STATISTICS IN TRANSITION-new series, October 2011
Vol. 12, No. 2, pp. 309—330

ESTIMATION OF QUADRATIC FINITE POPULATION FUNCTIONS USING CALIBRATION

Dalius Pumputis[1], **Andrius Čiginas**[2]

ABSTRACT

Since the quadratic finite population functions can be expressed as totals over a synthetic population consisting of some ordered pairs of elements of the initial population, the traditional and penalized calibration technique is used to derive some calibrated estimators of the quadratic finite population functions. A linear combination of estimators discussed is considered as well. A comparison of approximate variances of the calibrated estimators is also presented. A simulation study is performed to analyze the empirical properties of the calibrated estimators of the finite population variance and covariance which appear as special cases of the quadratic functions. It is shown also how the calibrated estimators of the population covariance (variance) can be applied in regression estimation of the finite population total.

Key words: calibrated estimator; penalized calibration; auxiliary variables; approximate variance.

1. Introduction

In many statistical offices and official statistics, auxiliary information becomes more and more important at the estimation stage seeking to increase the accuracy of estimators of finite population parameters. To this end, the calibration approach is often used. The idea of the calibration technique for estimating the finite population totals is presented by Deville and Särndal (1992).

Since the population totals or means are the most popular parameters in survey practice, there exists a lot of scientific literature which deals with the estimation of these parameters using calibration methods. Some of them are discussed in the paper of Särndal (2007), where an overview of the calibration theory and its application in survey sampling are given.

[1] Lithuanian University of Educational Sciences, Vilnius, Lithuania.
E-mail: dalius.pumputis@vpu.lt
[2] Vilnius University, Vilnius, Lithuania. E-mail: andrius.ciginas@mif.vu.lt

The topic on the estimation of some quadratic finite population functions, such as the finite population variance, covariance or variance of the Horvitz-Thompson estimator (see e.g. Särndal, Swensson and Wretman, 1992, p. 43), is not often met in the literature of survey statistics. Plikusas and Pumputis (2007) introduced the calibrated estimators of the population covariance (variance), which use one weighting system defined by various calibration equations and distance measures. In the paper (Plikusas and Pumputis, 2010), the estimation of population covariance (variance) is considered using several systems of calibrated weights. The estimators, derived here, are applied to improve the regression (GREG) estimators of the finite population total. A more detailed description about that application is given in the Subsection 2.4.

Singh, Horn, Chowdhury and Yu (1999) proposed calibrated estimators of the variance of the Horvitz-Thompson estimator. Sitter and Wu (2002) extended the model calibration and pseudoempirical likelihood methods to obtain efficient estimators of quadratic finite population functions. Using a general expression of the new estimators, they also derived the corresponding model calibrated estimators of the population variance, the covariance and variance of the Horvitz-Thompson estimator, and analyzed their properties.

The structure of this paper is as follows. In the next section we derive some calibrated estimators of the quadratic population functions by employing Sitter and Wu's (2002) idea to express the quadratic population functions as the population totals and by applying Deville and Särndal's (1992) calibration method as well as the penalized calibration approach (Farrell and Singh, 2002). Subsection 2.3 provides a slightly different calibration which leads to a linear combination of the Horvitz-Thompson type estimator and calibrated estimators mentioned above. In Section 3 we first derive the approximate variances of the calibrated estimators and then we present a comparison of them. Some numerical results are presented in Section 4. Here we compare by simulation the calibrated estimators of the finite population variance and covariance which are both special cases of the quadratic functions. Section 5 is devoted to concluding remarks.

2. Estimators of quadratic functions

2.1. Deville and Särndal's calibration

Consider a finite population $\mathbf{U}=\{u_1,u_2,\ldots,u_N\}$ of N elements. Without loss of generality, we can assume $\mathbf{U}=\{1,2,\ldots,N\}$. Let $y^{(k)}: y_1^{(k)}, y_2^{(k)},\ldots, y_N^{(k)}$, $k=1,2,\ldots,J$, be J study variables defined on the population $\mathbf{U}$ and taking fixed real values. The values of all variables are known only for sampled population elements. Denote $\mathbf{y}_i=(y_i^{(1)}, y_i^{(2)},\ldots, y_i^{(J)})$.

We are interested in the estimation of the quadratic finite population function

$$T=\sum_{i=1}^{N}\sum_{j=i+1}^{N}\phi(\mathbf{y}_i,\mathbf{y}_j), \tag{1}$$

under a probability sampling design (of fixed size) with strictly positive second and fourth order inclusion probabilities. Here $\phi(\bullet,\bullet)$ is a symmetric function (a kernel of degree 2 for a *U*-statistic).

Well known finite population parameters, such as the finite population variance of $y^{(k)}$,

$$S^2_{y^{(k)}} = \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{j=i+1}^{N} \left(y_i^{(k)} - y_j^{(k)} \right)^2 ;$$

the finite population covariance between two study variables $y^{(k)}$ and $y^{(l)}$,

$$C\left(y^{(k)}, y^{(l)}\right) = \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{j=i+1}^{N} \left(y_i^{(k)} - y_j^{(k)} \right)\left(y_i^{(l)} - y_j^{(l)} \right);$$

and the variance,

$$V\left(\hat{t}_{HT,y^{(k)}}\right) = \sum_{i=1}^{N} \sum_{j=i+1}^{N} \left(\pi_i \pi_j - \pi_{ij} \right)\left(y_i^{(k)} / \pi_i - y_j^{(k)} / \pi_j \right)^2 ,$$

of the Horvitz-Thompson estimator, $\hat{t}_{HT,y^{(k)}} = \sum_{i \in \mathbf{s}} d_i y_i^{(k)}$, of the population total $t_{y^{(k)}} = \sum_{i=1}^{N} y_i^{(k)}$, are as special cases of function *T*. Here π_i and π_{ij} are the first and second order inclusion probabilities, respectively; $\mathbf{s}$, $\mathbf{s} \subset \mathbf{U}$, denotes the probability sample set drawn from the population $\mathbf{U}$; $d_i = 1/\pi_i$ is the sample design weight of the element i, $i = 1,2,...,N$.

The presented alternative expressions of the finite population variance and covariance are useful in the context of our investigation. The variance $V\left(\hat{t}_{HT,y^{(k)}}\right)$ of the Horvitz-Thompson estimator $\hat{t}_{HT,y^{(k)}}$ is given in the Yates and Grundy (1953) form.

Let us arrange all the pairs (ij), $i < j$, of indexes of population elements in a sequence and number the elements of the sequence using $m = 1,2,...,N^*$, where $N^* = N(N-1)/2$ (For more details on the procedure see Sitter and Wu (2002)).

Then function *T* can be expressed in the following way:

$$T = \sum_{m=1}^{N^*} \phi_{ym} ,$$

where $\phi_{ym} = \phi(\mathbf{y}_i, \mathbf{y}_j)$ for a pair of indices $m = (ij)$. Now function T is viewed as a population total of the variable $\phi_y : \phi_{y1}, \phi_{y2}, ..., \phi_{yN^*}$, defined on a synthetic finite population $\mathbf{U}^* = \{1, 2, \ldots, N^*\}$ of size N^*.

Thus, some calibration methods can be easily employed to derive the estimators of function *T*. But first, some elements of the sampling design in the population $\mathbf{U}^*$ should be defined. The sampling design in the population $\mathbf{U}^*$ is defined so that the corresponding sample of pairs is $\mathbf{s}^* = \{m = (ij) \mid i < j, i, j \in \mathbf{s}\}$ and it is treated as if it were drawn from population $\mathbf{U}^*$; the first order inclusion probabilities over the synthetic population $\mathbf{U}^*$ are coincident with the second order inclusion probabilities over the population $\mathbf{U}$: $\pi_m^* = \pi_{ij}$ for $m = (ij)$, where π_{ij} are assumed to be strictly positive. Then the sample design weights over the population $\mathbf{U}^*$ are equal to the inverse of second order inclusion probabilities: $d_m^* = 1/\pi_m^* = 1/\pi_{ij}$ for $m = (ij)$. Denote $d_{ij} = d_m^*$.

When sample design weights are defined and there is no auxiliary information, the quadratic finite population function (1) can be estimated using the Horvitz-Thompson type estimator:

$$\hat{T}_{HT} = \sum_{m \in \mathbf{s}^*} d_m^* \phi_{ym} = \sum_{i \in \mathbf{s}} \sum_{j > i} d_{ij} \phi(\mathbf{y}_i, \mathbf{y}_j). \tag{2}$$

As it is known, estimator (2) is unbiased but its variance is often relatively large.

The weights d_{ij} of the estimator $\hat{T}_{HT}$ can be modified using auxiliary variables and calibration methods to obtain estimators with a smaller variance. Let $x^{(k)}$ serve as an auxiliary variable for the study variable $y^{(k)}$, $k = 1, 2, ..., J$. Denote $\mathbf{x}_i = (x_i^{(1)}, x_i^{(2)}, ..., x_i^{(J)})$. Assume also that the values of all auxiliary variables are known only for sampled population elements and that the total $T_{\phi_x} = \sum_{i=1}^{N} \sum_{j=i+1}^{N} \phi(\mathbf{x}_i, \mathbf{x}_j)$ is known.

Remark 1. The simple summary statistics of the auxiliary variables (e.g. a total of $x^{(k)}$) are independent of the survey and may be taken from an outside source, such as national statistical institutes. The second-order summary statistics T_{ϕ_x} are much more complicated and they are not often considered in real surveys. Thus, the direct access to such a type of auxiliary information is not very realistic. A situation when all the values $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N$ are known (referred to as complete auxiliary information) is more realistic and useful in practice. The auxiliary variables may be taken from the previous complete surveys of the same population, various administrative registers and databases. Knowing these variables, one can easily calculate T_{ϕ_x} and use it for the construction of the calibrated estimators.

We consider here the calibrated estimators of the quadratic finite population functions of the following shape

$$\hat{T}_{cal} = \sum_{i \in \mathbf{s}} \sum_{j > i} \omega_{ij}^{(cal)} \phi(\mathbf{y}_i, \mathbf{y}_j), \quad (3)$$

where new (calibrated) weights $\omega_{ij}^{(cal)}$ are defined under the following conditions:

The weights $\omega_{ij}^{(cal)}$ satisfy some calibration equation;

The distance between the weights d_{ij} and $\omega_{ij}^{(cal)}$ is minimal according to some distance measures.

First, by applying Deville and Särndal's (1992) calibration technique, we define the calibrated estimator of quadratic function T:

$$\hat{T}_{DS} = \sum_{i \in \mathbf{s}} \sum_{j > i} \omega_{ij}^{(DS)} \phi(\mathbf{y}_i, \mathbf{y}_j), \quad (4)$$

where the weights $\omega_{ij}^{(DS)}$ minimize the distance measure

$$L(\omega, d) = \sum_{i \in \mathbf{s}} \sum_{j > i} \frac{\left(\omega_{ij}^{(DS)} - d_{ij}\right)^2}{d_{ij} q_{ij}} \quad (5)$$

and satisfy the calibration equation

$$\sum_{i \in \mathbf{s}} \sum_{j > i} \omega_{ij}^{(DS)} \phi(\mathbf{x}_i, \mathbf{x}_j) = \sum_{i=1}^{N} \sum_{j=i+1}^{N} \phi(\mathbf{x}_i, \mathbf{x}_j) = T_{\phi_x}. \quad (6)$$

Here $q_{ij}, i, j \in \mathbf{s},\ i < j$, are free additional weights. The estimators can be modified by choosing q_{ij}.

Calibration equation (6) shows that the known quadratic function T_{ϕ_x} is estimated by $\hat{T}_{DS,\phi_x} = \sum_{i \in \mathbf{s}} \sum_{j > i} \omega_{ij}^{(DS)} \phi(\mathbf{x}_i, \mathbf{x}_j)$ without error. In the case of a quite high correlation between the variables ϕ_y and ϕ_x (where ϕ_x is defined similarly as ϕ_y), it is natural to expect that the estimates of the function T are more accurate when the new weights $\omega_{ij}^{(DS)}$ are applied in (4).

The weights $\omega_{ij}^{(DS)}$ of estimator (4) are given by the following proposition that is actually a corollary which follows from the derivation of weights of a calibrated estimator of the finite population total (see Deville and Särndal, 1992).

Proposition 1. The weights $\omega_{ij}^{(DS)}, i, j \in \mathbf{s}, i < j,$ which minimize the distance measure (5) and satisfy equation (6), are defined by the equations:

$$\omega_{ij}^{(DS)} = d_{ij}\left(1 + \frac{q_{ij}\phi(\mathbf{x}_i, \mathbf{x}_j)}{\sum_{u\in \mathbf{s}}\sum_{v>u} d_{uv}q_{uv}\phi^2(\mathbf{x}_u, \mathbf{x}_v)}\left(T_{\phi_x} - \sum_{u\in \mathbf{s}}\sum_{v>u} d_{uv}\phi(\mathbf{x}_u, \mathbf{x}_v)\right)\right).$$

A number of other calibrated estimators may be derived using different distance measures and calibration equations. In the following part of this paper, we will analyze some cases.

Remark 2. By replacing the values $\phi(\mathbf{y}_i, \mathbf{y}_j)$ and $\phi(\mathbf{x}_i, \mathbf{x}_j)$ in the expression of the estimator $\hat{T}_{DS}$ with $(\pi_i\pi_j - \pi_{ij})\left(\frac{y_i^{(k)}}{\pi_i} - \frac{y_j^{(k)}}{\pi_j}\right)^2$ and $(\pi_i\pi_j - \pi_{ij})\left(\frac{x_i^{(k)}}{\pi_i} - \frac{x_j^{(k)}}{\pi_j}\right)^2$, we obtain an estimator $\hat{V}_{DS}\left(\hat{t}_{HT,y^{(k)}}\right)$ of the variance of the Horvitz-Thompson estimator $\hat{t}_{HT,y^{(k)}} = \sum_{i\in \mathbf{s}} d_i y_i^{(k)}$. Assume $\pi_i\pi_j - \pi_{ij} > 0$ and let $q_{ij} = Q_{ij} / \left(\pi_i\pi_j - \pi_{ij}\right)$, where Q_{ij} are free additional constants. Then the estimator $\hat{V}_{DS}\left(\hat{t}_{HT,y^{(k)}}\right)$ reduces to that considered by Singh, Horn, Chowdhury and Yu (1999).

2.2. Penalized calibration estimators

Let us consider the estimator of quadratic finite population function T of the same form (3) and define the weights $\omega_{ij}^{(FS)}, i, j \in \mathbf{s}, i < j,$ of it, using the same calibration equation (6), but a different distance measure

$$L_P(w, d) = \sum_{i\in \mathbf{s}}\sum_{j>i} \frac{\left(\omega_{ij}^{(FS)} - d_{ij}\right)^2}{d_{ij}q_{ij}} + \varphi^2 \sum_{i\in \mathbf{s}}\sum_{j>i} \frac{\left(\omega_{ij}^{(FS)}\right)^2}{d_{ij}q_{ij}} \tag{7}$$

the analog of which is proposed in the papers of Farrell and Singh (2002) and Singh (2003), and called a penalized one. Minimization of this distance measure subject to calibration equation (6) leads to the estimator with interesting features that can be described by the words of Farrell and Singh (2002, p. 965): "… φ is a positive quantity that reflects a penalty to be decided by the investigator based on prior knowledge, or the desire for certain levels of efficiency and bias…increasing φ results in a decrease in the mean square error of the estimator; unfortunately has the side effect of increasing the bias".

Denote by $\hat{T}_{FS}$ the new, just defined estimator. Since the function L_p is coincident with L as $\varphi = 0$, the new group of estimators

$$\hat{T}_{FS} = \sum_{i\in \mathbf{s}}\sum_{j>i} \omega_{ij}^{(FS)} \phi(\mathbf{y}_i, \mathbf{y}_j) \tag{8}$$

includes the calibrated estimators $\hat{T}_{DS}$.

Proposition 2. The weights $\omega_{ij}^{(FS)}, i, j \in \mathbf{s}, i < j,$ which minimize the distance measure (7) and satisfy the calibration equation (6), are defined by the equations:

$$\omega_{ij}^{(FS)} = \frac{d_{ij}}{1+\varphi^2}\left(1 + \frac{q_{ij}\phi(\mathbf{x}_i, \mathbf{x}_j)}{\sum_{u\in \mathbf{s}}\sum_{v>u} d_{uv} q_{uv} \phi^2(\mathbf{x}_u, \mathbf{x}_v)}\left(\left(1+\varphi^2\right)T_{\phi_x} - \sum_{u\in \mathbf{s}}\sum_{v>u} d_{uv}\phi(\mathbf{x}_u, \mathbf{x}_v)\right)\right).$$

Proof. Let us take the distance measure (7) and calibration equation (6), and define the Lagrange function

$$\Lambda = \sum_{i\in \mathbf{s}}\sum_{j>i} \frac{\left(\omega_{ij}^{(FS)} - d_{ij}\right)^2}{d_{ij}q_{ij}} + \varphi^2 \sum_{i\in \mathbf{s}}\sum_{j>i} \frac{\left(\omega_{ij}^{(FS)}\right)^2}{d_{ij}q_{ij}} - \lambda\left(\sum_{i\in \mathbf{s}}\sum_{j>i} \omega_{ij}^{(DS)}\phi(\mathbf{x}_i, \mathbf{x}_j) - T_{\phi_x}\right).$$

By solving the equations

$$\frac{\partial \Lambda}{\partial \omega_{ij}^{(FS)}} = 0,\ i, j \in \mathbf{s},\ i < j,$$

we get

$$\omega_{ij}^{(FS)} = \frac{1}{1+\varphi^2} d_{ij}\left(1 + \frac{1}{2}\lambda q_{ij}\phi(\mathbf{x}_i, \mathbf{x}_j)\right). \tag{9}$$

Then, multiplying (9) by $\phi(\mathbf{x}_i, \mathbf{x}_j)$, summing over the sample $\mathbf{s}^*$ elements and taking into account calibration equation (6), we get an expression for *λ*. Substituting this expression into (9), we get an equation for $\omega_{ij}^{(FS)}$.

One can note that penalized calibration is usually used to penalize the magnitude of the calibrated weights when a lot of calibration constraints are used and the sample is particularly unbalanced so that negative or very large weights occur after the calibration procedure (see Guggemos and Tillé, 2010). In this paper, we consider penalized calibration in the case of only one calibration equation, because we are seeking only to find out if the penalized distance measure may be more advantageous than function (5) when the resulting estimators are derived using the same calibration equation.

As it is shown below (see Subsections 3.2 and 4.2), the penalized estimator $\hat{T}_{FS}$ has the lower approximate and empirical variances as compared to that of the calibrated estimator $\hat{T}_{DS}$, but according to the results of Farrell and Singh (2002), the bias of $\hat{T}_{FS}$ becomes relatively large when the parameter φ is increasing. This is not a desirable property that could inspire for the development of an improved penalized estimator.

2.3. Linear combination of estimators

We consider here a slightly different calibration when the weights of estimator (3) are derived by calibrating the original design weights d_{ij}, multiplied by some correction factor. The estimator under consideration is

$$\hat{T}_{lin} = \sum_{i \in \mathbf{s}} \sum_{j > i} \omega_{ij}^{(lin)} \phi(\mathbf{y}_i, \mathbf{y}_j), \tag{10}$$

where the weights $\omega_{ij}^{(lin)}$ minimize the distance measure

$$L_{new}(\omega, \tilde{d}) = \sum_{i \in \mathbf{s}} \sum_{j > i} \frac{\left(\omega_{ij}^{(lin)} - \tilde{d}_{ij}\right)^2}{d_{ij} q_{ij}},$$
$$\tilde{d}_{ij} = c d_{ij},\ c = \alpha + \beta + \gamma / (1 + \varphi^2),\ \alpha + \beta + \gamma = 1, \tag{11}$$

and satisfy the new calibration equation

$$\sum_{i \in \mathbf{s}} \sum_{j > i} \omega_{ij}^{(lin)} \phi(\mathbf{x}_i, \mathbf{x}_j) = (1 - \alpha) T_{\phi_x} + \alpha \sum_{i \in \mathbf{s}} \sum_{j > i} d_{ij} \phi(\mathbf{x}_i, \mathbf{x}_j). \tag{12}$$

Note that the right side of the calibration equation (12) consists of two terms: the first one is the true value of T_{ϕ_x} multiplied by $1-\alpha$, and the second one – the estimate of T_{ϕ_x} multiplied by the coefficient α.

Proposition 3. Minimization of the distance measure (11) subject to the calibration equation (12) leads to the calibrated weights given by

$$\omega_{ij}^{(lin)} = d_{ij}\left(c + \frac{q_{ij}\phi(\mathbf{x}_i,\mathbf{x}_j)\left((1-\alpha)T_{\phi_x} + (\alpha - c)\sum_{u\in\mathbf{s}}\sum_{v>u} d_{uv}\phi(\mathbf{x}_u,\mathbf{x}_v)\right)}{\sum_{u\in\mathbf{s}}\sum_{v>u} d_{uv}q_{uv}\phi^2(\mathbf{x}_u,\mathbf{x}_v)} \right). \tag{13}$$

The proof is similar to that of Proposition 2.

By inserting the weights (13) into (10), we get the estimator

$$\hat{T}_{lin} = \alpha\hat{T}_{HT} + \beta\hat{T}_{DS} + \gamma\hat{T}_{FS},\ \alpha+\beta+\gamma=1, \tag{14}$$

which is a linear combination of $\hat{T}_{HT}$, $\hat{T}_{DS}$ and $\hat{T}_{FS}$.

In expression (14) one can see that α can be interpreted as a weight of the Horvitz-Thompson estimate which is included into the expression of $\hat{T}_{lin}$. Therefore, the absolute value of α reflects the rate of an influence of the Horvitz-Thompson estimator on the accuracy of the estimator $\hat{T}_{lin}$. A similar discussion can be provided about the coefficients β and γ. As an example of a high influence of the Horvitz-Thompson estimator can be obtained by choosing a value close to one for α and the values close to zero for the coefficients β and γ ($\alpha+\beta+\gamma=1$). Then the estimator $\hat{T}_{lin}$ is almost unbiased with a variance similar to that of the Horvitz-Thompson estimator. The variance of $\hat{T}_{lin}$ can be reduced by choosing a value of α close to zero, but then the estimator $\hat{T}_{lin}$ may be more biased.

Thus, the statistical properties of $\hat{T}_{lin}$ can be controlled through the values of coefficients α, β and γ. Consequently, the (optimal) values of α, β and γ, which minimize the mean square error of the estimator $\hat{T}_{lin}$ subject to an unbiasedness constraint, are more preferable than any set of α, β and γ.

2.4. Some aspects from a practical perspective

The main purpose of this subsection is to present some possibilities for the practical applications of the calibrated estimators of some quadratic functions, such as the finite population variance and covariance.

Note that according to the formulation of our problem, there is only one auxiliary variable available when the estimated parameter is a finite population variance and two auxiliaries are used in the case of estimation of the finite population covariance. Further, for simplicity, we denote the study and auxiliary variables, corresponding to the cases of estimation of variance and covariance, by y and a, and by y, z and a, b.

By replacing the values $\phi(\mathbf{y}_i, \mathbf{y}_j)$ and $\phi(\mathbf{x}_i, \mathbf{x}_j)$ in the expressions of the estimators $\hat{T}_{HT}, \hat{T}_{DS}, \hat{T}_{FS}$ and $\hat{T}_{lin}$ with $\frac{1}{N(N-1)}(y_i - y_j)^2$ and $\frac{1}{N(N-1)}(a_i - a_j)^2$, we get four estimators of the finite population variance S_y^2. We denote them by $\hat{S}_{HT}^2$, $\hat{S}_{DS}^2$, $\hat{S}_{FS}^2$, and $\hat{S}_{lin}^2$, respectively. Analogously, the substitution $\frac{1}{N(N-1)}(y_i - y_j)(z_i - z_j)$ and $\frac{1}{N(N-1)}(a_i - a_j)(b_i - b_j)$ leads to the four estimators of the finite population covariance $C(y, z)$: $\hat{C}_{HT}, \hat{C}_{DS}, \hat{C}_{FS}$ and $\hat{C}_{lin}$.

The estimators derived can be useful in the following common situation. Let us say, we want to estimate a population total

$$t_y = \sum_{k=1}^{N} y_k .$$

In the case of only one known auxiliary variable, say a, one can take the simple regression estimator (see e.g. Särndal, Swensson and Wretman, 1992, p. 272)

$$\hat{t}_{yr} = \frac{N}{\hat{N}} \sum_{k \in \mathbf{s}} d_k y_k + \hat{B}\left(\sum_{k=1}^{N} a_k - \frac{N}{\hat{N}} \sum_{k \in \mathbf{s}} d_k a_k \right), \tag{15}$$

where $\hat{N} = \sum_{k \in \mathbf{s}} d_k$, $\hat{B} = \frac{\sum_{k \in \mathbf{s}} d_k \left(a_k - \sum_{l \in \mathbf{s}} d_l a_l \Big/ \hat{N} \right)\left(y_k - \sum_{l \in \mathbf{s}} d_l y_l \Big/ \hat{N} \right)}{\sum_{k \in \mathbf{s}} d_k \left(a_k - \sum_{l \in \mathbf{s}} d_l a_l \Big/ \hat{N} \right)^2}$.

If the variables y and a are well correlated, then estimator (15) is much more accurate as compared to the Horvitz-Thompson estimator. For the sample designs for which $\hat{N} = N$, regression estimator (15) reduces to

$$\hat{t}_{yr} = \sum_{k\in\mathbf{s}} d_k y_k + \frac{\hat{C}(y,a)}{\hat{S}_a^2}\left(\sum_{k=1}^{N} a_k - \sum_{k\in\mathbf{s}} d_k a_k\right), \tag{16}$$

where $\hat{C}(y,a) = \frac{1}{N-1}\sum_{k\in\mathbf{s}} d_k\left(a_k - \sum_{l\in\mathbf{s}} d_l a_l \Big/ N\right)\left(y_k - \sum_{l\in\mathbf{s}} d_l y_l \Big/ N\right)$ and $\hat{S}_a^2 = \frac{1}{N-1}\sum_{k\in\mathbf{s}} d_k\left(a_k - \sum_{l\in\mathbf{s}} d_l a_l \Big/ N\right)^2$ are standard only design based estimators of the population covariance $C(y,a)$ and variance S_a^2, respectively. As it is shown in (Plikusas and Pumputis, 2010), regression estimator (16) can be improved by replacing the standard estimators $\hat{C}(y,a)$ and $\hat{S}_a^2$ with more accurate ones. Thus the calibrated estimators $\hat{C}_{DS}$, $\hat{C}_{FS}$, $\hat{C}_{lin}$ and $\hat{S}_{DS}^2$, $\hat{S}_{FS}^2$, $\hat{S}_{lin}^2$ may be suitable for this purpose assuming that there are available two additional known variables x_y and x_a which serve as the auxiliaries for the variables y and a, respectively.

Beside that application of the calibrated estimators of the finite population variance and covariance, they can be used also to improve the estimates of other finite population parameters, such as a finite population correlation coefficient

$$\rho(y,z) = C(y,z)\big/\left(S_y \cdot S_z\right)$$

which is a ratio of the covariance $C(y,z)$ and product of the standard deviations $S_y = \sqrt{S_y^2}$ and $S_z = \sqrt{S_z^2}$. The simplest way to estimate the correlation coefficient $\rho(y,z)$ is to use the Horvitz–Thompson type estimators $\hat{C}_{HT}$, $\hat{S}_{HT,y}^2$ and $\hat{S}_{HT,z}^2$ for estimating covariance $C(y,z)$ and variances S_y^2, S_z^2, respectively, and to take the ratio

$$\hat{\rho}(y,z) = \hat{C}_{HT}\Big/\sqrt{\hat{S}_{HT,y}^2 \cdot \hat{S}_{HT,z}^2} \tag{17}$$

as the estimator of the correlation coefficient. More accurate estimates may be obtained using in (17) the calibrated estimators instead of the corresponding Horvitz-Thompson estimators of the covariance $C(y,z)$ and variances S_y^2 and S_z^2.

3. Comparison of estimators

3.1. Approximate variances

For practical and theoretical purposes, it is good to have expressions of the exact or approximate variances of estimators, or even more, to know which estimator has the lowest variance. Since the estimators $\hat{T}_{DS}, \hat{T}_{FS}$ and $\hat{T}_{lin}$ are nonlinear functions of the Horvitz-Thompson estimators

$$\hat{T}_{HT} = \sum_{i \in \mathbf{s}} \sum_{j>i} d_{ij} \phi(\mathbf{y}_i, \mathbf{y}_j), \; \hat{T}_{HT,\phi_x} = \sum_{i \in \mathbf{s}} \sum_{j>i} d_{ij} \phi(\mathbf{x}_i, \mathbf{x}_j),$$

$$\hat{T}_{HT,q\phi_x^2} = \sum_{i \in \mathbf{s}} \sum_{j>i} d_{ij} q_{ij} \phi^2(\mathbf{x}_i, \mathbf{x}_j), \; \hat{T}_{HT,q\phi_x\phi_y} = \sum_{i \in \mathbf{s}} \sum_{j>i} d_{ij} q_{ij} \phi(\mathbf{x}_i, \mathbf{x}_j) \phi(\mathbf{y}_i, \mathbf{y}_j),$$

of the population totals

$$T, \; T_{\phi_x}, \; T_{q\phi_x^2} = \sum_{i=1}^{N} \sum_{j=i+1}^{N} q_{ij} \phi^2(\mathbf{x}_i, \mathbf{x}_j), \; T_{q\phi_x\phi_y} = \sum_{i=1}^{N} \sum_{j=i+1}^{N} q_{ij} \phi(\mathbf{x}_i, \mathbf{x}_j) \phi(\mathbf{y}_i, \mathbf{y}_j),$$

respectively, we will use the Taylor linearization technique to derive expressions of the approximate variances.

According to the Result 6.6.1 of (Särndal, Swensson and Wretman, 1992, p. 235), the approximate variance of $\hat{T}_{DS}$ can be written as

$$AV(\hat{T}_{DS}) = V(\hat{T}_{HT} - B\hat{T}_{HT,\phi_x}), \tag{18}$$

where $B = T_{q\phi_x\phi_y} / T_{q\phi_x^2}$.

Proposition 4. The approximate variances of calibrated estimators $\hat{T}_{FS}$ and $\hat{T}_{lin}$ can be expressed as follows

$$AV(\hat{T}_{FS}) = \frac{1}{(1+\varphi^2)^2} AV(\hat{T}_{DS}), \tag{19}$$

$$AV(\hat{T}_{lin}) = \left(\frac{1 - \alpha + \beta\varphi^2}{1+\varphi^2} \right)^2 AV(\hat{T}_{DS}) + \alpha^2 V(\hat{T}_{HT})$$
$$+ 2\alpha \frac{1 - \alpha + \beta\varphi^2}{1+\varphi^2} C(\hat{T}_{HT} - B\hat{T}_{HT,\phi_x}, \hat{T}_{HT}).$$

Proof. By substituting the weights $\omega_{ij}^{(FS)}$ into (8), we obtain

$$\hat{T}_{FS} = \hat{T}_{HT} \big/ \left(1+\varphi^2\right) + \left(T_{\phi_x} - \hat{T}_{HT,\phi_x} \big/ \left(1+\varphi^2\right)\right) \cdot \hat{T}_{HT,q\phi_x^2}^{-1} \cdot \hat{T}_{HT,q\phi_x\phi_y}$$
$$= f\left(\hat{T}_{HT}, \hat{T}_{HT,\phi_x}, \hat{T}_{HT,q\phi_x^2}, \hat{T}_{HT,q\phi_x\phi_y}\right).$$

Thus, the estimator $\hat{T}_{FS}$ can be viewed as a nonlinear function depending on the Horvitz-Thompson estimators $\hat{T}_{HT}, \hat{T}_{HT,\phi_x}, \hat{T}_{HT,q\phi_x^2}$ and $\hat{T}_{HT,q\phi_x\phi_y}$ that are unbiased, i.e.

$$E\hat{T}_{HT} = T,\; E\hat{T}_{HT,\phi_x} = T_{\phi_x},\; E\hat{T}_{HT,q\phi_x^2} = T_{q\phi_x^2},\; E\hat{T}_{HT,q\phi_x\phi_y} = T_{q\phi_x\phi_y}\;.$$

Using the Taylor linearization method, we derive a linear approximation of the function $\hat{T}_{FS}$.
The expansion of this estimator in a Taylor series up to the first order terms at the point $\left(\hat{T}_{HT}, \hat{T}_{HT,\phi_x}, \hat{T}_{HT,q\phi_x^2}, \hat{T}_{HT,q\phi_x\phi_y}\right) = \left(T, (1+\varphi^2)T_{\phi_x}, T_{q\phi_x^2}, T_{q\phi_x\phi_y}\right)$ is

$$\hat{T}_{FS}^{(linear)} = \frac{1}{1+\varphi^2}\left[\hat{T}_{HT} - B\left(\hat{T}_{HT,\phi_x} - (1+\varphi^2)T_{\phi_x}\right)\right]. \tag{20}$$

The approximate variance of the estimator $\hat{T}_{FS}$ is equal to the exact variance of $\hat{T}_{FS}^{(linear)}$:

$$AV\left(\hat{T}_{FS}\right) = V\left(\hat{T}_{FS}^{(linear)}\right).$$

By calculating the variance $V\left(\hat{T}_{FS}^{(linear)}\right)$ and taking into account equality (18), we get the expression of $AV\left(\hat{T}_{FS}\right)$.

The approximate variance of the estimator $\hat{T}_{lin}$ is obtained similarly, using an expansion of $\hat{T}_{lin}$ in a Taylor series at the point

$$\left(\hat{T}_{HT}, \hat{T}_{HT,\phi_x}, \hat{T}_{HT,q\phi_x^2}, \hat{T}_{HT,q\phi_x\phi_y}\right) = \left(T, \frac{(1-\alpha)(1+\varphi^2)}{1-\alpha+\beta\varphi^2}T_{\phi_x}, T_{q\phi_x^2}, T_{q\phi_x\phi_y}\right). \qquad \square$$

Remark 3. The estimator of the variance of the estimator $\hat{T}_{DS}$ can be defined as follows. First, we write

$$AV\left(\hat{T}_{DS}\right) = V\left(\sum_{i\in\mathbf{s}}\sum_{j>i} d_{ij}\left(\phi(\mathbf{y}_i, \mathbf{y}_j) - B\phi(\mathbf{x}_i, \mathbf{x}_j)\right)\right) = V\left(\hat{T}_{HT,\phi_y - B\phi_x}\right),$$

where $\hat{T}_{HT,\phi_y - B\phi_x}$ is the Horvitz-Thompson estimator of the total of the variable $\phi_y - B\phi_x$ defined on the population $\mathbf{U}^*$ and taking values $\phi_{y1} - B\phi_{x1}$, $\phi_{y2} - B\phi_{x2}$, ..., $\phi_{yN^*} - B\phi_{xN^*}$. Here $\phi_{ym} = \phi(\mathbf{y}_i, \mathbf{y}_j)$, $\phi_{xm} = \phi(\mathbf{x}_i, \mathbf{x}_j)$ for a pair of indices $m = (ij)$. Then, using Result 2.8.1 from (Särndal, Swensson and Wretman, 1992, p. 43), we get the more convenient expression of the approximate variance

$$AV\left(\hat{T}_{DS}\right) = V\left(\hat{T}_{HT,\phi_y - B\phi_x}\right) = \sum_{r=1}^{N^*}\sum_{m=1}^{N^*}\left(\pi^*_{rm} - \pi^*_r\pi^*_m\right)\frac{\phi_{yr} - B\phi_{yr}}{\pi^*_r}\cdot\frac{\phi_{ym} - B\phi_{ym}}{\pi^*_m},$$

where π^*_m and π^*_{rm} are the first and second order inclusion probabilities over the synthetic population $\mathbf{U}^*$. In fact, π^*_m and π^*_{rm} coincide with the second and fourth order inclusion probabilities over the population $\mathbf{U}$: $\pi^*_m = \pi_{ij}$ and $\pi^*_{rm} = \pi_{ijkl}$ for $m = (ij)$ and $r = (kl)$, where π_{ij} and π_{ijkl} are assumed to be strictly positive (see Subsection 2.1).

As the estimator of the variances $V\left(\hat{T}_{DS}\right)$ and $V\left(\hat{T}_{HT,\phi_y - B\phi_x}\right)$, we take

$$\hat{V}\left(\hat{T}_{DS}\right) = \hat{V}\left(\hat{T}_{HT,\phi_y - B\phi_x}\right) = \sum_{r\in\mathbf{s}^*}\sum_{m\in\mathbf{s}^*}\left(1 - \frac{\pi^*_r\pi^*_m}{\pi^*_{rm}}\right)\frac{\phi_{yr} - \hat{B}\phi_{yr}}{\pi^*_r}\cdot\frac{\phi_{ym} - \hat{B}\phi_{ym}}{\pi^*_m},$$

where $\hat{B} = \hat{T}_{HT,q\phi_x\phi_y} \Big/ \hat{T}_{HT,q\phi_x^2}$.

The estimators of the variances $V\left(\hat{T}_{FS}\right)$ and $V\left(\hat{T}_{lin}\right)$ can be derived in a similar way.

Alternatively, the replication methods, such as the jackknife, bootstrap and balanced halfsamples (see e.g. Särndal, Swensson and Wretman, 1992), can be used for estimating the variances of the estimators $\hat{T}_{DS}$, $\hat{T}_{FS}$ and $\hat{T}_{lin}$.

3.2. Comparison of variances

Next, we will compare the variances of estimators by analyzing their differences. Let us start from the difference $AV\left(\hat{T}_{DS}\right) - V\left(\hat{T}_{HT}\right)$. Using equality (18), it can be easily expressed as follows

$$AV\left(\hat{T}_{DS}\right) - V\left(\hat{T}_{HT}\right) = B^2V\left(\hat{T}_{HT,\phi_x}\right) - 2BC\left(\hat{T}_{HT}, \hat{T}_{HT,\phi_x}\right).$$

Thus, $AV\left(\hat{T}_{DS}\right) \le V\left(\hat{T}_{HT}\right)$, if

$$\rho\left(\hat{T}_{HT}, \hat{T}_{HT,\phi_x}\right) \ge \frac{1}{2} B \sqrt{V\left(\hat{T}_{HT,\phi_x}\right) / V\left(\hat{T}_{HT}\right)}, \text{ as } B > 0, \text{ or}$$
$$\rho\left(\hat{T}_{HT}, \hat{T}_{HT,\phi_x}\right) \le \frac{1}{2} B \sqrt{V\left(\hat{T}_{HT,\phi_x}\right) / V\left(\hat{T}_{HT}\right)}, \text{ as } B < 0, \tag{21}$$

where $\rho\left(\hat{T}_{HT}, \hat{T}_{HT,\phi_x}\right)$ is the correlation coefficient between $\hat{T}_{HT}$ and $\hat{T}_{HT,\phi_x}$.

A negative value of $B = T_{q\phi_x\phi_y} / T_{q\phi_x^2}$ may occur if the range of the function $\phi(\bullet,\bullet)$ is **R**, e.g. when T is a covariance between two study variables.

In the case of simple random sampling without replacement, inequalities (21) reduce to

$$\rho\left(\phi_y, \phi_x\right) \ge \frac{1}{2} B \sqrt{S_{\phi_x}^2 / S_{\phi_y}^2}, \text{ as } B > 0,$$
$$\rho\left(\phi_y, \phi_x\right) \le \frac{1}{2} B \sqrt{S_{\phi_x}^2 / S_{\phi_y}^2}, \text{ as } B < 0,$$

where $\rho\left(\phi_y, \phi_x\right)$ is the correlation coefficient between the variables $\phi_y : \phi_{y1}, \phi_{y2}, ..., \phi_{yN^*}$ and $\phi_x : \phi_{x1}, \phi_{x2}, ..., \phi_{xN^*}$, defined on the population $\mathbf{U}^*$. The notation $S_{\phi_y}^2$, $S_{\phi_x}^2$ is used to denote variances of the variables ϕ_y and ϕ_x, respectively.

Comparison of calibrated and penalized calibration estimators. Equality (19) shows that the approximate variance of the penalized estimator $\hat{T}_{FS}$ is lower than that of the calibrated estimator $\hat{T}_{DS}$. Even more, according to the lines of Farrell and Singh (2002), the approximate mean square of the estimator $\hat{T}_{FS}$ is minimized when

$$\varphi^2 = AV\left(\hat{T}_{DS}\right) / \left(T - BT_{\phi_x}\right)^2, \tag{22}$$

and it equals

$$AMSE_{\min}\left(\hat{T}_{FS}\right) = \frac{1}{1+\varphi^2} AV\left(\hat{T}_{DS}\right).$$

One can easily verify this equality using the same expansion (20) of $\hat{T}_{FS}$ in a Taylor series.

Comparison of the calibrated estimators $\hat{T}_{DS}$ and $\hat{T}_{lin}$. First, we derive the optimal values of coefficients α and β which minimize the approximate

variance of the estimator $\hat{T}_{lin}$ under the condition $AE(\hat{T}_{lin}) = T$ that indicates the approximate unbiasedness of the estimator. Denote the optimal values of α and β by $\alpha_{\min}$ and $\beta_{\min}$, respectively.
Let us define the Lagrange function

$$\Lambda^* = AV(\hat{T}_{lin}) - \lambda^* \left(AE(\hat{T}_{lin}) - T\right).$$

By solving the equations

$$\frac{\partial \Lambda^*}{\partial \alpha} = 0, \ \frac{\partial \Lambda^*}{\partial \beta} = 0 \text{ and } AE(\hat{T}_{lin}) = T,$$

we find

$$\begin{aligned} \alpha_{\min} &= \frac{AV(\hat{T}_{DS}) - C(\hat{T}_{HT} - B\hat{T}_{HT,\phi_x}, \hat{T}_{HT})}{B^2 V(\hat{T}_{HT,\phi_x})}, \\ \beta_{\min} &= \frac{V(\hat{T}_{HT}) - C(\hat{T}_{HT} - B\hat{T}_{HT,\phi_x}, \hat{T}_{HT})}{B^2 V(\hat{T}_{HT,\phi_x})}. \end{aligned} \tag{23}$$

The second derivatives test for critical points shows that the values $\alpha_{\min}$ and $\beta_{\min}$ satisfy the condition of the minimal approximate variance which is equal to

$$\begin{aligned} AV_{\min}(\hat{T}_{lin}) &= \frac{AV(\hat{T}_{DS})V(\hat{T}_{HT}) - C^2(\hat{T}_{HT} - B\hat{T}_{HT,\phi_x}, \hat{T}_{HT})}{AV(\hat{T}_{DS}) + V(\hat{T}_{HT}) - 2C(\hat{T}_{HT} - B\hat{T}_{HT,\phi_x}, \hat{T}_{HT})} \\ &= \frac{AV(\hat{T}_{DS})V(\hat{T}_{HT})\left(1 - \rho^2(\hat{T}_{HT} - B\hat{T}_{HT,\phi_x}, \hat{T}_{HT})\right)}{B^2 V(\hat{T}_{HT,\phi_x})}, \end{aligned}$$

where $\rho(\hat{T}_{HT} - B\hat{T}_{HT,\phi_x}, \hat{T}_{HT})$ is the correlation coefficient between the estimators $\hat{T}_{HT,\phi_y - B\phi_x} = \hat{T}_{HT} - B\hat{T}_{HT,\phi_x}$ and $\hat{T}_{HT}$.

The difference

$$AV_{\min}(\hat{T}_{lin}) - AV(\hat{T}_{DS}) = -\frac{\left(AV(\hat{T}_{DS}) - C(\hat{T}_{HT} - B\hat{T}_{HT,\phi_x}, \hat{T}_{HT})\right)^2}{B^2 V(\hat{T}_{HT,\phi_x})}$$

is negative or is equal to zero. It means that the approximate variance of the estimator $\hat{T}_{lin}$ is not higher than that of the calibrated estimator $\hat{T}_{DS}$.

Note that $\alpha_{\min} + \beta_{\min} = 1$. Therefore, under such a setting of optimal coefficients, the estimator $\hat{T}_{FS}$ is not included into the linear combination (14), and, consequently, $\hat{T}_{lin} = \alpha_{\min} \hat{T}_{HT} + \beta_{\min} \hat{T}_{DS}$.

Comparison of the calibrated estimators $\hat{T}_{FS}$ and $\hat{T}_{lin}$. The difference between the approximate variances $AV_{\min}\left(\hat{T}_{lin}\right)$ and $AV\left(\hat{T}_{FS}\right)$ can be expressed as follows:

$$AV_{\min}\left(\hat{T}_{lin}\right) - AV\left(\hat{T}_{FS}\right) = \frac{AV\left(\hat{T}_{DS}\right)V\left(\hat{T}_{HT}\right)}{-\left(AV\left(\hat{T}_{DS}\right) + V\left(\hat{T}_{HT}\right) - 2C\left(\hat{T}_{HT} - B\hat{T}_{HT,\phi_x}, \hat{T}_{HT}\right)\right)\left(1+\varphi^2\right)^2}$$
$$\times \left[\rho^2\left(\hat{T}_{HT} - B\hat{T}_{HT,\phi_x}, \hat{T}_{HT}\right)\left(1+\varphi^2\right)^2 + \frac{\left(AV\left(\hat{T}_{DS}\right) - 2C\left(\hat{T}_{HT} - B\hat{T}_{HT,\phi_x}, \hat{T}_{HT}\right)\right)}{V\left(\hat{T}_{HT}\right)} - \varphi^2\left(2+\varphi^2\right)\right].$$

Since

$$AV\left(\hat{T}_{DS}\right) + V\left(\hat{T}_{HT}\right) - 2C\left(\hat{T}_{HT} - B\hat{T}_{HT,\phi_x}, \hat{T}_{HT}\right) = B^2 V\left(\hat{T}_{HT,\phi_x}\right),$$

the inequality $AV_{\min}\left(\hat{T}_{lin}\right) - AV\left(\hat{T}_{FS}\right) \le 0$ is equivalent to that

$$\rho^2\left(\hat{T}_{HT} - B\hat{T}_{HT,\phi_x}, \hat{T}_{HT}\right) \ge 1 - \frac{B^2 V\left(\hat{T}_{HT,\phi_x}\right)}{\left(1+\varphi^2\right)^2 V\left(\hat{T}_{HT}\right)}. \quad (24)$$

Under condition (24) the approximate variance of the estimator $\hat{T}_{lin}$ is not higher than that of the penalized calibration estimator $\hat{T}_{FS}$.

In the case of simple random sampling without replacement, inequality (24) reduces to

$$\rho^2\left(\phi_y - B\phi_x, \phi_y\right) \ge 1 - \frac{B^2 S^2_{\phi_x}}{(1+\varphi^2)^2 S^2_{\phi_y}}.$$

4. Simulation study

4.1. Simulation setup

The simulation study is performed to observe the efficiency of the Horvitz-Thompson type and calibrated estimators of the finite population variance and covariance. For that purpose, we consider a subset (of size 300) of the real population from the Lithuanian Enterprise Survey. The study variables y and z are the profit of an enterprise in a different time period, whereas the values of auxiliary variables a and b are the numbers of employees of the enterprise at the same periods. The correlation coefficients between the study and auxiliary variables are $\rho(y,a)=0.81$, $\rho(z,b)=0.90$, $\rho(y,b)=0.63$ and $\rho(z,a)=0.60$. The relationship between the variables ϕ_y and ϕ_x is also strong enough, because $\rho(\phi_y,\phi_x)=0.64$ (when the estimated parameter is a finite population variance) and $\rho(\phi_y,\phi_x)=0.88$ (in the case of estimation of a finite population covariance).

The population is stratified into two strata by the size of the survey variable y. The stratified simple random sample is used as a sample design. The sample size $n=100$ is allocated to strata, using Neyman's optimal allocation. $M=10000$ samples were drawn and for each of them the estimators $\hat{S}^2_{HT}$, $\hat{S}^2_{DS}$, $\hat{S}^2_{FS}$ and $\hat{S}^2_{lin}$ of the finite population variance S^2_y, and the estimators $\hat{C}_{HT}$, $\hat{C}_{DS}$, $\hat{C}_{FS}$ and $\hat{C}_{lin}$ of the finite population covariance $C(y,z)$ were computed.

The uniform weights $q_{ij}=1$ were used for the calibrated estimators. A value of the parameter φ was calculated using formula (22). The coefficients α and β that appear in the expression of $\hat{T}_{lin}$, were defined in the following way. We give here an explanation in the general case, where the parameter T is any quadratic function. First, we write

$$
\begin{aligned}
V\left(\hat{T}_{lin}\right) &= \alpha^2 V\left(\hat{T}_{HT}\right)+\beta^2 V\left(\hat{T}_{DS}\right)+(1-\alpha-\beta)^2 V\left(\hat{T}_{FS}\right)+2\alpha\beta C\left(\hat{T}_{HT},\hat{T}_{DS}\right)\\
&\quad +2\alpha(1-\alpha-\beta)C\left(\hat{T}_{HT},\hat{T}_{FS}\right)+2\beta(1-\alpha-\beta)C\left(\hat{T}_{DS},\hat{T}_{FS}\right), \qquad (25)\\
E\hat{T}_{lin} &= \alpha E\hat{T}_{HT}+\beta E\hat{T}_{DS}+(1-\alpha-\beta)E\hat{T}_{FS}.
\end{aligned}
$$

Then, by replacing in (25) true values of variances $V\left(\hat{T}_{HT}\right)$, $V\left(\hat{T}_{DS}\right)$, $V\left(\hat{T}_{FS}\right)$, expectations $E\hat{T}_{HT}$, $E\hat{T}_{DS}$, $E\hat{T}_{FS}$, and covariances between the estimators with the empirical ones, we minimize the empirical mean square error $MSE_{emp}\left(\hat{T}_{lin}\right)=V_{emp}\left(\hat{T}_{lin}\right)+\left(E_{emp}\hat{T}_{lin}-T\right)^2$ subject to the constraint $E_{emp}\hat{T}_{lin}=T$. The solution of this optimization problem ($\tilde{\alpha}_{\min}$, $\tilde{\beta}_{\min}$) is used in (14) for calculating estimates. This choice of the coefficients α and β often leads to a little bit more accurate estimates of $\hat{T}_{lin}$ as compared to those which are computed using $\alpha_{\min}$ and $\beta_{\min}$, defined by (23). Thus, using our data, we obtain

$\widetilde{\alpha}_{\min}=0.30$, $\widetilde{\beta}_{\min}=0.81$ ($\widetilde{\gamma}_{\min}=1-\widetilde{\alpha}_{\min}-\widetilde{\beta}_{\min}=-0.11$), if the parameter T is a population variance, and the values $\widetilde{\alpha}_{\min}=0.09$, $\widetilde{\beta}_{\min}=0.69$ ($\widetilde{\gamma}_{\min}=0.22$) were used in the case of estimation of a finite population covariance.

4.2. Simulation results

The empirical relative bias (RB), variance (V), mean square error (MSE), and the coefficient of variation (cv) have been calculated for each estimator (see Tables 1 and 2). For any estimator $\hat{\theta}$ of the finite population parameter θ, all these characteristics of accuracy are defined by the following equations:

$$RB\left(\hat{\theta}\right)=\frac{1}{M}\sum_{i=1}^{M}\frac{\hat{\theta}_i-\theta}{\theta},\ V\left(\hat{\theta}\right)=\frac{1}{M}\sum_{i=1}^{M}\left(\hat{\theta}_i-\frac{1}{M}\sum_{j=1}^{M}\hat{\theta}_j\right)^2,$$

$$MSE\left(\hat{\theta}\right)=\frac{1}{M}\sum_{i=1}^{M}\left(\hat{\theta}_i-\theta\right)^2,\ cv\left(\hat{\theta}\right)=\sqrt{Var\left(\hat{\theta}\right)}\cdot\left(\frac{1}{M}\sum_{i=1}^{M}\hat{\theta}_i\right)^{-1},$$

where $\hat{\theta}_i$ is the estimate of parameter θ, computed using data of the ith simulated sample.

Table 1. The main estimated characteristics of accuracy for the estimators of the finite population variance

Estimator	RB	$V\times10^{-17}$	$MSE\times10^{-17}$	cv
$\hat{S}^2_{HT}$	-0.0039	5.27	5.40	0.0991
$\hat{S}^2_{DS}$	0.0070	4.11	4.13	0.0855
$\hat{S}^2_{FS}[\varphi^2=0.09]$	-0.0053	3.95	3.96	0.0849
$\hat{S}^2_{lin}$	0.0037	3.64	3.64	0.0808

Table 2. The main estimated characteristics of accuracy for the estimators of the finite population covariance

Estimator	RB	$V\times10^{-13}$	$MSE\times10^{-13}$	cv
$\hat{C}_{HT}$	-0.0015	11.50	13.82	0.1752
$\hat{C}_{DS}$	0.0115	3.61	3.67	0.0899
$\hat{C}_{FS}[\varphi^2=0.10]$	-0.0066	3.32	3.34	0.0878
$\hat{C}_{lin}$	0.0010	3.27	3.27	0.0864

Due to quite a high correlation between the variables ϕ_y and ϕ_x, the calibrated estimators of population variance and covariance are much more accurate compared to the Horvitz-Thompson estimators of the same parameters. More significant differences between the characteristics of accuracy of the calibrated and Horvitz-Thompson estimators are observed in the case of estimation of the population covariance (see Table 2), where the coefficient of variation of the calibrated estimators $\hat{C}_{DS}$, $\hat{C}_{FS}$, $\hat{C}_{lin}$ is approximately two times lower than that of the only design-based estimator $\hat{C}_{HT}$. The variances and mean square errors differ about four times.

The estimators $\hat{S}^2_{lin}$ and $\hat{C}_{lin}$ which are the linear combinations of the estimators $\hat{S}^2_{HT}$, $\hat{S}^2_{DS}$, $\hat{S}^2_{FS}$ and $\hat{C}_{HT}$, $\hat{C}_{DS}$, $\hat{C}_{FS}$, respectively, outperform all the estimators from the corresponding group. The empirical analogs of the terms included into (24) do not satisfy this inequality. Thus, it seems that the approximate variances of the estimators $\hat{S}^2_{lin}$ and $\hat{C}_{lin}$ are higher than that of the corresponding penalized estimators $\hat{S}^2_{FS}$ and $\hat{C}_{FS}$, although the behaviour of empirical variances is contrary. The reason for that could be due to the linearization when only the first order Taylor approximations of the estimators are used for calculating approximate variances, and the remainder terms of a Taylor expansion are neglected.

Comparing the penalized calibration estimators $\hat{S}^2_{FS}$ and $\hat{C}_{FS}$ to the corresponding calibrated estimators $\hat{S}^2_{DS}$ and $\hat{C}_{DS}$, we note that not only the variance, but also the mean square error of the penalized estimators is lower than the variance of the calibrated estimators, as it is in the case of the approximate variance and mean square error (see Section 3).

5. Conclusion

Some of the estimators of finite population parameters can be treated as they are of quite a good quality in the sense of a small variance or small bias, but other characteristics of accuracy (e.g. mean square error) may not satisfy the survey statisticians and practitioners. In our case, the Horvitz-Thompson estimator $\hat{T}_{HT}$ is unbiased, but its variance may be relatively large. The calibrated estimator $\hat{T}_{DS}$ is preferable because of its lower variance (especially if the variables ϕ_y and ϕ_x

are well correlated), although it is slightly biased. Of course, the small bias has a minor impact on the results. The penalized calibration estimator $\hat{T}_{FS}$ has to be used carefully because an increase in the penalty (φ) may have a negative impact on the bias. The best properties of the estimators $\hat{T}_{HT}$, $\hat{T}_{DS}$ and $\hat{T}_{FS}$ are reflected by that of the estimator $\hat{T}_{lin}$. For some sets of coefficients α, β and γ, it is unbiased and may have the lowest variance among the estimators discussed in this paper.

REFERENCES

Deville, J.C. and Särndal, C. E., 1992. Calibration estimators in survey sampling. *Journal of the American Statistical Association,* 87, pp.376–382.

Farrell, P. and Singh, S., 2002. Penalized chi square distance function in survey sampling. *ASA Proceedings,* pp.963–968.

Guggemos, F. and Tillé, Y., 2010. Penalized calibration in survey sampling: Design-based estimation assisted by mixed models. *Journal of Statistical Planning and Inference,* 140, pp.3199–3212.

Plikusas, A. and Pumputis, D., 2007. Calibrated estimators of the population covariance. *Acta Applicandae Mathematicae*, 97, pp.177–187.

Plikusas, A. and Pumputis, D., 2010. Estimation of the finite population covariance using calibration. *Nonlinear Analysis: Modelling and Control*, 15(3), pp.325–340.

Särndal, C.E., 2007. The calibration approach in survey theory and practice. *Survey Methodology*, 33(2), pp.99–119.

Särndal, C.E. Swensson, B. and Wretman, J., 1992. *Model Assisted Survey Sampling*. New York: Springer-Verlag.

Singh, S., 2003. On Farrell and Singh's penalized chi square distance functions in survey sampling. *SCC Proceedings,* pp.173–178.

Singh, S. Horn, S. Chowdhury, S. and Yu, F., 1999. Calibration of the estimators of variance. *Austral. & New Zealand J. Statist.,* 41(2), pp.199–212.

Sitter, R.R. and Wu, C., 2002. Efficient estimation of quadratic finite population functions in the presence of auxiliary information. *Journal of the American Statistical Association,* 97(458), pp.535–543.

Yates, F. and Grundy, P., 1953. Selection without replacement from within strata with probability proportional to size. *Journal of the Royal Statistical Society*, 15(2), pp.253–261.