{b}_{\bullet,\diamond}^{(h)} - b_{\bullet}.$$

Moreover, following Kovar *et al.* (1988), we computed

$$\text{rel.bias}(\widehat{V}_{\bullet,\diamond}^{2}) := \frac{\sum_{h=1}^{1000}\widehat{V}_{\bullet,\diamond}^{(h)2}/1000}{\text{mse}} - 1$$

and

$$\text{rel.stab.}(\widehat{V}_{\bullet,\diamond}^{2}) := \frac{\sqrt{\sum_{h=1}^{1000}(\widehat{V}_{\bullet,\diamond}^{(h)2} - \text{mse}_{\bullet,\diamond})^2/1000}}{\text{mse}_{\bullet,\diamond}},$$

where mse $\text{mse}_{\bullet,\diamond}$ is computed from the first set of 1000 samples. Since for the conditional Poisson sampling design, and occasionally for the other designs as well, we obtained suspiciously large values for the relative bias and/or the relative stability, we also computed some quantiles of the observed relative estimation errors

$$\text{rel.est.error}_h := \frac{\widehat{V}_{\bullet,\diamond}^{(h)2}}{\text{mse}_{\bullet,\diamond}} - 1, \quad h = 1, 2, \ldots, 1000.$$

From the tables in Appendix B we see that the coverage rates of the confidence intervals for b_2 and b_3 are usually larger than their nominal level. In fact, in most cases the LTE and the RTE are both below their nominal level. In the samples from the populations with $\gamma = 0$ we usually observe $LTE < RTE$, while in the samples from the populations with $\gamma = 1$ we often observe the opposite inequality. Also, we note little difference between the coverage accuracy of the confidence intervals based on $\widehat{b}_{\bullet,Ha}$ and those based on $\widehat{b}_{\bullet,cal}$. For comparison, the confidence intervals for the mean usually suffer from undercoverage. For the confidence intervals based on $\overline{y}_{Ha}$ we usually observe $LTE < RTE$ when the sampling scheme is either srs or stratified srs and the opposite inequality for the conditional Poisson sampling designs. The LTE and RTE of the confidence intervals based on y_{cal} are more balanced than those based on $\overline{y}_{cal}$.

Next, we comment on the central and lower parts of the tables. In the central parts we see that the rmse's of the calibration estimators $\widehat{b}_{\bullet,cal}$ are usually close to those of the Hájek estimators $\widehat{b}_{\bullet,Ha}$ suggesting that $\widehat{F}_{cal}(t)$ takes little advantage of the auxiliary information. By contrast, the rmse of $\overline{y}_{cal}$ is (as expected) constantly smaller than the rmse of $\overline{y}_{Ha}$ with at times large observed differences for the populations generated from the model with $\gamma = 0$ (i.e. when the correlation between the x_i and y_i values is stronger).

From the avg($\widehat{V}$) column we can infer the average length of the confidence intervals. Consistently with our observations about the rmse's, we generally observe little difference between the avg($\widehat{V}$) values corresponding to $\widehat{b}_{\bullet,Ha}$ and $\widehat{b}_{\bullet,cal}$, while the difference between the avg($\widehat{V}$) values corresponding to $\overline{y}_{Ha}$ and $\overline{y}_{cal}$ is often quite large especially in the cases where the population was generated from the model with $\gamma = 0$. Moreover, we note that the avg$(\widehat{V}_{2,\diamond})$ and avg$(\widehat{V}_{3,\diamond})$ values are almost constantly larger than their corresponding rmse which is in line with our observation that the coverage rates of the confidence intervals for b_2 and b_3 are larger than nominal.

Finally, we analyze the relative bias and the relative stability of the variance estimators. We immediately note that the bias of the variance estimators $\widehat{V}_{\bullet,\diamond}$ is often a considerable fraction of the mean square error of the corresponding $\widehat{b}_{\bullet,\diamond}$ estimator. In fact, for $n = 40$ the relative bias of $\widehat{V}_{\bullet,\diamond}$ is occasionally larger than 0.70 but in most cases it decreases sharply in absolute value when passing from samples of size $n = 40$ to samples of size $n = 80$. However there are some exceptions to this rule. In fact, in three out of 24 cases ($\widehat{V}_{3,cal}$ in case of srs from the population generated with $\gamma = 0$, $\widehat{V}_{2,cal}$ in case of stratified srs from the population generated with $\gamma = 0$, $\widehat{V}_{3,cal}$ in case of conditional Poisson sampling from the population generated with $\gamma = 0$) the relative bias of $\widehat{V}_{\bullet,\diamond}$ increases slightly as the sample size passes from $n = 40$ to $n = 80$ (note that all of the three exceptional cases occur when $\widehat{F}_{cal}(t)$ is used, but in the case of conditional Poisson sampling from the population generated with $\gamma = 0$, the relative bias of $\widehat{V}_{3,Ha}$ reduces only by 0.001 when passing from $n = 40$ to $n = 80$). For comparison, Kovar et al. (1988) (KRW) consider the median estimator $\widehat{F}_{Ha}^{-1}(0.5)$ and compute the relative bias and relative stability of a corresponding variance estimator (which is similar to ours) from 1000 simulated samples (see Tables 4 and 6 in KRW). The samples where obtained by stratified srs with fixed strata sample size of either $n = 2$ or $n = 5$ from two populations with 32 strata each. Stratification is based on the auxiliary variable and the second population is needed to test for the effect of strong correlation between the auxiliary and the study variable. The variance estimator they consider is defined in such way that a corresponding $1 - \alpha$ normal confidence interval for the median matches the $1 - \alpha$ Woodruff confidence interval. Thus their variance estimator incorporates the idea underlying the density estimator (3) of the present article and can therefore be considered similar to our $\widehat{V}_{\bullet,\diamond}$ estimators. KRW consider variance estimators corresponding to α =0.01, 0.025, 0.05, 0.10, 0.20 in order to test for the dependence on α. The relative biases obtained by them are in the range 0.032 to 0.11 for the samples from the population with lower correlation between the auxiliary variable and the study variable (see Table 4 in KRW). For the samples from the population with high correlation they report results only for the case where the fixed strata sample size is equal to 2, and obtain negative relative biases in the range –0.901 to –0.757. Substantially, KRW ascribe this far from zero negative relative bias to a problem with the density estimator (3) which occurs in case of stratified sampling from highly correlated populations (see page 42 in KRW). The results in Tables 4 and 6 of KRW show that the relative bias of the Rao-Wu modified bootstrap variance estimator can be kept in the range –0.083 to 0.147 by using an appropriate resampling scheme. We did not test this bootstrap estimator with the populations and sampling designs of the present article and therefore we are not able to tell whether it would give rise to smaller relative bias for the larger strata sample sizes considered in the present article (for the definition of the modified bootstrap estimator in case of unequal probability sampling see Rao and Wu, 1988).

In exactly the same three cases where the relative bias does not decrease when passing from sample size $n = 40$ to sample size $n = 80$, the relative stability does not decrease either. In most of the remaining cases the relative stability decreases quite sharply. To provide a benchmark value, we note that in case of i.i.d. sampling from a normal distribution the true relative stability of the unbiased estimator for the variance of the sample mean is $\sqrt{2/(n-1)}$ which is 0.226 for $n = 40$ and 0.159 for $n = 80$. For the srs and the stratified designs we usually obtained relative stabilities between 2 and 6 times as large. Such multiples are in line with those that can be deduced from the simulation results in KRW for the several variance estimators of $\widehat{F}_{Ha}^{-1}(0.5)$ considered by them (the multipliers can be obtained from Table 4 and Table 6 in KRW using $n = 32 \times 2 = 64$ or $n = 32 \times 5 = 160$ since KRW consider stratified sampling with $L = 32$ strata and strata sample sizes of either 2 or 5). However, for the conditional Poisson designs, we also obtained much larger relative stabilities, in particular for the $\widehat{V}_{2,\diamond}^2$estimators. As can be deduced from the third part of the tables in Appendix B, these large values are due to pronounced asymmetry and heavy-tailness of the distributions of the variance estimators. In most cases (but not all), the large relative stability can be ascribed to the largest 10% of the simulated $\widehat{V}_{2,\diamond}^2$ values.

3.2 *Simulations with SHIW sample as population*

To further test the variance estimators $\widehat{V}_{\bullet,\diamond}^2$ and the corresponding confidence intervals, we also considered the 2022 sample of the Survey on Household Income and Wealth (SHIW) conducted by the Bank of Italy as population from which to simulate samples. The microdata of the 2022 wave and of all other waves of the SHIW can be freely downloaded from the Bank of Italy's official website. For details about the SHIW we refer to Banca d'Italia (2024) and references therein. The 2022 sample contains detailed information about $N = 9641$ Italian households. In particular, it contains information about household income and wealth. In our simulations we considered the former as study variable Y and the latter as auxiliary variable X. The population mean of the study variable Y and the corresponding b_2 and b_3 indices are given by $Y = 55{,}875.22$ euro, $b_2 = 0.211$ and $b_3 = 0.551$, respectively. In our simulations we considered two types of fixed size designs, each one with total total sample size $n = 200$ and $n = 400$:

i) Stratified simple random sampling with proportional allocation. As stratification variable we used the residence region variable "IREG" (we united the strata referring to the regions "Piemonte" and "Valle d'Aosta"). Table 1 shows the total number of households N_h for each of the 19 strata and the strata sample sizes n_h corresponding to the two considered total sample sizes.
ii) Two stage stratified cluster sampling. We used the stratification variable "IREG" to obtain 19 strata and we added an additional stratum which contains all households which reside in municipalities with more than 500k inhabitants (this information can be deduced from the "ACOM4C" variable). Then we randomly allo-

cated the households in each of the first 19 strata to artificial PSUs by generating for each household a random integer between zero and the largest integer which does not exceed 0.01 times the number of households in the stratum. To allocate the households in the 20th stratum to artificial PSUs, we generated for each household a dichotomous variable and allocated households from the same region "IREG" in one of two PSUs according to the outcome of the dichotomous variable. We resorted two random allocation of households to artificial PSUs because the SHIW dataset does not provide information about PSU membership. Table 2 shows the number of randomly generated PSUs and the number of households for each stratum. Given this population structure, we implemented a two stage cluster design by selecting a total of 40 artificial PSUs from the first 19 strata according to stratified simple random sampling with proportional allocation (the number of artificial PSUs to be selected from each stratum is proportional to the number of artificial PSUs within the stratum) and by selecting all of the 12 artificial PSUs from the 20th stratum (the stratum with large PSUs). Thus, after the first stage of sampling we end up with a total of $n_{PSU} = 40 + 12 = 52$ artificial PSUs. In order to get the final sample of households, we then sampled from each one of the selected PSUs by simple random sampling. In this second stage of sampling, we allocated the final sample size (either $n = 200$ or $n = 400$) in proportion to the number of households in the $n_{PSU} = 52$ PSUs from the first stage.

In the simulations according to sampling plan i) we computed all estimates by using the actual first and second order inclusion probabilities. On the other hand, with sampling plan ii) we tested the case where secondary data provides only first order inclusion probabilities only for the sample households and the remaining first order inclusion probabilities as well as the second order inclusion probabilities are unknown and cannot be recovered from the sample because PSU membership is not provided. Hence, in the simulations with sampling plan ii) we computed the second order inclusion probabilities (which are needed for variance estimation) according to the Hájek approximation (4) with the sample estimate $\widehat{d_0} := \sum_{i \in s}(1 - \pi_i)$ in place of $d_0 := \sum_{i=1}^{N} \pi_i(1 - \pi_i)$. Unfortunately, we cannot do the same with the entire SHIW sample because the latter does not provide first order inclusion probabilities, not even for the sample households, but it provides only sample weights from a calibration process. In fact, according to the accompanying documentation, the recommended method for variance estimation from the SHIW sample is the jackknife based on replicates computed from specifically provided weights. In the present article we did not test variance estimation based on replication methods and therefore we are not able to tell how the corresponding estimators compare to the Taylor-linearization estimators $\widehat{V}^2_{\bullet,\diamond}$.

TABLE 1. - *Strata sizes N_h and strata sample sizes n_h for the stratified srs sampling designs i)*

Region (stratum)	N_h	$n_h^{(1)}$	$n_h^{(2)}$
Piemonte e Valle d'Aosta	657	14	27
Lombardia	937	19	39
Trentino	390	8	16
Veneto	625	13	26
Friuli	198	4	8
Liguria	234	5	10
Emilia Romagna	727	15	30
Toscana	433	9	18
Umbria	289	6	12
Marche	281	6	12
Lazio	508	11	21
Abruzzo	273	6	11
Molise	139	3	6
Campania	963	20	40
Puglia	914	19	38
Basilicata	136	3	6
Calabria	266	6	11
Sicilia	1126	23	47
Sardegna	545	10	22
Total	9641	200	400

The simulation results corresponding to the sampling designs i) and ii) are reported in Tables B13-B16. We note that the coverage rates of the confidence intervals for b_2 and b_3 are often larger than the corresponding nominal confidence levels, while the confidence intervals for the mean suffer from undercoverage. Perhaps surprisingly, we also note that with sampling plan i) the relative bias of the variance estimators $\widehat{V}^2_{\bullet,\diamond}$ is farther away from zero and often a much larger fraction of relative stability than with sampling plan ii), even though with sampling plan ii) we did not use the actual second order inclusion probabilities. In this respect it is also worth noting that the relative bias with sampling plan i) is always positive, while with sampling plan ii) it is sometimes negative as well.

TABLE 2. - *Number of PSU's and number of households per stratum for the stratified two stage cluster design in ii)*

Region (stratum)	Number of PSU's	Number of households
Piemonte e Valle d'Aosta	6	492
Lombardia	7	597
Trentino	5	390
Veneto	7	625
Friuli	3	198
Liguria	2	112
Emilia Romagna	8	727
Toscana	5	433
Umbria	4	289
Marche	4	281
Lazio	3	205
Abruzzo	4	273
Molise	2	139
Campania	8	735
Puglia	10	914
Basilicata	2	136
Calabria	4	266
Sicilia	8	715
Sardegna	6	545
large PSU's from all regions	12	1569
Total	110	9641

4. CONCLUSIONS

In this article we explored finite population inference for Bowley's skewness measure b_2 and for the Groeneveld-Meeden index b_3. In particular we considered plug-in estimators based on the Hàjek cdf-estimator and on a Deville-Särndal type calibration estimator. We employed a heuristic argument to derive asymptotic variance formulae and proposed corresponding estimators.

We tested the estimators and corresponding normal confidence intervals in a simulation study. In the simulations we usually obtained larger than nominal coverage rates for the confidence intervals. Moreover we observed little difference in terms of coverage rate, bias and rmse between the plug-in estimators based on the Hájek cdf-estimator and the plug-in estimators based on the calibration cdf- estimator. This result is likely due to the fact that the calibration cdf-estimator does not exploit the auxiliary information in efficient way.

As for the variance estimators, we observed large bias relative to the empirical mse's of the estimators for b_2 and b_3 which is likely due to the density estimator on which the variance estimators depend. However, for srs and stratified srs the relative stability behaves quite reasonably and in line with the relative stability of the variance estimators for the median estimator considered by Kovar et al. (1988). On the other hand, for the conditional Poisson designs we also obtained very large relative stabilities, especially for the variance estimators of the two b_2 estimators. Our simulation results show that in some cases, but not all, large relative stabilities can be ascribed to only 10% of the largest simulated variance estimates.

We conclude this paper with some suggestions. First, we suggest to explore the use of a variance stabilizing transformation as in Staudte (2014) which may lead to more efficient confidence intervals and/or to a more efficient variance estimator. Second, we suggest to test the bootstrap methods in Rao and Wu (1988) and to compare the results with those of this article. Finally, we observe that there are several alternative cdf-estimators which, under linear superpopulation models, use auxiliary information in more efficient way than the calibration estimator considered in this article (see e.g. Chambers and Dunstan, 1986; Rao et al., 1990 or Rueda, Martínez, Martínez and Arcos, 2007). Alternative cdf-estimators may therefore give rise to more efficient plug-in estimators for b_2 and/or b_3. However, plug-in estimates obtained from different cdf-estimates may not be consistent with one another.

ACKNOWLEDGEMENTS

The author wishes to thank the editor and both anonymous referees for the swift review process. He also is indebted to both referees for several helpful comments and suggestions.

APPENDIX A

Asymptotic normality of $\widehat{b}_3$. For $\lambda \in [0, 1]$ define $F_\lambda(t) := F(t) + \lambda[\widehat{F}(t) - F(t)]$, $\mu_\lambda := \int t dF_\lambda(t)$, $\nu_\lambda := F_\lambda^{-1}(0.5)$

$$\delta_\lambda := \int |t - \nu_\lambda| dF_\lambda(t)$$

$$= \int_{-\infty}^{\nu_\lambda} (\nu_\lambda - t) dF_\lambda(t) + \int_{\nu_\lambda}^{\infty} (t - \nu_\lambda) dF_\lambda(t)$$

$$= \int_{-\infty}^{\nu_\lambda} F_\lambda(t) dt + \int_{\nu_\lambda}^{\infty} [1 - F_\lambda(t)] dt$$

and $b_{3;\lambda} := \frac{\mu_\lambda - \nu_\lambda}{\delta_\lambda}$. For the moment we do not impose any restriction on the distribution functions $F(t)$ and $\widehat{F}(t)$, but later we will assume that $F(t)$ is the limit of a sequence $\{F_N(t)\}_{N=1}^{\infty}$ of population distribution functions of the study variable Y and that $\widehat{F}(t)$ is either $\widehat{F}_{Ha}(t)$ or $\widehat{F}_{cal}(t)$.

The Gateaux derivative of b_3 in direction $\widehat{F}(t) - F(t)$ is given by

$$\frac{\partial b_{3;\lambda}}{\partial \lambda} = \frac{1}{\delta_\lambda} \left(\frac{\partial \mu_\lambda}{\partial \lambda} - \frac{\partial \nu_\lambda}{\partial \lambda} \right) - \frac{\mu_\lambda - \nu_\lambda}{\delta_\lambda^2} \frac{\partial \delta_\lambda}{\partial \lambda}$$

provided that the derivatives on the rhs exist. Of course

$$\frac{\partial \mu_\lambda}{\partial \lambda} = \int t d\widehat{F}(t) - \int t dF(t) := \widehat{\mu} - \mu$$

always exists. However, $\frac{\partial \nu_\lambda}{\partial \lambda}$ and $\frac{\partial \delta_\lambda}{\partial \lambda}$ do not necessarily exist unless we impose some condition. One such condition is

A) $F(t)$ and $\widehat{F}(t)$ are absolutely continuous and differentiable with strictly positive derivatives $f(t)$ and $\widehat{f}(t)$, respectively.

If condition A holds we obtain

$$\frac{\partial \nu_\lambda}{\partial \lambda} = -\frac{\widehat{F}(\nu_\lambda) - F(\nu_\lambda)}{f_\lambda(\nu_\lambda)},$$

where $f_\lambda(t) := f(t) + \lambda[\widehat{f}(t) - f(t)]$ is the derivative of $F_\lambda(t)$ (which must be continuous and strictly positive by condition A), and

$$\frac{\partial \delta_\lambda}{\partial \lambda} = [2F_\lambda(\nu_\lambda) - 1] \frac{\partial \nu_\lambda}{\partial \lambda} + \int_{-\infty}^{\nu_\lambda} (\widehat{F}(t) - F(t)) dt - \int_{\nu_\lambda}^{\infty} (\widehat{F}(t) - F(t)) dt$$

$$= [2F_\lambda(\nu_\lambda) - 1]\frac{\partial \nu_\lambda}{\partial \lambda} + 2\nu_\lambda[\widehat{F}(\nu_\lambda) - F(\nu_\lambda)] +$$

$$- \int_{-\infty}^{\nu_\lambda} td[\widehat{F}(t) - F(t)] + \int_{\nu_\lambda}^{\infty} td[\widehat{F}(t) - F(t)].$$

Thus, using the notation $\widehat{b_3} := b_{3;\lambda=1}$ and $b_3 := b_{3;\lambda=0}$, we may write

$$\begin{aligned} \widehat{b_3} &= b_3 + \left.\frac{\partial b_{3;\lambda}}{\partial \lambda}\right|_{\lambda=0} + \int_0^1 \frac{\partial^2 b_{3;\lambda}}{\partial \lambda^2}(1-\lambda)d\lambda \\ &= b_3 + \int g(t)d[\widehat{F}(t) - F(t)] + R, \end{aligned} \tag{A.1}$$

where

$$g(t) := \frac{1}{\delta}\left\{t(1-b_3) + I(t \leq \nu)\left(\frac{1}{f(\nu)} - 2\nu b_3 + 2tb_3\right)\right\}, \quad t \in \mathbb{R}, \tag{A.2}$$

$\frac{\partial^2 b_{3;\lambda}}{\partial \lambda^2}$ is a quadratic function of $\widehat{F}(\nu_\lambda) - F(\nu_\lambda)$ and

$$R := \int_0^1 \frac{\partial^2 b_{3;\lambda}}{\partial \lambda^2}(1-\lambda)d\lambda.$$

As anticipated above, from now on we assume that $F(t) = \lim_{N\to\infty} F_N(t)$ and that $\widehat{F}(t)$ is either $\widehat{F}_{Ha}(t)$ or $\widehat{F}_{cal}(t)$. Of course, assumption A does not hold in this case, but nevertheless we assume that the following conditions, which resemble assumption A and (A.1), hold:

B1) $N, n \to \infty$, $F_N(t) \to F(t) := \int_{-\infty}^{t} f(y)dy$ where $f(t)$ is continuous and strictly positive in a neighborhood of $F^{-1}(0.5)$

B2)

$$\widehat{b_3} = b_3 + \int g(t)d[\widehat{F}(t) - F(t)] + o_p(n^{-1/2}).$$

In addition to B1 and B2 we also assume that

B3) $V_N^2 := var(\int g(t)d\widehat{F}(t)) = O(n^{-1})$ and $\frac{1}{V_N}\int g(t)d[\widehat{F}(t) - F(t)] \xrightarrow{\mathcal{L}} N(0,1).$

It is immediately seen that B1 - B3 imply

$$\frac{\widehat{b}_3 - b_3}{V_N} \xrightarrow{\mathcal{L}} N(0,1)$$

which justifies the use of normal confidence intervals. It is also worth noting that

$$V_N^2 := var\left(\int g(t)d\widehat{F}(t)\right) = \frac{1}{N^2} var\left(\sum_{i\in s} d_i g(y_i)\right) \tag{A.3}$$

where

$$d_i = \begin{cases} N/(\widehat{N}\pi_i) & \text{if } \widehat{F}(t) = \widehat{F}_{Ha}(t), \\ w_i \text{ defined in (1) and (2)} & \text{if } \widehat{F}(t) = \widehat{F}_{cal}(t), \end{cases} \tag{A.4}$$

Note that assumption B1 is certainly satisfied when the (x_i, y_i) couples are realizations of i.i.d. random vectors with a continuous and strictly positive joint density function. Of course this is by far not the only case where assumption B1 holds. As for assumptions B2 and B3, in this work we do not investigate sufficient conditions under which they hold. However, we conjecture that in a broad range of situations of practical interest it should be possible to prove B1 - B3 perhaps by using results from Conti and Marella (2015), Han and Wellner (2021) and/or Dey and Chaudhuri (2024). The simulation results in Appendix B support this claim.

Of course, the above reasoning can be applied to $b_2(r)$ as well. For $b_2(r)$ we get

$$g_2(t) := \frac{1}{\nu_{1-r} - \nu_r}\left[I(t \le \nu_{1-r})\frac{b_2 - 1}{f(\nu_{1-r})} - I(t \le \nu_r)\frac{1 + b_2}{f(\nu_r)} + 2I(t \le \nu)\frac{1}{f(\nu)}\right], \tag{A.5}$$

where $\nu_r := F^{-1}(r)$.

Groeneveld (1991) provides expressions for the influence functions of $b_2(r)$ and b_3. Obviously, they are closely related to our $g_2(t)$ and $g_3(t)$ functions. In fact, the relation between the influence function $IF(t; F, b_\bullet)$ and $g_\bullet(t)$ is given by

$$IF(t; F, b_\bullet) := \int g_\bullet(x) d[I(x \le t) - F(x)].$$

APPENDIX B

TABLE B1. - *Simple random sampling: $N = 800$, $\gamma = 0$, $n = 40$*

	nominal coverage			left tail error			right tail error		
confidence interval	0.90	0.95	0.99	0.05	0.025	0.005	0.05	0.025	0.005
ci for μ based on $\widehat{F}_{Ha}$	0.863	0.914	0.971	0.032	0.011	0.003	0.105	0.075	0.026
ci for μ based on $\widehat{F}_{cal}$	0.878	0.939	0.987	0.06	0.028	0.005	0.062	0.033	0.008
ci for b_2 based on $\widehat{F}_{Ha}$	0.955	0.981	0.995	0.02	0.007	0.001	0.025	0.012	0.004
ci for b_2 based on $\widehat{F}_{cal}$	0.931	0.967	0.99	0.025	0.011	0.003	0.044	0.022	0.007
ci for b_3 based on $\widehat{F}_{Ha}$	0.915	0.957	0.986	0.042	0.021	0.006	0.043	0.022	0.008
ci for b_3 based on $\widehat{F}_{cal}$	0.882	0.925	0.974	0.019	0.01	0.002	0.099	0.065	0.024

estimator	bias	rmse	bias^2/mse	avg$(\widehat{V}_{\bullet,\diamond})$	rel.bias$(\widehat{V}^2_{\bullet,\diamond})$	rel.stab.$(\widehat{V}^2_{\bullet,\diamond})$
$\bar{y}_{Ha}$	0.007	0.337	0	0.321	-0.002	0.759
$\bar{y}_{cal}$	0.004	0.156	0.001	0.157	0.026	0.248
$\widehat{b}_{2,Ha}$	-0.013	0.199	0.004	0.269	0.89	1.233
$\widehat{b}_{2,cal}$	-0.016	0.202	0.006	0.242	0.258	0.458
$\widehat{b}_{3,Ha}$	0.011	0.16	0.004	0.178	0.474	0.767
$\widehat{b}_{3,cal}$	-0.04	0.159	0.063	0.159	0.044	0.421

	percentiles of relative estimation errors of variance estimates								
estimator	min	0.05	0.10	0.25	0.50	0.75	0.90	0.95	max
$\widehat{V}^2(\bar{y}_{Ha})$	-0.795	-0.609	-0.558	-0.446	-0.215	0.143	0.618	1.379	3.75
$\widehat{V}^2(\bar{y}_{cal})$	-0.564	-0.329	-0.263	-0.14	-0.002	0.163	0.351	0.475	1.54
$\widehat{V}^2_{2,Ha}$	-0.363	0.028	0.139	0.381	0.699	1.13	1.773	2.38	6.861
$\widehat{V}^2_{2,cal}$	-0.487	-0.232	-0.125	0.086	0.372	0.733	1.155	1.457	5.456
$\widehat{V}^2_{3,Ha}$	-0.513	-0.237	-0.161	-0.014	0.191	0.46	0.764	1.007	1.738
$\widehat{V}^2_{3,cal}$	-0.709	-0.507	-0.411	-0.252	-0.037	0.275	0.593	0.893	2.082

TABLE B2. - *Simple random sampling: $N = 800$, $\gamma = 0$, $n = 80$*

	nominal coverage			left tail error			right tail error		
confidence interval	0.90	0.95	0.99	0.05	0.025	0.005	0.05	0.025	0.005
ci for μ based on $\widehat{F}_{Ha}$	0.888	0.932	0.98	0.028	0.012	0.003	0.084	0.056	0.017
ci for μ based on $\widehat{F}_{cal}$	0.898	0.942	0.98	0.06	0.038	0.009	0.042	0.02	0.011
ci for b_2 based on $\widehat{F}_{Ha}$	0.947	0.976	0.995	0.024	0.012	0.002	0.029	0.012	0.003
ci for b_2 based on $\widehat{F}_{cal}$	0.925	0.966	0.993	0.033	0.015	0.004	0.042	0.019	0.003
ci for b_3 based on $\widehat{F}_{Ha}$	0.901	0.961	0.994	0.047	0.02	0.003	0.052	0.019	0.003
ci for b_3 based on $\widehat{F}_{cal}$	0.905	0.945	0.987	0.025	0.012	0.003	0.07	0.043	0.01

estimator	bias	rmse	bias^2/mse	avg($\widehat{V}_{\bullet,\diamond}$)	rel.bias($\widehat{V}^2_{\bullet,\diamond}$)	rel.stab.($\widehat{V}^2_{\bullet,\diamond}$)
$\bar{y}_{Ha}$	0.012	0.24	0.002	0.228	-0.042	0.505
$\bar{y}_{cal}$	0.007	0.113	0.004	0.109	-0.065	0.16
$\widehat{b}_{2,Ha}$	-0.01	0.143	0.005	0.165	0.355	0.55
$\widehat{b}_{2,cal}$	-0.011	0.143	0.006	0.156	0.209	0.427
$\widehat{b}_{3,Ha}$	0.001	0.109	0	0.119	0.232	0.49
$\widehat{b}_{3,cal}$	-0.021	0.1	0.046	0.107	0.206	0.559

	percentiles of relative estimation errors of variance estimates								
estimator	min	0.05	0.10	0.25	0.50	0.75	0.90	0.95	max
$\widehat{V}^2(\bar{y}_{Ha})$	-0.69	-0.533	-0.48	-0.362	-0.177	0.053	0.581	1.236	2.023
$\widehat{V}^2(\bar{y}_{cal})$	-0.477	-0.276	-0.241	-0.165	-0.072	0.023	0.111	0.177	1.209
$\widehat{V}^2_{2,Ha}$	-0.433	-0.166	-0.093	0.059	0.285	0.567	0.889	1.106	3.86
$\widehat{V}^2_{2,cal}$	-0.586	-0.315	-0.226	-0.076	0.141	0.44	0.835	1.057	2.571
$\widehat{V}^2_{3,Ha}$	-0.437	-0.24	-0.179	-0.052	0.123	0.381	0.715	0.945	2.352
$\widehat{V}^2_{3,cal}$	-0.627	-0.421	-0.34	-0.171	0.087	0.472	0.91	1.222	3.179

TABLE B3. - *Simple random sampling: N = 800, γ = 1, n = 40*

confidence interval	nominal coverage 0.90	0.95	0.99	left tail error 0.05	0.025	0.005	right tail error 0.05	0.025	0.005
ci for μ based on $\widehat{F}_{Ha}$	0.846	0.903	0.957	0.017	0.005	0	0.137	0.092	0.043
ci for μ based on $\widehat{F}_{cal}$	0.821	0.886	0.949	0.081	0.049	0.024	0.098	0.065	0.027
ci for b_2 based on $\widehat{F}_{Ha}$	0.969	0.987	0.998	0.021	0.009	0.002	0.01	0.004	0
ci for b_2 based on $\widehat{F}_{cal}$	0.946	0.971	0.992	0.039	0.025	0.007	0.015	0.004	0.001
ci for b_3 based on $\widehat{F}_{Ha}$	0.916	0.953	0.993	0.055	0.035	0.006	0.029	0.012	0.001
ci for b_3 based on $\widehat{F}_{cal}$	0.882	0.934	0.984	0.054	0.03	0.008	0.064	0.036	0.008

estimator	bias	rmse	bias^2/mse	avg($\widehat{V}_{\bullet,\diamond}$)	rel.bias($\widehat{V}^2_{\bullet,\diamond}$)	rel.stab.($\widehat{V}^2_{\bullet,\diamond}$)
$\overline{y}_{Ha}$	-0.004	0.403	0	0.395	0.061	0.703
$\overline{y}_{cal}$	0	0.355	0	0.316	-0.146	0.55
$\widehat{b}_{2,Ha}$	-0.026	0.199	0.017	0.291	1.363	2.467
$\widehat{b}_{2,cal}$	-0.021	0.198	0.011	0.251	0.255	0.506
$\widehat{b}_{3,Ha}$	0.005	0.153	0.001	0.168	0.777	1.748
$\widehat{b}_{3,cal}$	-0.031	0.162	0.037	0.16	0.02	0.58

estimator	percentiles of relative estimation errors of variance estimates: min	0.05	0.10	0.25	0.50	0.75	0.90	0.95	max
$\widehat{V}^2(\overline{y}_{Ha})$	-0.834	-0.692	-0.635	-0.474	-0.179	0.47	1.069	1.429	2.946
$\widehat{V}^2(\overline{y}_{cal})$	-0.885	-0.667	-0.605	-0.47	-0.284	-0.012	0.387	0.924	3.546
$\widehat{V}^2_{2,Ha}$	-0.638	-0.113	0.041	0.352	0.813	1.472	3.038	4.631	17.835
$\widehat{V}^2_{2,cal}$	-0.624	-0.331	-0.178	0.041	0.422	0.992	1.899	2.821	22.411
$\widehat{V}^2_{3,Ha}$	-0.644	-0.349	-0.226	-0.063	0.191	0.512	0.825	1.081	3.181
$\widehat{V}^2_{3,cal}$	-0.81	-0.548	-0.435	-0.269	-0.053	0.214	0.518	0.762	13.41

TABLE B4. - *Simple random sampling: N = 800, γ = 1, n = 80*

	nominal coverage			left tail error			right tail error		
confidence interval	0.90	0.95	0.99	0.05	0.025	0.005	0.05	0.025	0.005
ci for μ based on $\widehat{F}_{Ha}$	0.878	0.928	0.971	0.026	0.007	0	0.096	0.065	0.029
ci for μ based on $\widehat{F}_{cal}$	0.862	0.928	0.972	0.062	0.034	0.016	0.076	0.038	0.012
ci for b_2 based on $\widehat{F}_{Ha}$	0.926	0.963	0.994	0.046	0.024	0.006	0.028	0.013	0
ci for b_2 based on $\widehat{F}_{cal}$	0.918	0.958	0.992	0.05	0.027	0.008	0.032	0.015	0
ci for b_3 based on $\widehat{F}_{Ha}$	0.893	0.938	0.986	0.069	0.043	0.01	0.038	0.019	0.004
ci for b_3 based on $\widehat{F}_{cal}$	0.9	0.944	0.981	0.047	0.03	0.011	0.053	0.026	0.008

estimator	bias	rmse	bias^2/mse	avg$(\widehat{V}_{\bullet,\diamond})$	rel.bias$(\widehat{V}^2_{\bullet,\diamond})$	rel.stab.$(\widehat{V}^2_{\bullet,\diamond})$
$\bar{y}_{Ha}$	0.005	0.291	0	0.28	-0.023	0.444
$\bar{y}_{cal}$	0.003	0.259	0	0.227	-0.195	0.458
$\widehat{b}_{2,Ha}$	-0.01	0.144	0.005	0.165	0.37	0.773
$\widehat{b}_{2,cal}$	-0.002	0.139	0	0.151	0.045	0.308
$\widehat{b}_{3,Ha}$	0.004	0.11	0.002	0.111	0.223	0.543
$\widehat{b}_{3,cal}$	-0.013	0.113	0.013	0.109	-0.041	0.312

	percentiles of relative estimation errors of variance estimates								
estimator	min	0.05	0.10	0.25	0.50	0.75	0.90	0.95	max
$\widehat{V}^2(\bar{y}_{Ha})$	-0.817	-0.602	-0.537	-0.373	-0.096	0.257	0.602	0.81	1.714
$\widehat{V}^2(\bar{y}_{cal})$	-0.764	-0.588	-0.545	-0.448	-0.286	-0.068	0.227	0.498	5.06
$\widehat{V}^2_{2,Ha}$	-0.659	-0.289	-0.205	-0.039	0.218	0.573	1.051	1.489	6.21
$\widehat{V}^2_{2,cal}$	-0.702	-0.383	-0.268	-0.098	0.115	0.431	0.835	1.137	3.89
$\widehat{V}^2_{3,Ha}$	-0.726	-0.386	-0.303	-0.18	0.021	0.223	0.435	0.561	2.009
$\widehat{V}^2_{3,cal}$	-0.815	-0.458	-0.389	-0.258	-0.078	0.131	0.328	0.49	2.684

TABLE B5. - *Stratified simple random sampling: $N = 800$, $\gamma = 0$, $n = 40$*

confidence interval	nominal coverage 0.90	0.95	0.99	left tail error 0.05	0.025	0.005	right tail error 0.05	0.025	0.005
ci for μ based on $\widehat{F}_{Ha}$	0.873	0.932	0.984	0.04	0.026	0.003	0.087	0.042	0.013
ci for μ based on $\widehat{F}_{cal}$	0.895	0.938	0.98	0.058	0.037	0.014	0.047	0.025	0.006
ci for b_2 based on $\widehat{F}_{Ha}$	0.943	0.971	0.996	0.021	0.008	0.001	0.036	0.021	0.003
ci for b_2 based on $\widehat{F}_{cal}$	0.931	0.968	0.996	0.027	0.012	0.002	0.042	0.02	0.002
ci for b_3 based on $\widehat{F}_{Ha}$	0.924	0.965	0.989	0.021	0.013	0.002	0.055	0.022	0.009
ci for b_3 based on $\widehat{F}_{cal}$	0.918	0.956	0.987	0.018	0.009	0.002	0.064	0.035	0.011

estimator	bias	rmse	bias^2/mse	avg$(\widehat{V}_{\bullet,\diamond})$	rel.bias$(\widehat{V}^2_{\bullet,\diamond})$	rel.stab.$(\widehat{V}^2_{\bullet,\diamond})$
$\bar{y}_{Ha}$	-0.004	0.226	0	0.21	-0.056	0.722
$\bar{y}_{cal}$	0.002	0.163	0	0.158	-0.052	0.207
$\widehat{b}_{2,Ha}$	-0.016	0.2	0.006	0.237	0.44	0.658
$\widehat{b}_{2,cal}$	-0.014	0.201	0.005	0.239	0.367	0.625
$\widehat{b}_{3,Ha}$	-0.01	0.139	0.006	0.16	0.46	0.705
$\widehat{b}_{3,cal}$	-0.026	0.132	0.04	0.151	0.369	0.66

estimator	percentiles of relative estimation errors of variance estimates: min	0.05	0.10	0.25	0.50	0.75	0.90	0.95	max
$\widehat{V}^2(\bar{y}_{Ha})$	-0.7	-0.543	-0.492	-0.394	-0.257	-0.026	0.297	2.109	3.413
$\widehat{V}^2(\bar{y}_{cal})$	-0.534	-0.389	-0.312	-0.188	-0.053	0.081	0.205	0.28	0.64
$\widehat{V}^2_{2,Ha}$	-0.52	-0.21	-0.122	0.08	0.37	0.699	1.111	1.346	2.563
$\widehat{V}^2_{2,cal}$	-0.607	-0.24	-0.124	0.082	0.38	0.741	1.115	1.439	4.188
$\widehat{V}^2_{3,Ha}$	-0.467	-0.281	-0.204	-0.008	0.288	0.643	1.04	1.299	2.578
$\widehat{V}^2_{3,cal}$	-0.626	-0.351	-0.236	-0.04	0.29	0.667	1.071	1.387	3.183

TABLE B6. - *Stratified simple random sampling: N = 800, γ = 0, n = 80*

confidence interval	nominal coverage 0.90	0.95	0.99	left tail error 0.05	0.025	0.005	right tail error 0.05	0.025	0.005
ci for μ based on $\widehat{F}_{Ha}$	0.888	0.944	0.985	0.036	0.015	0.002	0.076	0.041	0.013
ci for μ based on $\widehat{F}_{cal}$	0.902	0.962	0.994	0.049	0.021	0.003	0.049	0.017	0.003
ci for b_2 based on $\widehat{F}_{Ha}$	0.939	0.964	0.991	0.029	0.015	0.002	0.032	0.021	0.007
ci for b_2 based on $\widehat{F}_{cal}$	0.918	0.957	0.991	0.032	0.016	0.003	0.05	0.027	0.006
ci for b_3 based on $\widehat{F}_{Ha}$	0.927	0.964	0.992	0.033	0.012	0.002	0.04	0.024	0.006
ci for b_3 based on $\widehat{F}_{cal}$	0.918	0.96	0.989	0.023	0.007	0.002	0.059	0.033	0.009

estimator	bias	rmse	bias^2/mse	avg($\widehat{V}_{\bullet,\diamond}$)	rel.bias($\widehat{V}^2_{\bullet,\diamond}$)	rel.stab.($\widehat{V}^2_{\bullet,\diamond}$)
$\bar{y}_{Ha}$	0.006	0.153	0.002	0.154	0.089	0.69
$\bar{y}_{cal}$	0.001	0.111	0	0.109	-0.026	0.146
$\widehat{b}_{2,Ha}$	-0.013	0.138	0.009	0.156	0.321	0.556
$\widehat{b}_{2,cal}$	-0.013	0.14	0.008	0.156	0.411	0.662
$\widehat{b}_{3,Ha}$	0.002	0.094	0.001	0.11	0.267	0.492
$\widehat{b}_{3,cal}$	-0.019	0.09	0.044	0.101	0.316	0.617

estimator	percentiles of relative estimation errors of variance estimates: min	0.05	0.10	0.25	0.50	0.75	0.90	0.95	max
$\widehat{V}^2(\bar{y}_{Ha})$	-0.637	-0.417	-0.384	-0.296	-0.14	0.06	1.662	1.843	2.283
$\widehat{V}^2(\bar{y}_{cal})$	-0.445	-0.248	-0.205	-0.129	-0.034	0.059	0.173	0.221	0.476
$\widehat{V}^2_{2,Ha}$	-0.646	-0.266	-0.203	-0.009	0.242	0.568	0.923	1.176	2.418
$\widehat{V}^2_{2,cal}$	-0.47	-0.254	-0.186	-0.044	0.196	0.493	0.794	1.016	2.736
$\widehat{V}^2_{3,Ha}$	-0.527	-0.262	-0.178	0.004	0.336	0.732	1.115	1.38	2.375
$\widehat{V}^2_{3,cal}$	-0.528	-0.358	-0.278	-0.073	0.223	0.601	1.052	1.337	2.633

TABLE B7. - *Stratified simple random sampling: $N = 800$, $\gamma = 1$, $n = 40$*

	nominal coverage			left tail error			right tail error		
confidence interval	0.90	0.95	0.99	0.05	0.025	0.005	0.05	0.025	0.005
ci for μ based on $\widehat{F}_{Ha}$	0.883	0.929	0.979	0.047	0.026	0.007	0.07	0.045	0.014
ci for μ based on $\widehat{F}_{cal}$	0.873	0.916	0.977	0.062	0.043	0.014	0.065	0.041	0.009
ci for b_2 based on $\widehat{F}_{Ha}$	0.951	0.977	0.994	0.031	0.015	0.006	0.018	0.008	0
ci for b_2 based on $\widehat{F}_{cal}$	0.958	0.986	0.995	0.029	0.01	0.005	0.013	0.004	0
ci for b_3 based on $\widehat{F}_{Ha}$	0.908	0.962	0.987	0.049	0.023	0.008	0.043	0.015	0.005
ci for b_3 based on $\widehat{F}_{cal}$	0.911	0.96	0.988	0.04	0.017	0.007	0.049	0.023	0.005

estimator	bias	rmse	bias^2/mse	avg$(\widehat{V}_{\bullet,\diamond})$	rel.bias$(\widehat{V}^2_{\bullet,\diamond})$	rel.stab.$(\widehat{V}^2_{\bullet,\diamond})$
$\bar{y}_{Ha}$	-0.003	0.334	0	0.329	0.023	0.49
$\bar{y}_{cal}$	0.007	0.327	0	0.32	0.016	0.519
$\widehat{b}_{2,Ha}$	-0.016	0.207	0.006	0.25	0.59	1.373
$\widehat{b}_{2,cal}$	-0.018	0.205	0.008	0.251	0.151	0.462
$\widehat{b}_{3,Ha}$	-0.004	0.149	0.001	0.157	0.633	1.428
$\widehat{b}_{3,cal}$	-0.017	0.148	0.014	0.157	0.174	0.5

	percentiles of relative estimation errors of variance estimates								
estimator	min	0.05	0.10	0.25	0.50	0.75	0.90	0.95	max
$\widehat{V}^2(\bar{y}_{Ha})$	-0.814	-0.575	-0.512	-0.356	-0.086	0.303	0.699	1.038	2.335
$\widehat{V}^2(\bar{y}_{cal})$	-0.787	-0.575	-0.515	-0.36	-0.103	0.257	0.725	1.028	2.527
$\widehat{V}^2_{2,Ha}$	-0.673	-0.359	-0.268	-0.028	0.318	0.766	1.464	2.269	13.322
$\widehat{V}^2_{2,cal}$	-0.653	-0.342	-0.253	-0.016	0.364	0.847	1.694	2.221	14.794
$\widehat{V}^2_{3,Ha}$	-0.757	-0.458	-0.34	-0.158	0.078	0.401	0.735	0.954	1.913
$\widehat{V}^2_{3,cal}$	-0.719	-0.447	-0.355	-0.152	0.102	0.409	0.802	1.023	2.337

TABLE B8. - *Stratified simple random sampling: $N = 800$, $\gamma = 1$, $n = 80$*

	nominal coverage			left tail error			right tail error		
confidence interval	0.90	0.95	0.99	0.05	0.025	0.005	0.05	0.025	0.005
ci for μ based on $\widehat{F}_{Ha}$	0.877	0.934	0.983	0.046	0.025	0.006	0.077	0.041	0.011
ci for μ based on $\widehat{F}_{cal}$	0.878	0.945	0.984	0.052	0.027	0.007	0.07	0.028	0.009
ci for b_2 based on $\widehat{F}_{Ha}$	0.928	0.966	0.992	0.044	0.023	0.007	0.028	0.011	0.001
ci for b_2 based on $\widehat{F}_{cal}$	0.934	0.965	0.994	0.039	0.022	0.005	0.027	0.013	0.001
ci for b_3 based on $\widehat{F}_{Ha}$	0.901	0.943	0.99	0.056	0.032	0.006	0.043	0.025	0.004
ci for b_3 based on $\widehat{F}_{cal}$	0.911	0.95	0.992	0.045	0.028	0.004	0.044	0.022	0.004

estimator	bias	rmse	bias^2/mse	avg($\widehat{V}_{\bullet,\diamond}$)	rel.bias($\widehat{V}^2_{\bullet,\diamond}$)	rel.stab.($\widehat{V}^2_{\bullet,\diamond}$)
$\bar{y}_{Ha}$	-0.003	0.246	0	0.239	-0.031	0.321
$\bar{y}_{cal}$	-0.005	0.239	0.001	0.227	-0.07	0.355
$\widehat{b}_{2,Ha}$	-0.015	0.143	0.011	0.153	0.196	0.555
$\widehat{b}_{2,cal}$	-0.01	0.14	0.006	0.151	-0.004	0.287
$\widehat{b}_{3,Ha}$	-0.008	0.11	0.005	0.108	0.209	0.568
$\widehat{b}_{3,cal}$	-0.015	0.107	0.018	0.108	0.023	0.296

	percentiles of relative estimation errors of variance estimates								
estimator	min	0.05	0.10	0.25	0.50	0.75	0.90	0.95	max
$\widehat{V}^2(\bar{y}_{Ha})$	-0.693	-0.495	-0.423	-0.27	-0.068	0.194	0.376	0.518	1.371
$\widehat{V}^2(\bar{y}_{cal})$	-0.687	-0.493	-0.438	-0.317	-0.127	0.085	0.39	0.616	2.369
$\widehat{V}^2_{2,Ha}$	-0.718	-0.368	-0.291	-0.133	0.104	0.414	0.783	1.033	7.351
$\widehat{V}^2_{2,cal}$	-0.679	-0.344	-0.268	-0.11	0.115	0.395	0.779	1.062	8.196
$\widehat{V}^2_{3,Ha}$	-0.788	-0.45	-0.37	-0.205	-0.013	0.184	0.387	0.49	0.945
$\widehat{V}^2_{3,cal}$	-0.802	-0.427	-0.348	-0.19	0.007	0.209	0.416	0.538	1.165

TABLE B9. - *Conditional Poisson sampling: $N = 800$, $\gamma = 0$, $n = 40$*

	nominal coverage			left tail error			right tail error		
confidence interval	0.90	0.95	0.99	0.05	0.025	0.005	0.05	0.025	0.005
ci for μ based on $\widehat{F}_{Ha}$	0.819	0.868	0.924	0.174	0.132	0.076	0.007	0	0
ci for μ based on $\widehat{F}_{cal}$	0.796	0.878	0.961	0.112	0.069	0.022	0.092	0.053	0.017
ci for b_2 based on $\widehat{F}_{Ha}$	0.893	0.943	0.983	0.047	0.027	0.011	0.06	0.03	0.006
ci for b_2 based on $\widehat{F}_{cal}$	0.905	0.954	0.99	0.046	0.026	0.006	0.049	0.02	0.004
ci for b_3 based on $\widehat{F}_{Ha}$	0.917	0.961	0.994	0.036	0.019	0.004	0.047	0.02	0.002
ci for b_3 based on $\widehat{F}_{cal}$	0.924	0.963	0.989	0.027	0.012	0.003	0.049	0.025	0.008

estimator	bias	rmse	bias^2/mse	avg($\widehat{V}_{\bullet,\diamond}$)	rel.bias($\widehat{V}^2_{\bullet,\diamond}$)	rel.stab.($\widehat{V}^2_{\bullet,\diamond}$)
$\bar{y}_{Ha}$	0.076	0.412	0.034	0.36	-0.13	0.835
$\bar{y}_{cal}$	0.009	0.264	0.001	0.219	-0.256	0.549
$\widehat{b}_{2,Ha}$	0.019	0.333	0.003	0.395	2.464	32.374
$\widehat{b}_{2,cal}$	0.008	0.338	0.001	0.35	3.639	54.873
$\widehat{b}_{3,Ha}$	0.038	0.2	0.035	0.254	0.269	4.711
$\widehat{b}_{3,cal}$	0.023	0.213	0.012	0.239	0.334	0.697

	percentiles of relative estimation errors of variance estimates								
estimator	min	0.05	0.10	0.25	0.50	0.75	0.90	0.95	max
$\widehat{V}^2(\bar{y}_{Ha})$	-0.857	-0.713	-0.678	-0.564	-0.352	0.014	0.463	1.136	7.403
$\widehat{V}^2(\bar{y}_{cal})$	-0.898	-0.731	-0.684	-0.572	-0.388	-0.088	0.304	0.685	4.097
$\widehat{V}^2_{2,Ha}$	-0.928	-0.653	-0.561	-0.363	-0.038	0.561	1.507	2.804	781.189
$\widehat{V}^2_{2,cal}$	-0.857	-0.534	-0.443	-0.28	-0.026	0.335	0.741	1.167	146.594
$\widehat{V}^2_{3,Ha}$	-0.832	-0.59	-0.49	-0.28	0.166	0.9	1.964	3.328	1182.168
$\widehat{V}^2_{3,cal}$	-0.779	-0.465	-0.356	-0.097	0.255	0.656	1.106	1.533	3.737

TABLE B10. - *Conditional Poisson sampling: $N = 800$, $\gamma = 0$, $n = 80$*

confidence interval	nominal coverage 0.90	0.95	0.99	left tail error 0.05	0.025	0.005	right tail error 0.05	0.025	0.005
ci for μ based on $\widehat{F}_{Ha}$	0.857	0.898	0.955	0.135	0.1	0.045	0.008	0.002	0
ci for μ based on $\widehat{F}_{cal}$	0.855	0.918	0.975	0.076	0.04	0.017	0.069	0.042	0.008
ci for b_2 based on $\widehat{F}_{Ha}$	0.887	0.933	0.977	0.044	0.023	0.008	0.069	0.044	0.015
ci for b_2 based on $\widehat{F}_{cal}$	0.894	0.933	0.97	0.041	0.027	0.007	0.065	0.04	0.023
ci for b_3 based on $\widehat{F}_{Ha}$	0.923	0.957	0.992	0.023	0.008	0	0.054	0.035	0.008
ci for b_3 based on $\widehat{F}_{cal}$	0.901	0.945	0.981	0.03	0.012	0	0.069	0.043	0.019

estimator	bias	rmse	bias^2/mse	avg($\widehat{V}_{\bullet,\diamond}$)	rel.bias($\widehat{V}^2_{\bullet,\diamond}$)	rel.stab.($\widehat{V}^2_{\bullet,\diamond}$)
$\overline{y}_{Ha}$	0.006	0.29	0	0.267	-0.071	0.704
$\overline{y}_{cal}$	0.003	0.185	0	0.166	-0.145	0.543
$\widehat{b}_{2,Ha}$	0.014	0.228	0.004	0.262	0.485	1.932
$\widehat{b}_{2,cal}$	0.005	0.231	0	0.251	0.717	2.693
$\widehat{b}_{3,Ha}$	0.011	0.139	0.006	0.169	0.268	1.002
$\widehat{b}_{3,cal}$	0.007	0.144	0.002	0.166	0.387	0.739

estimator	percentiles of relative estimation errors of variance estimates min	0.05	0.10	0.25	0.50	0.75	0.90	0.95	max
$\widehat{V}^2(\overline{y}_{Ha})$	-0.822	-0.654	-0.598	-0.462	-0.248	0.028	0.588	1.243	4.275
$\widehat{V}^2(\overline{y}_{cal})$	-0.818	-0.635	-0.58	-0.458	-0.266	-0.017	0.391	0.782	4.096
$\widehat{V}^2_{2,Ha}$	-0.8	-0.517	-0.389	-0.182	0.163	0.701	1.412	2.082	40.752
$\widehat{V}^2_{2,cal}$	-0.697	-0.458	-0.386	-0.199	0.072	0.493	0.997	1.475	19.978
$\widehat{V}^2_{3,Ha}$	-0.706	-0.447	-0.344	-0.083	0.277	0.884	1.575	2.184	32.56
$\widehat{V}^2_{3,cal}$	-0.738	-0.394	-0.299	-0.072	0.286	0.707	1.201	1.534	4.477

TABLE B11. - *Conditional Poisson sampling: $N = 800$, $\gamma = 1$, $n = 40$*

	nominal coverage			left tail error			right tail error		
confidence interval	0.90	0.95	0.99	0.05	0.025	0.005	0.05	0.025	0.005
ci for μ based on $\widehat{F}_{Ha}$	0.866	0.911	0.965	0.1	0.072	0.033	0.034	0.017	0.002
ci for μ based on $\widehat{F}_{cal}$	0.884	0.937	0.984	0.051	0.033	0.008	0.065	0.03	0.008
ci for b_2 based on $\widehat{F}_{Ha}$	0.953	0.974	0.989	0.016	0.011	0.008	0.031	0.015	0.003
ci for b_2 based on $\widehat{F}_{cal}$	0.964	0.98	0.996	0.013	0.008	0.002	0.023	0.012	0.002
ci for b_3 based on $\widehat{F}_{Ha}$	0.95	0.971	0.993	0.011	0.003	0	0.039	0.026	0.007
ci for b_3 based on $\widehat{F}_{cal}$	0.95	0.972	0.992	0.015	0.01	0.002	0.035	0.018	0.006

estimator	bias	rmse	bias^2/mse	avg($\widehat{V}_{\bullet,\diamond}$)	rel.bias($\widehat{V}^2_{\bullet,\diamond}$)	rel.stab.($\widehat{V}^2_{\bullet,\diamond}$)
$\bar{y}_{Ha}$	0.066	0.378	0.03	0.342	-0.147	0.432
$\bar{y}_{cal}$	0.006	0.252	0.001	0.238	-0.097	0.266
$\widehat{b}_{2,Ha}$	0.014	0.314	0.002	0.659	7.174	25.094
$\widehat{b}_{2,cal}$	0.011	0.323	0.001	0.585	5.207	71.55
$\widehat{b}_{3,Ha}$	0.011	0.162	0.004	0.247	3.538	7.831
$\widehat{b}_{3,cal}$	0.005	0.177	0.001	0.217	0.573	0.905

	percentiles of relative estimation errors of variance estimates								
estimator	min	0.05	0.10	0.25	0.50	0.75	0.90	0.95	max
$\widehat{V}^2(\bar{y}_{Ha})$	-0.726	-0.559	-0.506	-0.393	-0.232	0.004	0.287	0.506	4.87
$\widehat{V}^2(\bar{y}_{cal})$	-0.685	-0.441	-0.374	-0.267	-0.115	0.037	0.189	0.31	2.4
$\widehat{V}^2_{2,Ha}$	-1	-0.533	-0.383	-0.086	0.909	5.19	18.457	30.902	395.177
$\widehat{V}^2_{2,cal}$	-1	-0.31	-0.122	0.208	1.003	3.254	9.548	17.407	63.205
$\widehat{V}^2_{3,Ha}$	-0.783	-0.469	-0.317	-0.023	0.568	1.705	3.667	6.163	1823.474
$\widehat{V}^2_{3,cal}$	-0.622	-0.282	-0.169	0.061	0.435	0.939	1.459	1.912	4.056

TABLE B12. - *Conditional Poisson sampling: $N = 800$, $\gamma = 1$, $n = 80$*

confidence interval	nominal coverage 0.90	0.95	0.99	left tail error 0.05	0.025	0.005	right tail error 0.05	0.025	0.005
ci for μ based on $\widehat{F}_{Ha}$	0.887	0.938	0.981	0.082	0.049	0.016	0.031	0.013	0.003
ci for μ based on $\widehat{F}_{cal}$	0.898	0.954	0.99	0.048	0.021	0.003	0.054	0.025	0.007
ci for b_2 based on $\widehat{F}_{Ha}$	0.93	0.967	0.992	0.031	0.014	0.002	0.039	0.019	0.006
ci for b_2 based on $\widehat{F}_{cal}$	0.912	0.966	0.99	0.049	0.022	0.006	0.039	0.012	0.004
ci for b_3 based on $\widehat{F}_{Ha}$	0.932	0.963	0.989	0.013	0.003	0	0.055	0.034	0.011
ci for b_3 based on $\widehat{F}_{cal}$	0.921	0.963	0.991	0.038	0.015	0.001	0.041	0.022	0.008

estimator	bias	rmse	bias^2/mse	avg$(\widehat{V}_{\bullet,\diamond})$	rel.bias$(\widehat{V}^2_{\bullet,\diamond})$	rel.stab.$(\widehat{V}^2_{\bullet,\diamond})$
$\bar{y}_{Ha}$	0.007	0.247	0.001	0.24	-0.022	0.372
$\bar{y}_{cal}$	0.003	0.154	0	0.16	0.093	0.223
$\widehat{b}_{2,Ha}$	0.015	0.221	0.004	0.32	2.638	12.141
$\widehat{b}_{2,cal}$	0.021	0.225	0.008	0.253	0.874	2.952
$\widehat{b}_{3,Ha}$	0.007	0.117	0.004	0.149	0.498	2.933
$\widehat{b}_{3,cal}$	0.01	0.129	0.006	0.146	0.321	0.641

estimator	percentiles of relative estimation errors of variance estimates min	0.05	0.10	0.25	0.50	0.75	0.90	0.95	max
$\widehat{V}^2(\bar{y}_{Ha})$	-0.615	-0.428	-0.364	-0.251	-0.096	0.105	0.372	0.597	2.491
$\widehat{V}^2(\bar{y}_{cal})$	-0.369	-0.211	-0.152	-0.037	0.074	0.216	0.361	0.433	1.038
$\widehat{V}^2_{2,Ha}$	-0.788	-0.461	-0.359	-0.123	0.222	1.044	3.815	10.055	187.483
$\widehat{V}^2_{2,cal}$	-0.812	-0.432	-0.341	-0.165	0.107	0.501	1.013	1.589	55.124
$\widehat{V}^2_{3,Ha}$	-0.671	-0.425	-0.298	-0.025	0.38	1.088	1.967	2.612	58.781
$\widehat{V}^2_{3,cal}$	-0.619	-0.362	-0.27	-0.083	0.208	0.596	1.069	1.312	4.289

TABLE B13. - *Stratified simple random sampling: $N = 9641$, $\gamma = SHIW$, $n = 200$*

	nominal coverage			left tail error			right tail error		
confidence interval	0.90	0.95	0.99	0.05	0.025	0.005	0.05	0.025	0.005
ci for μ based on $\widehat{F}_{Ha}$	0.874	0.913	0.956	0.002	0	0	0.124	0.087	0.044
ci for μ based on $\widehat{F}_{cal}$	0.83	0.884	0.948	0.134	0.09	0.043	0.036	0.026	0.009
ci for b_2 based on $\widehat{F}_{Ha}$	0.957	0.983	1	0.014	0.007	0	0.029	0.01	0
ci for b_2 based on $\widehat{F}_{cal}$	0.956	0.984	0.999	0.023	0.01	0.001	0.021	0.006	0
ci for b_3 based on $\widehat{F}_{Ha}$	0.885	0.933	0.985	0.024	0.015	0.003	0.091	0.052	0.012
ci for b_3 based on $\widehat{F}_{cal}$	0.916	0.956	0.981	0.043	0.028	0.015	0.041	0.016	0.004

estimator	bias	rmse	bias^2/mse	avg$(\widehat{V}_{\bullet,\diamond})$	rel.bias$(\widehat{V}^2_{\bullet,\diamond})$	rel.stab.$(\widehat{V}^2_{\bullet,\diamond})$
$\bar{y}_{Ha}$	1.557	9.848	0.025	8.656	0.518	4.38
$\bar{y}_{cal}$	4.232	8.435	0.252	6.013	-0.225	2.114
$\widehat{b}_{2,Ha}$	-0.024	0.094	0.065	0.111	0.4	0.538
$\widehat{b}_{2,cal}$	-0.01	0.092	0.011	0.11	0.243	0.718
$\widehat{b}_{3,Ha}$	-0.03	0.099	0.089	0.108	0.442	0.6
$\widehat{b}_{3,cal}$	-0.008	0.082	0.01	0.093	0.376	0.951

	percentiles of relative estimation errors of variance estimates								
estimator	min	0.05	0.10	0.25	0.50	0.75	0.90	0.95	max
$\widehat{V}^2(\bar{y}_{Ha})$	-0.954	-0.896	-0.873	-0.799	-0.654	0.088	1.223	3.014	25.933
$\widehat{V}^2(\bar{y}_{cal})$	-0.959	-0.908	-0.889	-0.833	-0.715	-0.353	0.699	1.616	50.534
$\widehat{V}^2_{2,Ha}$	-0.366	-0.094	-0.017	0.143	0.355	0.592	0.855	1.053	2.413
$\widehat{V}^2_{2,cal}$	-0.403	-0.078	0.014	0.17	0.384	0.625	0.972	1.181	3.284
$\widehat{V}^2_{3,Ha}$	-0.552	-0.381	-0.321	-0.205	0.014	0.476	1.194	1.641	4.223
$\widehat{V}^2_{3,cal}$	-0.73	-0.365	-0.285	-0.109	0.138	0.521	1.386	1.95	8.576

TABLE B14. - *Stratified simple random sampling: $N = 9641$, $\gamma = SHIW$, $n = 400$*

	nominal coverage			left tail error			right tail error		
confidence interval	0.90	0.95	0.99	0.05	0.025	0.005	0.05	0.025	0.005
ci for μ based on $\widehat{F}_{Ha}$	0.905	0.933	0.975	0.003	0	0	0.092	0.067	0.025
ci for μ based on $\widehat{F}_{cal}$	0.854	0.904	0.968	0.101	0.071	0.025	0.045	0.025	0.007
ci for b_2 based on $\widehat{F}_{Ha}$	0.924	0.966	0.998	0.014	0.006	0	0.062	0.028	0.002
ci for b_2 based on $\widehat{F}_{cal}$	0.932	0.973	0.996	0.018	0.009	0.002	0.05	0.018	0.002
ci for b_3 based on $\widehat{F}_{Ha}$	0.88	0.935	0.981	0.01	0.002	0	0.11	0.063	0.019
ci for b_3 based on $\widehat{F}_{cal}$	0.908	0.948	0.991	0.025	0.012	0.003	0.067	0.04	0.006

estimator	bias	rmse	bias^2/mse	avg$(\widehat{V}_{\bullet,\diamond})$	rel.bias$(\widehat{V}^2_{\bullet,\diamond})$	rel.stab.$(\widehat{V}^2_{\bullet,\diamond})$
$\bar{y}_{Ha}$	1.86	7.499	0.062	6.704	0.385	2.837
$\bar{y}_{cal}$	3.082	5.746	0.288	4.521	-0.233	0.813
$\widehat{b}_{2,Ha}$	-0.032	0.067	0.222	0.075	0.272	0.392
$\widehat{b}_{2,cal}$	-0.019	0.064	0.091	0.074	0.105	0.709
$\widehat{b}_{3,Ha}$	-0.025	0.081	0.097	0.082	0.344	0.455
$\widehat{b}_{3,cal}$	-0.012	0.058	0.043	0.069	0.49	0.885

	percentiles of relative estimation errors of variance estimates								
estimator	min	0.05	0.10	0.25	0.50	0.75	0.90	0.95	max
$\widehat{V}^2(\bar{y}_{Ha})$	-0.941	-0.874	-0.851	-0.788	-0.619	-0.072	0.957	9.142	11.945
$\widehat{V}^2(\bar{y}_{cal})$	-0.919	-0.857	-0.831	-0.768	-0.592	-0.001	0.822	1.482	4.477
$\widehat{V}^2_{2,Ha}$	-0.334	-0.108	-0.048	0.071	0.237	0.41	0.645	0.752	1.868
$\widehat{V}^2_{2,cal}$	-0.245	-0.055	0.002	0.132	0.309	0.509	0.755	0.891	1.98
$\widehat{V}^2_{3,Ha}$	-0.627	-0.523	-0.462	-0.369	-0.145	0.325	1.07	1.669	3.535
$\widehat{V}^2_{3,cal}$	-0.566	-0.275	-0.196	-0.029	0.274	0.834	1.429	1.869	5.085

TABLE B15. - *Two stage stratified cluster sampling: N = 9641, γ = SHIW , n = 200*

	nominal coverage			left tail error			right tail error		
confidence interval	0.90	0.95	0.99	0.05	0.025	0.005	0.05	0.025	0.005
ci for μ based on $\widehat{F}_{Ha}$	0.782	0.833	0.914	0.003	0.001	0	0.215	0.166	0.086
ci for μ based on $\widehat{F}_{cal}$	0.76	0.825	0.912	0.144	0.102	0.053	0.096	0.073	0.035
ci for b_2 based on $\widehat{F}_{Ha}$	0.932	0.97	0.998	0.04	0.017	0.002	0.028	0.013	0
ci for b_2 based on $\widehat{F}_{cal}$	0.9	0.954	0.989	0.068	0.031	0.008	0.032	0.015	0.003
ci for b_3 based on $\widehat{F}_{Ha}$	0.88	0.927	0.988	0.019	0.012	0	0.101	0.061	0.012
ci for b_3 based on $\widehat{F}_{cal}$	0.871	0.919	0.968	0.054	0.035	0.018	0.075	0.046	0.014

estimator	bias	rmse	bias^2/mse	avg($\widehat{V}_{\bullet,\diamond}$)	rel.bias($\widehat{V}^2_{\bullet,\diamond}$)	rel.stab.($\widehat{V}^2_{\bullet,\diamond}$)
$\bar{y}_{Ha}$	-1.153	7.958	0.021	5.835	-0.015	4.377
$\bar{y}_{cal}$	2.028	5.865	0.12	4.013	-0.439	0.762
$\widehat{b}_{2,Ha}$	-0.001	0.099	0	0.115	0.362	0.521
$\widehat{b}_{2,cal}$	0.008	0.102	0.006	0.106	-0.043	0.324
$\widehat{b}_{3,Ha}$	-0.039	0.097	0.16	0.094	0.102	0.346
$\widehat{b}_{3,cal}$	-0.015	0.079	0.037	0.08	0.044	0.387

	percentiles of relative estimation errors of variance estimates								
estimator	min	0.05	0.10	0.25	0.50	0.75	0.90	0.95	max
$\widehat{V}^2(\bar{y}_{Ha})$	-0.934	-0.868	-0.844	-0.782	-0.66	-0.295	0.225	0.662	42.568
$\widehat{V}^2(\bar{y}_{cal})$	-0.913	-0.853	-0.824	-0.766	-0.653	-0.392	0.208	0.754	7.419
$\widehat{V}^2_{2,Ha}$	-0.467	-0.14	-0.057	0.103	0.301	0.583	0.854	1.044	1.927
$\widehat{V}^2_{2,cal}$	-0.666	-0.353	-0.26	-0.123	0.045	0.274	0.528	0.693	1.834
$\widehat{V}^2_{3,Ha}$	-0.66	-0.427	-0.389	-0.27	-0.095	0.103	0.344	0.533	2.132
$\widehat{V}^2_{3,cal}$	-0.699	-0.418	-0.352	-0.214	-0.02	0.213	0.51	0.78	3.282

TABLE B16. - *Two stage stratified cluster sampling: N = 9641, γ = SHIW , n = 400*

confidence interval	nominal coverage 0.90	0.95	0.99	left tail error 0.05	0.025	0.005	right tail error 0.05	0.025	0.005
ci for μ based on $\widehat{F}_{Ha}$	0.767	0.816	0.907	0.001	0.001	0	0.232	0.183	0.093
ci for μ based on $\widehat{F}_{cal}$	0.775	0.831	0.928	0.096	0.072	0.023	0.129	0.097	0.049
ci for b_2 based on $\widehat{F}_{Ha}$	0.92	0.956	0.994	0.047	0.029	0.005	0.033	0.015	0.001
ci for b_2 based on $\widehat{F}_{cal}$	0.906	0.94	0.986	0.061	0.042	0.013	0.033	0.018	0.001
ci for b_3 based on $\widehat{F}_{Ha}$	0.877	0.926	0.982	0.016	0.005	0	0.107	0.069	0.018
ci for b_3 based on $\widehat{F}_{cal}$	0.881	0.942	0.985	0.045	0.024	0.011	0.074	0.034	0.004

estimator	bias	rmse	bias^2/mse	avg$(\widehat{V}_{\bullet,\diamond})$	rel.bias$(\widehat{V}^2_{\bullet,\diamond})$	rel.stab.$(\widehat{V}^2_{\bullet,\diamond})$
$\bar{y}_{Ha}$	-1.317	5.675	0.054	4.536	0.23	3.907
$\bar{y}_{cal}$	0.713	4.844	0.022	3.049	-0.542	0.691
$\widehat{b}_{2,Ha}$	0.001	0.071	0	0.079	0.238	0.378
$\widehat{b}_{2,cal}$	0.009	0.071	0.015	0.073	-0.064	0.519
$\widehat{b}_{3,Ha}$	-0.031	0.073	0.184	0.069	0.083	0.277
$\widehat{b}_{3,cal}$	-0.01	0.061	0.029	0.058	-0.067	0.304

estimator	percentiles of relative estimation errors of variance estimates min	0.05	0.10	0.25	0.50	0.75	0.90	0.95	max
$\widehat{V}^2(\bar{y}_{Ha})$	-0.896	-0.837	-0.807	-0.75	-0.623	-0.374	0.062	0.412	26.15
$\widehat{V}^2(\bar{y}_{cal})$	-0.908	-0.858	-0.841	-0.792	-0.697	-0.438	-0.084	0.214	2.575
$\widehat{V}^2_{2,Ha}$	-0.428	-0.161	-0.093	0.022	0.192	0.415	0.629	0.762	1.536
$\widehat{V}^2_{2,cal}$	-0.517	-0.277	-0.217	-0.111	0.042	0.232	0.446	0.582	1.365
$\widehat{V}^2_{3,Ha}$	-0.615	-0.461	-0.406	-0.313	-0.182	0.008	0.255	0.518	3.99
$\widehat{V}^2_{3,cal}$	-0.714	-0.463	-0.389	-0.268	-0.113	0.087	0.289	0.446	2.15

REFERENCES

Banca d'Italia (2024). L'indagine sui bilanci delle famiglie italiane. Metodi e fonti: Note metodologiche.

Bowley A. (1901). *Elements of Statistics*. P.S. King.

Bowley A.L. (1920). *Elements of statistics*. P.S. King & Son, Ltd.; C. Scribner's Sons, London, New York, 4th ed edition.

Conti P.L., Marella D. (2015). Inference for quantiles of a finite population: Asymptotic versus resampling results. *Scandinavian Journal of Statistics*, **42**(2), 545-561.

Deville J.-C., Särndal C.-E. (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association*, **87**(418), 376-382.

Dey A., Chaudhuri P. (2024). Quantile processes and their applications in finite populations.

Francisco C.A., Fuller W.A. (1986). Estimation of the distribution func- tion with a complex survey. *Proc. Sec. Survey Res. Methods* (pp. 37-45), Washington D.C. Amer. Statist. Assoc.

Groeneveld R.A. (1991). An influence function approach to describing the skewness of a distribution. *The American Statistician*, **45**(2), 97-102.

Groeneveld R.A., Meeden G. (1984). Measuring skewness and kurtosis. *Journal of the Royal Statistical Society. Series D (The Statistician)*, **33**(4), 391-399.

Hajek J. (1964). Asymptotic Theory of Rejective Sampling with Varying Prob- abilities from a Finite Population. *The Annals of Mathematical Statistics*, **35**(4), 1491-1523.

Han Q., Wellner J.A. (2021). Complex sampling designs: Uniform limit theorems and applications. *The Annals of Statistics*, **49**(1), 459-485.

Hinkley D.V. (1975). On power transformations to symmetry. *Biometrika*, **62**(1), 101-111.

Kovar J.G., Rao J.N.K., Wu C.F.J. (1988). Bootstrap and other methods to measure errors in survey estimates. *Canadian Journal of Statistics*, **16**(S1), 25-45.

Pearson, K. (1895). Contributions to the mathematical theory of evolution. ii. skew variation in homogeneous material. *Philosophical Transactions of the Royal Society of London. A*, **186**, 343-414.

Rao J.N.K., Wu C.F.J. (1988). Resampling inference with complex survey data. *Journal of the American Statistical Association*, **83**(401), 231-241.

Staudte R.G. (2014). Inference for quantile measures of skewness. *TEST*, **23**(4), 751-768.

Tillé Y. (2006). *Sampling algorithms*. Springer series in statistics. Springer, New York.

van Zwet W. (1964). *Convex Transformations of Random Variables*. Mathematical Centre tracts. Mathematish Centrum.

Vijayan, K. (1975). On estimating the variance in unequal probability sampling. *Journal of the American Statistical Association*, **70**(351), 713-716.