Krakow Review of Economics and Management
Zeszyty Naukowe Uniwersytet Ekonomiczny w Krakowie
2 (1004)
2024

KREM, 2024, 2(1004): 121–138
ISSN 1898-6447
e-ISSN 2545-3238
https://doi.org/10.15678/KREM.2024.1004.0207

On Bootstrap Algorithms in Survey Sampling

Tomasz Żądło

University of Economics in Katowice, Department of Statistics, Econometrics and Mathematics, 1 Maja 50, 40-287 Katowice, Poland, e-mail: tomasz.zadlo@uekat.pl,
ORCID: https://orcid.org/0000-0003-0638-0748

Suggested citation: Żądło, T. (2024). On Bootstrap Algorithms in Survey Sampling. *Krakow Review of Economics and Management / Zeszyty Naukowe Uniwersytetu Ekonomicznego w Krakowie*, 2(1004), 121–138. https://doi.org/10.15678/KREM.2024.1004.0207

ABSTRACT

Objective: The aim of this paper is to present bootstrap algorithms for measuring the accuracy of estimation and prediction in design-based and model-based approaches in survey sampling and small area estimation. Three proposals of prediction-mean squared error estimators are also examined.
Research Design & Methods: Various bootstrap procedures are shown and used to estimate the design- and prediction-mean squared errors based on real data. Computations are supported by two R packages.
Findings: Three prediction-mean squared error estimators are proposed.
Implications/Recommendations: The bootstrap algorithms used in the design-based approach give similar results for the considered data for the variance estimates of the considered estimator, implying that the speed of the algorithms may be important for practitioners in cases of similar properties. The proposed estimators of the prediction mean squared error produce higher estimates than other estimators in the model-based approach, indicating a positive bias that can be interpreted as a pessimistic accuracy estimate.
Contribution: All the presented bootstrap algorithms are easily applicable using two R packages available at R CRAN and GitHub. Three double bootstrap prediction-MSE estimators are proposed and analysed in the real-data application.

Article type: original article.

Keywords: survey sampling, small area estimation, bootstrap, estimation and prediction accuracy.

JEL Classification: C83, C88.

1. Introduction

In survey sampling, two main approaches are considered – the classic design-based approach usually used to estimate the population parameters and the model-based approach typically used to predict subpopulation characteristics. In both approaches, accuracy estimation is among the crucial problems. Many estimators of accuracy measures are based on the various bootstrap algorithms discussed in this paper.

In the design-based approach, the inference is based on a random sample denoted by s. The variable of interest, typically denoted by y, is treated as fixed (non-random). Let $\mathbf{P}$, called the sampling space, be the set of all sample sets s. The sampling design is defined as the probability distribution $P(s)$ defined on the sampling space $\mathbf{P}$, such that $P(s) \geq 0$ and $\sum_{s \in \mathbf{P}} P(s) = 1$ for all $s \in \mathbf{P}$. As pointed out by Cassel, Särndal and Wretman (1977, p. 19), the data obtained based on a sample s and the associated observations of y's can be written as $d = ((k, y_k): k \in s)$. In the design-based approach, d is treated as the realisation of $D = ((k, y_k): k \in S)$, where S is a random variable whose realisations are possible samples drawn from the population. The statistic, denoted by $Z = u(D)$, is a function defined on the sample space of the random variable D, such that it depends on the variable of interest only through y_k for which $k \in s$ (Cassel, Särndal & Wretman, 1977, p. 20). The estimator of the parameter θ is a statistic $\hat{\theta}$ with values which belong to the parameter space. Hence, the sampling design is usually assumed to be the only source of the estimator's randomness. Before proceeding, it is necessary to define the inclusion probability of order r, denoted by $\pi_{k_1 \ldots k_r}$. It was defined by Cassel, Särndal and Wretman (1977, p. 11) as the probability that population elements $k_1, k_2, \ldots, k_r$ are included in the sample: $\pi_{k_1 \ldots k_r} = \sum_{s \in A(k_1 \ldots k_r)} P(s)$ where $A(k_1 \ldots k_r) = \{s: k_i \in s, \text{for } i = 1, \ldots, r\}$. Hence, the first order inclusion probability of an element k is the sum of probabilities of selecting those samples s that contain the element. Therefore, its inverse can be interpreted as the number of population elements represented by the sample element k. The measure of the design-based accuracy is the design-MSE (mean squared error) defined as (Cassel, Särndal & Wretman, 1977, p. 26):

$$MSE_p(\hat{\theta}) = E_p(\hat{\theta} - \theta)^2 = Var_p(\hat{\theta}) + B_p^2(\hat{\theta}), \tag{1}$$

where θ is a parameter, $\hat{\theta}$ is an estimator, $E_p(\,.\,)$ is the expectation with respect to the sampling design, $Var_p(\hat{\theta}) = E_p\left(\hat{\theta} - E_p(\hat{\theta})\right)^2$ is the measure of the design--precision and is called the design-variance, and $B_p(\hat{\theta}) = E_p(\hat{\theta}) - \theta$ is the design-bias.

In the model-based approach, the inference can be made based on both random and purposive samples. It is assumed that the observed vector of the variable of interest is the realisation of a certain vector of random variables $\mathbf{Y} = [Y_1, Y_2, \ldots, Y_N]^T$ with some joint distribution denoted by ζ. The superpopulation model is the set of conditions that define the class of distributions to which distribution ζ is assumed to belong. In this approach (Cassel, Särndal & Wretman, 1977, p. 91), the observed data $d = ((k, y_k): k \in s)$ are treated as the realisation of $D' = ((k, Y_k): k \in s)$. What is more, the statistic denoted by $Z = u(D')$ is a function, that for any given sample s depends on $Y_1, Y_2, \ldots, Y_N$ only through those Y_k such that $k \in s$ (Cassel, Särndal & Wretman, 1977, p. 91). If a statistic $\hat{\theta}$ is used to predict θ, which is a function of $Y_1, Y_2, \ldots, Y_N$, it is called a predictor of θ (Cassel, Särndal & Wretman, 1977, p. 91). Because values of the variable of interest are assumed to be realisations of the random variables, the predicted characteristic $\theta = \theta(\mathbf{Y})$ is random as well. Finally, it is commonly assumed that the distribution of the variable of interest is the only source of a predictor's randomness.

Some basic notations are introduced to present the linear mixed model, which is widely used in survey sampling. The population Ω of size N is divided into two sets: the sample s of size n and the set of non-sampled elements $\Omega_r = \Omega \backslash s$ of size $N_r = N - n$. The values of the variable of interest $y_1, y_2, \ldots, y_N$ are assumed to be realisations of random variables $Y_1, Y_2, \ldots, Y_N$. The population vector of size N of the random variables is denoted by $\mathbf{Y}$. There are also two known population matrices of auxiliary variables of sizes $N \times p$ and $N \times q$, denoted by $\mathbf{X}$ and $\mathbf{Z}$ and associated with fixed and random effects, respectively.

Rao and Molina (2015, p. 98) assume that the population data obey the following linear mixed model (LMM) assumptions:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{v} + \mathbf{e}, \tag{2}$$

where $\mathbf{v} \sim N(\mathbf{0}, \mathbf{G}(\boldsymbol{\delta}))$ and $\mathbf{e} \sim N(\mathbf{0}, \mathbf{R}(\boldsymbol{\delta}))$ are called vectors of random effects and random components, and are assumed to be independent. The vector of unknown parameters of the LMM is denoted by $\boldsymbol{\Psi} = [\boldsymbol{\beta}^T \ \boldsymbol{\delta}^T]^T$, where $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$ are vectors of fixed effects and variance components, respectively.

Without loss of generality, we assume that only first n realisations of Y_i are observed in the sample, and thus, we can decompose $\mathbf{Y}$ as follows: $\mathbf{Y} = \left[\mathbf{Y}_s^T \ \mathbf{Y}_r^T\right]^T$, where $\mathbf{Y}_s$ and $\mathbf{Y}_r$ are of sizes $n \times 1$ and $N_r \times 1$, respectively. In all notations, the subscript "s" is used for the sample and "r" for non-sampled elements. Similarly to the decomposition of the $\mathbf{Y}$ vector, matrices of auxiliary information in (2) can be decomposed

as follows: $\mathbf{X}=\left[\mathbf{X}_s^T\ \mathbf{X}_r^T\right]^T$ and $\mathbf{Z}=\left[\mathbf{Z}_s^T\ \mathbf{Z}_r^T\right]^T$. Matrices $\mathbf{R}(\boldsymbol{\delta})$ and $\mathbf{G}(\boldsymbol{\delta})$ are variance--covariance matrices of random vectors $\mathbf{e}$ and $\mathbf{v}$, respectively. The variance-covariance matrix of the dependent variable is denoted by $Var(\mathbf{Y})=\mathbf{V}(\boldsymbol{\delta})=\mathbf{Z}\mathbf{G}(\boldsymbol{\delta})\mathbf{Z}^T+\mathbf{R}(\boldsymbol{\delta})$. $\mathbf{V}$ and $\mathbf{R}$ can be decomposed as follows: $\mathbf{V}=\begin{bmatrix}\mathbf{V}_{ss} & \mathbf{V}_{sr}\\ \mathbf{V}_{rs} & \mathbf{V}_{rr}\end{bmatrix}$, $\mathbf{R}=\begin{bmatrix}\mathbf{R}_{ss} & \mathbf{R}_{sr}\\ \mathbf{R}_{rs} & \mathbf{R}_{rr}\end{bmatrix}$, where $\mathbf{V}_{ss}=Var(\mathbf{Y}_s)$, $\mathbf{V}_{rr}=Var(\mathbf{Y}_r)$, $\mathbf{V}_{sr}=Cov(\mathbf{Y}_s,\mathbf{Y}_r)$, $\mathbf{V}_{rs}=Cov(\mathbf{Y}_r,\mathbf{Y}_s)$ and similarly $\mathbf{R}_{ss}=$ $=Var(\mathbf{e}_s)$, $\mathbf{R}_{rr}=Var(\mathbf{e}_r)$, $\mathbf{R}_{sr}=Cov(\mathbf{e}_s,\mathbf{e}_r)$, $\mathbf{R}_{rs}=Cov(\mathbf{e}_r,\mathbf{e}_s)$.

The prediction-MSE (mean squared error) is a measure of prediction accuracy, defined as follows (Rao & Molina, 2015, p. 98):

$$MSE_\zeta(\hat{\theta})=E_\zeta(\hat{\theta}-\theta)^2=Var_\zeta(\hat{\theta}-\theta)+B_\zeta^2(\hat{\theta}), \tag{3}$$

where θ is a predicted random characteristic, $\hat{\theta}$ is a predictor, $E_\zeta(\,.\,)$ is the expectation with respect to the ζ distribution, $Var_\zeta(\hat{\theta})=E_\zeta(\hat{\theta}-\theta-E_\zeta(\hat{\theta}-\theta))^2$ is the measure of the prediction-precision known as the prediction-variance, $B_\zeta(\hat{\theta})=$ $=E_\zeta(\hat{\theta}-\theta)$ is the prediction-bias.

In this paper we present bootstrap algorithms for estimating the accuracy and precision of any estimator and predictor in both approaches. In sections 2 and 3, two bootstrap approaches in the design-based inference are studied. Different bootstrap approaches used in the model-based inference are introduced in sections 4, 5, and 6, including our new proposals of three prediction-MSE estimators based on the double bootstrap. In sections 7 and 8, real data applications are presented using two R packages: *pipsboot* (Kucharski & Żądło, 2021) and *qape* (Wolny-Dominiak & Żądło, 2022b).

2. The ad-hoc Approach in Design-based Inference

The classic Efron's bootstrap (Efron, 1979) procedure, where simple random samples are drawn with replacement from the original sample, is correct under the independence of random variables. In the case of complex sampling designs, appropriate modifications are required.

Let S_k^* be a number of replications of element k in the bootstrap sample. In the case of methods based on the *ad-hoc* approach, whether the following equalities hold for the proposed method (Antal & Tillé, 2014, p. 1349) is usually checked:

$$E^*(S_k^*)=1,\ k\in S, \tag{4}$$

$$Var^*(S_k^*)=1-\pi_k,\ k\in S, \tag{5}$$

$$Cov^*(S_k^*,S_l^*)=(\pi_{kl}-\pi_k\pi_l)\pi_{kl}^{-1},\ k\neq l\in S. \tag{6}$$

If these equality conditions are met for a bootstrap method, the bootstrap variance estimator of the Horvitz-Thompson estimator is equal to the unbiased classic vari-

ance estimator proposed by Horvitz and Thompson (1952). A modification of (5) and (6) can also be considered (Antal & Tillé, 2014, p. 1349) to ensure that the bootstrap variance estimator of the Horvitz-Thompson estimator is equal to the unbiased SYG variance estimator proposed by Sen (1953) and Yates and Grundy (1953). Usually, only the first two conditions (4) and (5) are considered because it is difficult to meet (6) exactly for fixed sampling designs with unequal inclusion probabilities, as stated by Antal and Tillé (2014, p. 1349). It should also be noted that meeting these conditions ensures unbiasedness of the bootstrap variance estimator of the Horvitz-Thompson estimator. However, if they are not met it does not imply that the bootstrap variance estimator of the Horvitz-Thompson estimator is less accurate than any other (including classic) variance estimator. Furthermore, these conditions are defined for the Horvitz-Thompson estimator, whereas in practice, other estimators can also be used (for example, those that use auxiliary information).

According to Ranalli and Mecatti (2012), the majority of bootstrap methods for complex sampling designs can be classified into one of two approaches. The first, the ad-hoc approach, is usually based on iid resampling and rescaling sample data. Methods using this approach include the rescaling bootstrap (Rao & Wu, 1988), the mirror-match bootstrap (Sitter, 1992), and the generalised weighted bootstrap (Beaumont & Patak, 2012). Proposals presented by Antal and Tillé (2011, 2014) are also considered within this approach.

One of these algorithms is presented below in detail as an example of the ad-hoc approach. The authors mix several sampling designs for resampling to meet two conditions: (4) and (5).

Algorithm 1. The bootstrap algorithm proposed by Antal and Tillé (2014, pp. 1355–1356) for the proportional-to-size sampling is presented below. Let function $H(.;.)$ be a function used to compute first order inclusion probabilities for the original sampling design, i.e., $\pi_k = H(x_1, x_2, \ldots, x_N; n)$.

1. Selecting the initial bootstrap sample. Use the Poisson sampling design with original inclusion probabilities π_k to choose units from S, where $X_k\,(k \in S)$ will denote the initial sample membership indicator. The units selected once form the initial bootstrap sample: $S_k^* = X_k$. Let $m = \sum_{k \in S} X_k$, and hence, $E_p(m) = \sum_{k \in S} \pi_k$.

2. Completing the bootstrap sample.

a) If $n - m \geq 2$, the double half sampling design is used among units such that $X_k = 0$ to draw $n - m$ elements. If $n - m$ is even, the procedure of the double half sampling design is defined as a sampling of $0.5(n-m)$ elements with simple random sampling without replacement, and then, each selected unit is taken twice. If $n - m$ is odd, the double half sampling design is slightly modified, as described by Antal and Tillé (2014, pp. 1351–1352).

b) If $n - m = 1$, the S_k^* is redefined as follows:

– with 0.5 probability: $S_k^* = 1$ for $k \in S$, i.e., the bootstrap sample is defined as the original sample,

– with 0.5 probability:

compute $\pi_{k|n-1} = E\left(X_k \middle| m = n-1\right) = 1 - (1-\pi_k)\pi_k^{-1}\left(\sum_{l \in S}(1-\pi_l)\pi_l^{-1}\right)^{-1}$. Then, using an unequal inclusion probabilities sampling design with a fixed sample size, draw $n - 2$ elements from S with probabilities $\psi_k = 1 - H_k(1 - \pi_{k|n-1}, k \in S; 2)$ and take them to the bootstrap sample once. Finally, draw a double half sample from the two remaining units.

Bootstrap estimators of the design-variance and the design-bias are defined as follows (e.g., Rao & Wu, 1988):

$$\widehat{Var}_p^{boot}(\hat{\theta}) = \frac{1}{B-1}\sum_{b=1}^{B}\left(\hat{\theta}_b^* - \frac{1}{B}\sum_{b=1}^{B}\hat{\theta}_b^*\right)^2, \tag{7}$$

$$\widehat{B}_p^{boot}(\hat{\theta}) = \frac{1}{B}\sum_{b=1}^{B}\hat{\theta}_b^* - \hat{\theta}, \tag{8}$$

where $\hat{\theta}$ is the value of the considered estimator computed based on the original sample and $\hat{\theta}_b^*$ is the value of the estimator computed for the bootstrap sample obtained in the bth iteration.

3. Plug-in Approach in Design-based Inference

The second approach is the plug-in approach. It is based on the concept of a pseudopopulation, though the pseudopopulation is not physically generated in some methods.

Algorithm 2. The general algorithm of typical bootstrap procedures for complex sampling designs based on the plug-in approach is as follows (compare Barbiero & Mecatti, 2010):

1. Define (possibly non-integer) weights w_k that determine how many population elements are represented by a given sample element.
2. If weights $w_k (k \in S)$ are not integers, replace them by integers denoted by $\tilde{w}_k$.
3. Build a pseudopopulation $\Omega^* = \{1, 2, \ldots, k^*, \ldots, N^{pseudo}\}$ by replicating the original sample elements $\tilde{w}_k$-times $(k \in S)$ each.
4. Draw a bootstrap sample S^* of size n (the original sample size) from Ω^* mimicking the original sampling design.
5. Compute the value of an estimator $\hat{\theta}$ on bootstrap sample s^*. Denote it by $\hat{\theta}^*$.
6. Steps 4 and 5 are iterated B times, which gives $\hat{\theta}_b^*$ for $b = 1, 2, \ldots, B$.

Many bootstrap algorithms can be described using Algorithm 2 with various definitions of w_k and $\tilde{w}_k$. Holmberg (1988) defines integer weights as follows: $w_k = \tilde{w}_k = \lfloor \pi_k^{-1} \rfloor + \epsilon_k$, where $\lfloor \pi_k^{-1} \rfloor$ is rounded down value of π_k^{-1} and ϵ_k is generated once from Bernoulli distribution with probability $\pi_k^{-1} - \lfloor \pi_k^{-1} \rfloor$. Barbiero and Mecatti (2010)

propose the 0.5 bootstrap with $w_k = \pi_k^{-1}$ and $\tilde{w}_k$ defined as w_k rounded to the nearest integer. Barbiero, Manzi and Mecatti (2015) present a similar proposal, but w_k are calibration weights (see e.g. Deville & Särndal, 1992) and $\tilde{w}_k$ are defined as w_k rounded to the nearest integer. Barbiero and Mecatti's (2010) two proposals for an x-balanced bootstrap are slightly different. The numbers of replications $w_k = \tilde{w}_k$ are rounded-down inverses of first order inclusion probabilities, to which additional pseudoelements are added to achieve the minimum absolute difference between total values of an auxiliary variable in the real population and the pseudopopulation.

Żądło (2021) proposed a fast procedure – a generalisation of the bootstrap algorithm presented by Quatember (2014) for probability proportional-to-size sampling, inspired by the plug-in approach. Here, though, the pseudopopulation is not physically constructed. The fact that the pseudo-population does not have to be created is crucial for the calculation time in this iterative algorithm. The proposed procedure is a draw-by-draw algorithm. In the *b*th bootstrap sample $(b = 1, 2, \ldots, B)$ element k is drawn from the original sample in the *j*th draw $(j = 1, 2, \ldots, n)$ with probability:

$$(w_k - h_{k,j-1}) \times x_k \left(t_x - \sum_{i \in s_{b,j-1}} x_i \right)^{-1}, \tag{9}$$

where w_k's are some calibration weights such that $\sum_{i \in s} w_i x_i = \sum_{i \in \Omega} x_i$ (see Deville & Särndal, 1992), $h_{k,j-1}$ is the number of replications of element k selected in the bootstrap procedure in the first $j - 1$ draws, and $s_{b,j-1}$ is the subset of the *b*th bootstrap sample after draw $j - 1$.

4. Parametric Bootstrap in Model-based Inference

In the two preceding sections we have presented bootstrap algorithms for the design-based inference. In this and subsequent sections, the model-based inference is analysed. Let us consider the problem of predicting a function of the population vector $\mathbf{Y}$ denoted by $\theta = \theta(\mathbf{Y})$ using a predictor $\hat{\theta} = \hat{\theta}(\mathbf{Y}_s)$, as well as the estimating its prediction accuracy.

Algorithm 3. The parametric bootstrap algorithm for the linear mixed model, presented by Butar and Lahiri (2003, p. 67) and Rao and Molina (2015, pp. 183–186), is as follows:

1. Based on n sample observations, estimate a vector of model parameters Ψ and obtain its estimate $\hat{\Psi}$.

2. Generate B realisations $y_i^{*(b)}$ (where $b = 1, 2, \ldots, B$) of Y_i, under model (2), where parameters Ψ are replaced by $\hat{\Psi}$, to obtain a bootstrap realisation of the population vector $\mathbf{Y}$ denoted by $\mathbf{y}^{*(b)} = \left[y_1^{*(b)} \ldots y_i^{*(b)} \ldots y_N^{*(b)} \right]^T$, where $i = 1, 2, \ldots, N$ and $b = 1, 2, \ldots, B$.

3. Decompose vectors $\mathbf{y}^{*(b)}$, where $b = 1, 2, \ldots, B$, as follows $\left[\mathbf{y}_s^{*(b)T} \quad \mathbf{y}_r^{*(b)T} \right]^T$.

4. For $b = 1, 2, \ldots, B$ do

a) Compute the bootstrap realisation of the characteristic of interest θ denoted by $\theta^{*(b)} = \theta^{*(b)}\left(\mathbf{y}^{*(b)}, \hat{\Psi}\right)$.

b) Estimate Ψ based on generated data $\mathbf{y}_s^{*(b)}$ to obtain $\hat{\Psi}^{*(b)}$.

c) Compute the bootstrap realisation of $\hat{\theta}$ denoted by $\hat{\theta}^{*(b)}\left(\mathbf{y}_s^{*(b)}, \hat{\Psi}^{*(b)}\right)$ based on (2), where Ψ is replaced with $\hat{\Psi}^{*(b)}$.

d) Compute bootstrap realisations of prediction error U^* denoted by u^* and for the bth iteration given by:

$$u^{*(b)} = \hat{\theta}^{*(b)}\left(\mathbf{y}_s^{*(b)}, \hat{\Psi}^{*(b)}\right) - \theta^{*(b)}\left(\mathbf{y}^{*(b)}, \hat{\Psi}\right) = \hat{\theta}^{*(b)} - \theta^{*(b)}. \tag{10}$$

EndFor.

5. Estimate the prediction mean squared error:

$$\widehat{MSE}^{parametric}(\hat{\theta}) = B^{-1} \sum_{b=1}^{B} u^{*(b)^2}. \tag{11}$$

If the model assumptions are met, the bootstrap estimator (11) has good properties compared with other MSE estimators (e.g., Krzciuk, 2018).

5. Residual Bootstrap in Model-based Inference

Before the introduction of the algorithm, we introduce the correction procedure of predicted random effects and random components. The correction procedure aims to avoid the problem of the underdispersed bootstrap distributions of parameter estimates and downwardly biased variance parameter estimates (Carpenter, Goldstein & Rasbash, 2003, p. 435).

Without losing the generality, we can rewrite model (2) as follows:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_1\mathbf{v}_1 + \ldots + \mathbf{Z}_l\mathbf{v}_l + \ldots + \mathbf{Z}_L\mathbf{v}_L + \mathbf{e}, \tag{12}$$

where $\mathbf{v}_1, \ldots, \mathbf{v}_l, \ldots, \mathbf{v}_L$ are independent vectors for different divisions of $\mathbf{Y}$ vector,

$$\mathbf{v}_l = \left[\mathbf{v}_{l1}^T \ldots \mathbf{v}_{lk}^T \ldots \mathbf{v}_{lK_l}^T\right]^T \tag{13}$$

is of size $K_lJ_l \times 1$, where $\mathbf{v}_{lk}$ is $J_l \times 1$ for all $k = 1, \ldots, K_l$ and K_l is the number of random effects at lth level of grouping, $\mathbf{Z}_l$ is of size $N \times K_lJ_l$.

Firstly, the correction procedure of random components is presented based on Chambers and Chandra (2013, p. 455). It is assumed that $Var(\mathbf{e}) = \mathbf{R} = \sigma_e^2 diag_{1 \le i \le N}(a_i)$ with known weights a_i. The corrected estimates of random components are defined as follows:

$$\hat{e}_{(cor)i} = \hat{\sigma}_e \sqrt{a_i}\, \hat{e}_i \left(n^{-1} \sum_{k=1}^{n} \hat{e}_i\right)^{-0.5}, \tag{14}$$

where $i = 1, 2, \ldots, n$, $\hat{\sigma}_e^2$ is an estimate (e.g., REML estimate) of σ_e^2 and $\hat{e}_i$ are estimated random components computed under model (2).

Secondly, the random effects' correction procedure based on Thai *et al*. (2013, p. 132) and Carpenter, Goldstein and Rasbash (2003, pp. 435–436) is introduced. We consider vector (13) and its covariance matrix of size $K_l \times K_l$ defined as $\boldsymbol{\Sigma}_l = Var\left(\left[v_{l1j} \ldots v_{lkj} \ldots v_{lK_l j}\right]^T\right)$, where v_{lkj} is the jth element of $\mathbf{v}_{lk}$. The estimated matrix $\boldsymbol{\Sigma}_l$, obtained by replacing unknown variance components with their (e.g. Restricted Maximum Likelihood) estimates is denoted by $\hat{\boldsymbol{\Sigma}}_l$, while the empirical covariance matrix is given by $\boldsymbol{\Sigma}_{(emp)l} = J_l^{-1} \begin{bmatrix} \hat{\mathbf{v}}_{l1}^T \\ \cdots \\ \hat{\mathbf{v}}_{lk}^T \\ \cdots \\ \hat{\mathbf{v}}_{lK_l}^T \end{bmatrix} \begin{bmatrix} \hat{\mathbf{v}}_{l1}^T \\ \cdots \\ \hat{\mathbf{v}}_{lk}^T \\ \cdots \\ \hat{\mathbf{v}}_{lK_l}^T \end{bmatrix}^T$, where $\hat{\mathbf{v}}_{lk}$ are estimated best linear unbiased predictors of $\mathbf{v}_{lk}$. Using Cholesky decomposition, these covariance matrices can be written as $\hat{\boldsymbol{\Sigma}}_l = \mathbf{L}_{(est)l} \mathbf{L}_{(est)l}^T$ and $\boldsymbol{\Sigma}_{(emp)l} = \mathbf{L}_{(emp)l} \mathbf{L}_{(emp)l}^T$. Let $\mathbf{A}_l = \left(\mathbf{L}_{(est)l} \mathbf{L}_{(emp)l}^{-1}\right)^T$; the corrected version of the estimated best linear unbiased predictor $\hat{\mathbf{v}}_l$ is then defined as:

$$\hat{\mathbf{v}}_{(cor)l} = \hat{\mathbf{v}}_l \mathbf{A}_l. \tag{15}$$

To introduce the algorithm, *srswr* ($\mathbf{A}$, m) can be defined as the outcome of taking a random sample with replacement of size m of rows of matrix $\mathbf{A}$. In the special case, when $\mathbf{A}$ is a vector, *srswr* ($\mathbf{A}$, m) is a simple random sample with size m of elements of the vector replaced.

Algorithm 4. The residual bootstrap algorithm (with the correction procedure to avoid the problem of the underdispersed bootstrap distributions) presented by Carpenter, Goldstein and Rasbash (2003, pp. 435–436), Chambers and Chandra (2013, p. 455) and Thai *et al*. (2013, p. 132) for the LMM is given by Algorithm 3 of the parametric bootstrap, where step 2 is replaced by:

2. Generate B population vectors of the variable of interest, denoted by $\mathbf{y}^{*(b)}$ as:

$$\mathbf{y}^{*(b)} = \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{Z}_1 \mathbf{v}_1^{*(b)} + \ldots + \mathbf{Z}_l \mathbf{v}_l^{*(b)} + \ldots + \mathbf{Z}_L \mathbf{v}_L^{*(b)} + \mathbf{e}^{*(b)}, \tag{16}$$

where $\hat{\boldsymbol{\beta}}$ is an estimator (e.g. REML estimator) of $\boldsymbol{\beta}$, $\mathbf{e}^{*(b)}$ is a vector of size $N \times 1$ defined as $srswr\left(col_{1 \le i \le n} \hat{e}_{(cor)i}, N\right)$ with $\hat{e}_{(cor)i}$ given by (14), and $\mathbf{v}_l^{*(b)}$ (for 1, 2, …, L) is a vector of size $K_l J_l \times 1$ built from columns of the matrix: $srswr\left(\left[\hat{\mathbf{v}}_{(cor)l1} \ldots \hat{\mathbf{v}}_{(cor)lk} \ldots \hat{\mathbf{v}}_{(cor)lK_l}\right], J_l\right)$ of size $J_l \times K_l$ with $\hat{\mathbf{v}}_{(cor)l}$ given by (15).

Then, the prediction-MSE estimator has the same form it has in the parametric bootstrap – given by (11).

6. Double Bootstrap in Model-based Inference

The double bootstrap algorithm can be used to obtain bias-corrected MSE estimators. In this procedure, the parametric bootstrap procedure is used in both levels

of the double bootstrap. The first-level bootstrap prediction errors given by (10) are obtained as the result of the first level. The second level of the bootstrap is then performed within the first level bootstrap loop, as described below in Algorithm 5.

Algorithm 5. Algorithm 3 of the parametric bootstrap provides the first level of the double-bootstrap procedure. According to Hall and Maiti (2006) and Erciulescu and Fuller (2014), in the 4th step of Algorithm 3 we add the following point e) to obtain the double-bootstrap algorithm. For the *b*th iteration of the first level, we add the following second-level bootstrap loop:

e) For $c = 1, 2, \ldots, C$ do

– Under model (2), where parameters Ψ are replaced by first level estimates $\hat{\Psi}^{*(b)}$, generate second-level realisations of Y_i denoted by $y_i^{**(b,c)}$, where $i = 1, 2, \ldots, N$ to obtain $\mathbf{y}^{**(b,c)} = \left[y_1^{**(b,c)} \ldots y_i^{**(b,c)} \ldots y_N^{**(b,c)}\right]^T$, which is decomposed as follows: $\mathbf{y}^{**(b,c)} = \left[\mathbf{y}_s^{**(b,c)T} \ \mathbf{y}_r^{**(b,c)T}\right]^T$.

– Compute the second-level bootstrap realisation of θ denoted by $\theta^{**(b,c)} = \theta^{**(b,c)}\left(\mathbf{y}^{**(b,c)}, \hat{\Psi}^{*(b)}\right)$.

– Use $\mathbf{y}_s^{**(b,c)}$ to obtain the vector of estimates $\hat{\Psi}^{**(b,c)}$, and compute second-level bootstrap realisation of $\hat{\theta}$, denoted by $\hat{\theta}^{**(b,c)}\left(\mathbf{y}_s^{**(b,c)}, \hat{\Psi}^{**(b,c)}\right)$.

– Compute the second-level bootstrap realisation of the prediction error given by:

$$\begin{aligned} u^{**(b,c)} &= \\ &= \hat{\theta}^{**(b,c)}\left(\mathbf{y}_s^{**(b,c)}, \hat{\Psi}^{**(b,c)}\right) - \theta^{**(b,c)}\left(\mathbf{y}^{**(b,c)}, \hat{\Psi}^{*(b)}\right) = \\ &= \hat{\theta}^{**(b,c)} - \theta^{**(b,c)}. \end{aligned} \tag{17}$$

EndFor.

Furthermore, in step 5 of Algorithm 3, we introduce additional MSE estimators. The first one is the classic bias-corrected double bootstrap MSE estimator (Erciulescu & Fuller, 2014, p. 12):

$$\widehat{MSE}^{db-B2}(\hat{\theta}) = 2\widehat{MSE}^{pboot} - \widehat{MSE}^{db-2lev} = B^{-1}\sum_{b=1}^{B}\tilde{u}_1^{**(b)}, \tag{18}$$

where $\widehat{MSE}^{pboot}$ is given by (11), $\widehat{MSE}^{db-2lev} = B^{-1}C^{-1}\sum_{b=1}^{B}\sum_{c=1}^{C}u^{**(b,c)^2}$,

$$\tilde{u}_1^{**(b)} = 2u^{*(b)^2} - C^{-1}\sum_{c=1}^{C}u^{**(b,c)^2} \tag{19}$$

and $u^{*(b)}$ and $u^{**(b,c)}$ are given by (10) and (17), respectively. The special case of (18) for $C = 1$ is as follows:

$$\widehat{MSE}^{db-1}(\hat{\theta}) = B^{-1}\sum_{b=1}^{B}\tilde{u}_2^{**(b)}, \tag{20}$$

where

$$\tilde{u}_2^{**(b)} = 2u^{*(b)^2} - u^{**(b,c)^2}. \tag{21}$$

The following MSE estimator, called the telescoping bootstrap MSE estimator, is a modification of (20) (Erciulescu & Fuller, 2014, p. 12):

$$\widehat{MSE}^{db-telesc}(\hat{\theta}) = B^{-1}\sum_{b=1}^{B}\tilde{u}_3^{**(b)}, \tag{22}$$

where

$$\tilde{u}_3^{**(b)} = u^{*(b)^2} + u^{*(b+1)^2} - u^{**(b,c)^2}. \tag{23}$$

It should be noted that one additional iteration of the first-level bootstrap should be performed using (22) in comparison to (20).

Because simulation studies showed that the properties of the above double bootstrap MSE estimators can suffer due to high bias corrections, leading even to the negative values of the MSE estimators, modifications were proposed. Hall and Maiti (2006, p. 228) provided the first:

$$\widehat{MSE}^{db-B2-HM}(\hat{\theta}) = \\ = \begin{cases} 2\widehat{MSE}^{pboot} - \widehat{MSE}^{db-2lev} & \text{if } \widehat{MSE}^{pboot} \geq \widehat{MSE}^{db-2lev} \\ \widehat{MSE}^{pboot}\exp\left[\dfrac{\widehat{MSE}^{pboot} - \widehat{MSE}^{db-2lev}}{\widehat{MSE}^{db-2lev}}\right] & \text{if } \widehat{MSE}^{pboot} < \widehat{MSE}^{db-2lev} \end{cases} \tag{24}$$

Erciulescu and Fuller (2014, p. 3311) provided the second:

$$\widehat{MSE}^{db-1-EF}(\hat{\theta}) = \\ = \begin{cases} q \times \widehat{MSE}^{pboot} & \text{if } (\widehat{MSE}^{pboot})^{-1} B^{-1}\sum_{b=1}^{B} u^{**(b,c)^2} < q \\ \widehat{MSE}^{db-1} & \text{otherwise} \end{cases} \tag{25}$$

where the authors' proposal is to set $q = 0.77$ and $C = 1$. The third modification, itself a modification of (22), can be written as follows (Erciulescu & Fuller, 2014, p. 3312):

$$\widehat{MSE}^{db-telesc-EF}(\hat{\theta}) = \\ = \begin{cases} \widehat{MSE}^{pboot} & \text{if } (\widehat{MSE}^{pboot})^{-1} B^{-1}\sum_{b=1}^{B} u^{**(b,c)^2} < q \\ \widehat{MSE}^{db-telesc} & \text{otherwise} \end{cases} \tag{26}$$

where the authors' proposal is to set $q = 0.77$ and $C = 1$.

Estimators (24), (25) and (26) are modifications of (18), (20) and (22), proposed to avoid negative MSE estimates. We propose to solve this problem differently: not by correcting the final results, but by correcting potentially negative elements $\tilde{u}_1^{**(b)}$, $\tilde{u}_2^{**(b)}$ and $\tilde{u}_3^{**(b)}$, given by (19), (21) and (23), functioning as squared predictions errors. Here we define the following modifications of (19), (21) and (23) to avoid possible negative values:

$$\tilde{u}_{h\,mod}^{**(b)} = \begin{cases} \tilde{u}_h^{**(b)} & \text{if } \tilde{u}_h^{**(b)} \geq 0 \\ u^{*(b)^2} & \text{if } \tilde{u}_h^{**(b)} < 0 \end{cases}, \; h = 1, 2, 3 \tag{27}$$

This leads to the following proposals of modifications of (18), (20) and (22):

$$\widehat{MSE}^{db-B2-mod}(\hat{\theta}) = B^{-1} \sum_{b=1}^{B} \tilde{u}_{1\,mod}^{**(b)}, \tag{28}$$

$$\widehat{MSE}^{db-1-mod}(\hat{\theta}) = B^{-1} \sum_{b=1}^{B} \tilde{u}_{2\,mod}^{**(b)}, \tag{29}$$

$$\widehat{MSE}^{db-telesc-mod}(\hat{\theta}) = B^{-1} \sum_{b=1}^{B} \tilde{u}_{3\,mod}^{**(b)}, \tag{30}$$

where $\tilde{u}_{h\,mod}^{**(b)}$, $h = 1, 2, 3$ are given by (27) for $h = 1, 2, 3$.

7. Applications – Part 1

We consider a real dataset from Särndal, Swensson and Wretman (1992) on 281 Swedish municipalities, which is available in the *sampling* R package (Tillé & Matei, 2021). The variable of interest is the revenue from municipal taxation in millions of kronor in 1985 (denoted by RMT85). The auxiliary variable is the population in 1975 in thousands of people (denoted by P75). It is used in this section to compute first-order inclusion probabilities for probability-proportional-to-size sampling using the Brewer sampling scheme (e.g., Tillé, 2006, p. 114). The estimation of the population total is considered, but the estimation of the design-accuracy of other population parameters, as presented by e.g., Antal and Tillé (2014) and Stachurski (2018), can also be analysed. Using the sampling scheme mentioned above, one sample of a size equal to of the population size $N = 281$ is drawn. The value of the Horvitz-Thompson (HT) estimator of the population total of municipal taxation revenues is computed based on the sample (53,861.61 Swedish kronor).

We estimate the design-standard error using various bootstrap algorithms available in *pipsboot* R package (Kucharski & Żądło, 2021). The number of iterations set for a bootstrap algorithm should be high enough to ensure stable results – some proposals have been put forward by Chwila and Żądło (2020). We use 1,000 iterations. The following algorithms are considered:

– BM0.5 – Barbiero and Mecatti (2010) 0.5 bootstrap,
– AT2011 – Antal and Tillé (2011) direct bootstrap,
– AT2014 – Antal and Tillé (2014) direct bootstrap,
– BMM – Barbiero, Manzi and Mecatti (2015) bootstrap,
– H – Holmberg (1988) bootstrap,
– Z – Żądło (2021) bootstrap,
– Q – Quatember (2014) bootstrap,

– SP – Sverchkov and Pfeffermann (2004) bootstrap,
– BMxb1 – Barbiero and Mecatti (2010) x-balanced 1 bootstrap,
– BMxb2 – Barbiero and Mecatti (2010) x-balanced 2 bootstrap.

Table 1. The Estimated Design-precision of the HT Estimator of the Population Total

Bootstrap	Estimated Design-standard Error	Estimated Relative Design-standard Error (in %)	Execution Time of 1,000 Iterations (in sec.)
BM0.5	1,338.6	2.49	0.019
AT2011	1,347.8	2.50	0.018
AT2014	1,352.4	2.51	0.001
BMM	1,294.4	2.40	0.052
H	1,327.7	2.47	0.050
Z	1,340.0	2.49	0.007
Q	1,380.2	2.56	0.001
SP	1,357.8	2.52	0.033
BMxb1	1,331.8	2.47	0.018
BMxb2	1,370.1	2.54	0.019

Source: the author's own computations using *pipsboot* R package (Kucharski & Żądło, 2021).

The estimated relative design-standard errors presented in Table 1 are similar to each other. This suggests that they have similar stochastic properties, though that should be confirmed in the simulation studies. In such a case, the time--consuming algorithms can be of lower interest, especially for large populations, where bootstrap algorithms based on the plug-in approach require the physical creation of a pseudo-population. Therefore, faster algorithms including Żądło (2021) may be preferable.

8. Applications – Part 2

In this section we use the same dataset as in the previous one. In this case, the inference is based on a single stratified sample of a size 20% of the population size ($N = 281$) to have enough sample observations to estimate all model parameters. The population total is predicted under the special case of model (2), where RMT85 logarithms are explained by P75 logarithms (the variables are described in the previous section). This model is the nested error linear mixed model with subpopulation-specific random effects, with subpopulations defined as regions (e.g., Żądło, 2015, p. 43). The model parameters are significant at a 0.05 significance level (the results are based on permutation tests). These tests are known to have good properties both for testing fixed effects (Krzciuk & Żądło, 2014a) and random

components (Krzciuk & Żądło, 2014b). What is more, there is no evidence to reject the null hypothesis on the normality of random effect and random components (the test procedure proposed by Jacqmin-Gadda *et al*. (2007) has been conducted).

The plug-in predictor is used here to predict the population total under the model being considered (see e.g., Chwila & Żądło, 2022, p. 20). However, its potential application is greater – it allows one to predict any given function of the population vector of the variable of interest, such as quantiles and distribution functions (e.g., Stachurski, 2021). Based on the sample under consideration, the value of the predictor of the population total of municipal taxation revenues is computed (it comes out to 54,022.03 Swedish kronor). Although we consider the use of bootstrap for linear mixed models, it also has successful applications for more general models, as shown e.g., by Wolny-Dominiak (2017) and Wolny-Dominiak and Żądło (2022a).

To estimate the prediction-accuracy, we use different bootstrap algorithms available in *qape* R package (Wolny-Dominiak & Żądło, 2022b). The assumed number of iterations is 1,000 for parametric, residual, and the first level of the double bootstrap, and 500 for the second level of the double bootstrap.

In sections 4, 5 and 6, various bootstrap MSE estimators were presented. The superscripts used in their notations are used as the names of bootstrap algorithms in Table 2, along with the appropriate equation numbers.

Table 2. Estimated Prediction Accuracy of the Plug-in Predictor of the Population Total

Bootstrap	Estimated Prediction-RMSE	Estimated Relative Prediction-RMSE (in %)	Execution Time[a] (in sec.)
Parametric – equation (11)	725.5	1.35	23.17
Residual – as in Algorithm 4	716.8	1.34	36.01
db-B2 – equation (18)	651.6	1.22	11,516[b]
db-B2-mod – equation (28)	811.8	1.51	
db-B2-HM – equation (24)	655.3	1.22	
db-1 – equation (20)	633.9	1.18	
db-1-mod – equation (29)	854.1	1.60	
db-1-EF – equation (25)	633.9	1.18	
db-telesc – equation (22)	634.1	1.18	
db-telesc-mod – equation (30)	830.8	1.55	
db-telesc-EF – equation (22)	634.1	1.18	

[a] Of 1,000 iterations for parametric, residual, and the first level of the double bootstrap, and 500 for the second level of the double bootstrap (parallel computing is used). [b] The time of the execution of doubleBootFuture{qape}function which computes all double bootstrap MSE estimators.

Source: the author's own computations using *qape* R package (Wolny-Dominiak & Żądło, 2022b).

Firstly, the results of db-1 and its modified version db-1-EF MSE estimators are identical, as are the results of db-telesc and its modified version db-telesc-EF. This means that the conditions implying possible corrections in (25) and (26) have not been met. Secondly, the results of the parametric and residual MSE estimators are as follows. The first one is based on the normality assumptions of random effects and random components, while the second is not. However, the normality assumption is met, and the results are comparable. Thirdly, the results of all double bootstrap estimators (except our proposals) are similar and smaller than the value of the parametric bootstrap estimators, since they are bias-corrected versions of the parametric bootstrap estimator. Finally, the values of our proposals given by (28), (29), and (30) are higher than the rest of the results, which may indicate that they are positively biased. Even if they overestimate the prediction MSE on average, they may be useful proposals because they provide pessimistic estimates of prediction accuracy. However, further Monte Carlo simulation analyses in this area should be conducted.

9. Conclusion

We have presented a variety of bootstrap algorithms used in survey sampling to estimate design- and prediction-accuracy. I use their computer implementations, available as two R packages on R CRAN and GitHub. Three double bootstrap MSE estimators are proposed and analysed in the application of the real data. Most of the application results within each approach are similar, suggesting that the stochastic properties of variance and MSE estimators are similar. However, they cannot be generalised for other population datasets directly, though the R packages *pipsboot* and *qape* make it easy to use these methods for any dataset. Further research should include Monte Carlo simulation studies, in which the biases and accuracies of variance and MSE estimators should be analysed. What is more, design-based considerations can be extended to other sampling designs, and the model-based results to other models, including the problem of model misspecification. The analyses could also be extended to consider the estimation and prediction of other characteristics than population total.

Conflict of Interest

The author declares no conflict of interest.

References

Antal, E., & Tillé, Y. (2011). A Direct Bootstrap Method for Complex Sampling Designs from a Finite Population. *Journal of the American Statistical Association*, *106*(494), 534–543. https://doi.org/10.1198/jasa.2011.tm09767

Antal, E., & Tillé, Y. (2014). A New Resampling Method for Sampling Designs without Replacement: The Doubled Half Bootstrap. *Computational Statistics*, *29*(5), 1345–1363. https://doi.org/10.1007/s00180-014-0495-0

Barbiero, A., Manzi, G., & Mecatti, F. (2015). Bootstrapping Probability-proportional--to-size Samples via Calibrated Empirical Population. *Journal of Statistical Computation and Simulation*, *85*(3), 608–620. https://doi.org/10.1080/00949655.2013.833204

Barbiero, A., & Mecatti, F. (2010). Bootstrap Algorithms for Variance Estimation in πPS Sampling. In: P. Mantovan, P. Secchi (Eds), *Complex Data Modeling and Computationally Intensive Statistical Methods*. *Contribution to Statistics* (pp. 57–69). Springer. https://doi.org/10.1007/978-88-470-1386-5_5

Beaumont, J.-F., & Patak, Z. (2012). On the Generalized Bootstrap for Sample Surveys with Special Attention to Poisson Sampling. *International Statistical Review*, *80*(1), 127–148. https://doi.org/10.1111/j.1751-5823.2011.00166.x

Butar, F. B., & Lahiri, P. (2003). On Measures of Uncertainty of Empirical Bayes Small--area Estimators. *Journal of Statistical Planning and Inference*, *112*(1–2), 63–76. https://doi.org/10.1016/S0378-3758(02)00323-3

Carpenter, J. R., Goldstein, H., & Rasbash, J. (2003). A Novel Bootstrap Procedure for Assessing the Relationship between Class Size and Achievement. *Journal of the Royal Statistical Society: Series C* (*Applied Statistics*), *52*(4), 431–443. https://doi.org/10.1111/1467-9876.00415

Cassel, C. M., Särndal, C.-E., & Wretman, J. H. (1977). *Foundations of Inference in Survey Sampling*. Wiley-Interscience.

Chambers, R., & Chandra, H. (2013). A Random Effect Block Bootstrap for Clustered Data. *Journal of Computational and Graphical Statistics*, *22*(2), 452–470. https://doi.org/10.1080/10618600.2012.681216

Chwila, A., & Żądło, T. (2020). On the Choice of the Number of Monte Carlo Iterations and Bootstrap Replicates in Empirical Best Prediction. *Statistics in Transition New Series*, *21*(2), 35–60. https://doi.org/10.21307/stattrans-2020-013

Chwila, A., & Żądło, T. (2022). On Properties of Empirical Best Predictors. *Communications in Statistics – Simulation and Computation*, *51*(1), 220–253. https://doi.org/10.1080/03610918.2019.1649422

Deville, J.-C., & Särndal, C.-E. (1992). Calibration Estimators in Survey Sampling. *Journal of the American Statistical Association*, *87*(418), 376–382. https://doi.org/10.1080/01621459.1992.10475217

Efron, B. (1979). Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*, *7*(1), 1–26. https://doi.org/10.1214/aos/1176344552

Erciulescu, A. L., & Fuller, W. A. (2014). Parametric Bootstrap Procedures for Small Area Prediction Variance. In: *JSM 2014 – Survey Research Methods Section* (pp. 3307–3318). American Statistical Association.

Hall, P., & Maiti, T. (2006). On Parametric Bootstrap Methods for Small Area Prediction. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *68*(2), 221–238. https://doi.org/10.1111/j.1467-9868.2006.00541.x

Holmberg, A. (1988). A Bootstrap Approach to Probability Proportional-to-size Sampling. In: *Proceedings of Section on Survey Research Methods* (pp. 378–383). American Statistical Association.

Horvitz, D. G., & Thompson, D. J. (1952). A Generalization of Sampling without Replacement from a Finite Universe. *Journal of the American Statistical Association*, *47*(260), 663–685. https://doi.org/10.1080/01621459.1952.10483446

Jacqmin-Gadda, H., Sibillot, S., Proust, C., Molina, J.-M., & Thiébaut, R. (2007). Robustness of the Linear Mixed Model to Misspecified Error Distribution. *Computational Statistics & Data Analysis*, *51*(10), 5142–5154. https://doi.org/10.1016/j.csda.2006.05.021

Krzciuk, M., & Żądło, T. (2014a). On Some Tests of Fixed Effects for Linear Mixed Models. *Studia Ekonomiczne*, 189, 49–57.

Krzciuk, M., & Żądło, T. (2014b). On Some Tests of Variance Components for Linear Mixed Models. *Studia Ekonomiczne*, 189, 77–85.

Krzciuk, M. K. (2018). On the Simulation Study of Jackknife and Bootstrap MSE Estimators of a Domain Mean Predictor for Fay-Herriot Model. *Acta Universitatis Lodziensis. Folia Oeconomica*, *5*(331), 169–183. https://doi.org/10.18778/0208-6018.331.11

Kucharski, R., & Żądło, T. (2021). *pipsboot: Bootstrap for Probability Proportional to Size Sampling, R package*. Retrieved from: https://github.com/kucharsky/pipsboot (accessed: 29.03.2024).

Quatember, A. (2014). The Finite Population Bootstrap – from the Maximum Likelihood to the Horvitz-Thompson Approach. *Austrian Journal of Statistics*, *43*(2), 93–102. https://doi.org/10.17713/ajs.v43i2.10

Ranalli, M. G., & Mecatti, F. (2012). Comparing Recent Approaches for Bootstrapping Sample Survey Data: A First Step towards a Unified Approach. In: *Section on Survey Research Methods – JSM 2012* (pp. 4088–4099). American Statistical Association.

Rao, J. N. K., & Molina, I. (2015). *Small Area Estimation* (2nd ed.). John Wiley & Sons.

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. https://doi.org/10.1080/01621459.1988.10478591

Särndal, C.-E., Swensson, B., & Wretman, J. (1992). *Model Assisted Survey Sampling*. Springer.

Sen, A. R. (1953). On the Estimate of the Variance in Sampling with Varying Probabilities. *Journal of the Indian Society of Agricultural Statistics*, *5*(1194).

Sitter, R. R. (1992). Comparing Three Bootstrap Methods for Survey Data. *Canadian Journal of Statistics*, *20*(2), 135–154. https://doi.org/10.2307/3315464

Stachurski, T. (2018). A Simulation Analysis of the Accuracy of Median Estimators for Different Sampling Designs. In: L. Váchová, V. Kratochvíl (Eds), *Proceedings of the 36th*

International Conference Mathematical Methods in Economics MME 2018 (pp. 509–514). MatfyzPress, Publishing House of the Faculty of Mathematics and Physics Charles University.

Stachurski, T. (2021). Small Area Quantile Estimation Based on Distribution Function Using Linear Mixed Models. *Economics and Business Review*, *7*(2), 97–114. https://doi.org/10.18559/ebr.2021.2.7

Sverchkov, M., & Pfeffermann, D. (2004). Prediction of Finite Population Totals Based on the Sample Distribution. *Survey Methodology*, *30*(1), 79–92.

Thai, H.-T., Mentré, F., Holford, N. H. G., Veyrat-Follet, C., & Comets, E. (2013). A Comparison of Bootstrap Approaches for Estimating Uncertainty of Parameters in Linear Mixed-effects Models. *Pharmaceutical Statistics*, *12*(3), 129–140. https://doi.org/10.1002/pst.1561

Tillé, Y. (2006). *Sampling Algorithms*. Springer.

Tillé, Y., & Matei, A. (2021). *sampling: Survey Sampling, R package*. Retrieved from: https://CRAN.R-project.org/package=sampling (accessed: 29.03.2024).

Wolny-Dominiak, A. (2017). Bootstrap Mean Squared Error of Prediction in Loss Reserving. In: K. Jajuga, L. T. Orlowski, K. Staehr (Eds), *Contemporary Trends and Challenges in Finance*. *Springer Proceedings in Business and Economics* (pp. 213–220). Springer International Publishing. https://doi.org/10.1007/978-3-319-54885-2_20

Wolny-Dominiak, A., & Żądło, T. (2022a). On Bootstrap Estimators of Some Prediction Accuracy Measures of Loss Reserves in a Non-life Insurance Company. *Communications in Statistics – Simulation and Computation*, *51*(8), 4225–4240. https://doi.org/10.1080/03610918.2020.1740263

Wolny-Dominiak, A., & Żądło, T. (2022b). *qape: Quantile of Absolute Prediction Errors, R package*. Retrieved from: https://CRAN.R-project.org/package=qape (accessed: 29.03.2024).

Yates, F., & Grundy, P. M. (1953). Selection without Replacement from within Strata with Probability Proportional to Size. *Journal of the Royal Statistical Society: Series B* (*Methodological*), *15*(2), 253–261. https://doi.org/10.1111/j.2517-6161.1953.tb00140.x

Żądło, T. (2015). *Statystyka małych obszarów w badaniach ekonomicznych*. *Podejście modelowe i mieszane*. Wydawnictwo Uniwersytetu Ekonomicznego w Katowicach.

Żądło, T. (2021). On the Generalisation of Quatember's Bootstrap. *Statistics in Transition New Series*, *22*(1), 163–178. https://doi.org/10.21307/stattrans-2021-009