

**ON MSE ESTIMATORS OF EBLUP
OF DOMAIN TOTAL
UNDER SOME LONGITUDINAL MODEL**

Tomasz Żądło

Abstract. Żądło (2012) proposed a certain unit-level longitudinal model which was a special case of the General Linear Mixed Model. Two vectors of random components included in the model obey assumptions of simultaneous spatial autoregressive process (SAR) and temporal first-order autoregressive process (AR(1)) respectively. Moreover, it is assumed that the population can change in time and the population elements can change its domains' (subpopulations') affiliation in time. Under the proposed model, Żądło (2012) derived the Empirical Best Linear Unbiased Predictor (EBLUP) of the domain total. What is more (based on the theorem proved by Żądło (2009)), the approximate equation of the mean squared error (MSE) was derived and its estimator based on the Taylor approximation was proposed. The proposed MSE estimator was derived under some assumptions including that the variance-covariance matrix can be decomposed into linear combination of variance components. The assumption was not met under the proposed model. In the paper the jackknife MSE estimator for the derived EBLUP will be proposed based on the results presented by Jiang, Lahiri, Wan (2002). The bias of the jackknife MSE estimator will be compared in the simulation study with the bias of the MSE estimator based on the Taylor approximation.

Keywords: small area estimation, MSE estimation, jackknife.

JEL Classification: C42.

DOI: 10.15611/me.2013.9.09.

1. Superpopulation model

Let us introduce the notation presented earlier by Żądło (2012). We assume that longitudinal data from $t = 1, \dots, M$ periods are studied. In the period t the population is denoted by Ω_t ($\bar{\bar{\Omega}}_t = N_t$). Let $\bigcup_{t=1}^M \Omega_t = \Omega$ and $\bar{\bar{\Omega}} = N$. The population in the period t is divided into D disjoint subpopula-

Tomasz Żądło

Department of Statistics, University of Economics in Katowice, 1 Maja Street 50, 40-287 Katowice, Poland.

E-mail: tomasz.zadlo@ue.katowice.pl

tions (domains) Ω_{dt} ($\bar{\bar{\Omega}}_{dt} = N_{dt}$), where $d = 1, \dots, D$. Let $\bigcup_{t=1}^M \Omega_{dt} = \Omega_d$ and $\bar{\bar{\Omega}}_d = N_d$. Let the set of population elements for which observations are available in the period t be denoted by s_t ($\bar{\bar{s}}_t = n_t$). Let $\bigcup_{t=1}^M s_t = s$ and $\bar{\bar{s}} = n$. Moreover, let $\Omega_{dt} \cap s_t = s_{dt}$, $\bar{\bar{s}}_{dt} = n_{dt}$ and $\bigcup_{t=1}^M s_{dt} = s_d$, $\bar{\bar{s}}_d = n_d$. Let: $\Omega_{rdt} = \Omega_{dt} - s_{dt}$, $N_{rdt} = N_{dt} - n_{dt}$. The d^* -th domain of interest in the period of interest t^* will be denoted by $\Omega_{d^*t^*}$ and $\bigcup_{t=1}^M \Omega_{d^*t} = \Omega_{d^*}$. Let M_{id} be the number of periods when element i is in the domain d , and m_{id} be the number of periods when element i is observed in the domain d . Let $\mathbf{Y}_{id} = [Y_{idj}]$, where $j = 1, \dots, M_{id}$, be the vector of size $M_{id} \times 1$ of random variables Y_{idj} for the i -th population element which belongs to the d -th domain. Let $\mathbf{Y}_{sid} = [Y_{sidj}^*]$, where $j^* = 1, \dots, m_{id}$, be a subvector of \mathbf{Y}_{id} for m_{id} observed in the sample realizations of random variables Y_{sidj}^* .

Let us assume that population longitudinal data obey the following model (proposed by Żądło (2012)):

$$\mathbf{Y}_d = \mathbf{X}_d \boldsymbol{\beta}_d + \mathbf{Z}_d \mathbf{v}_d + \mathbf{e}_d, \quad (1)$$

where $\mathbf{Y}_d = \text{col}_{1 \leq i \leq N_d}(\mathbf{Y}_{id})$, where \mathbf{Y}_{id} is a random vector, called profile, of size $M_{id} \times 1$, and \mathbf{Y}_d ($d = 1, \dots, D$) are assumed to be independent, $\mathbf{X}_d = \text{col}_{1 \leq i \leq N_d}(\mathbf{X}_{id})$, where \mathbf{X}_{id} is the known matrix of size $M_{id} \times p$, $\mathbf{Z}_d = \text{diag}_{1 \leq i \leq N_d}(\mathbf{Z}_{id})$, where \mathbf{Z}_{id} is the known vector of size $M_{id} \times 1$, vectors of random components \mathbf{v}_d and \mathbf{e}_d are assumed to be independent.

Let $\mathbf{v}_d = \text{col}_{1 \leq i \leq N_d}(v_{id})$, where v_{id} is a profile-specific random component and \mathbf{v}_d ($d = 1, 2, \dots, D$) are assumed to be independent and that they obey the assumptions of the simultaneously spatial autoregressive (SAR) process:

$$\mathbf{v}_d = \rho_{(sp)} \mathbf{W}_d \mathbf{v}_d + \mathbf{u}_d, \quad (2)$$

where \mathbf{W}_d is the spatial weight matrix for profiles \mathbf{Y}_{id} , $\mathbf{u}_d \sim (\mathbf{0}, \sigma_u^2 \mathbf{I}_{N_d})$. Usually it is defined as row-standardized matrix of values of some distance function between population elements or between subpopulations (in this case between profiles). Hence,

$$\mathbf{v}_d \sim (\mathbf{0}, \mathbf{R}_d), \quad (3)$$

where $\mathbf{R}_d = \sigma_u^2 \mathbf{C}_d^{-1}$ and $\mathbf{C}_d = (\mathbf{I}_{N_d} - \rho_{(sp)} \mathbf{W}_d)(\mathbf{I}_{N_d} - \rho_{(sp)} \mathbf{W}_d^T)$.

Moreover, $\mathbf{e}_d = col_{1 \leq i \leq N_d}(\mathbf{e}_{id})$, where \mathbf{e}_{id} is a random component vector of size $M_{id} \times 1$ and \mathbf{e}_{id} ($i = 1, \dots, N; d = 1, \dots, D$) are assumed to be independent and their elements obey assumptions of autoregressive process AR(1):

$$e_{idj} = \rho_{(t)} e_{idj-1} + \varepsilon_{idj}. \quad (4)$$

Hence,

$$e_{id} \sim (\mathbf{0}, \Sigma_{id}), \quad (5)$$

where elements of Σ_{id} are given by $\sigma_\varepsilon^2 \rho_{(t)}^{|k-l|} (1 - \rho_{(t)}^2)^{-1}$.

2. Predictor

Under the model (1), based on the theorem presented by Royall (1976), the best linear unbiased predictor is given by (as derived by Żądło (2012)):

$$\hat{\theta}_{d^{*t^*}}^{BLU} = \sum_{i \in S_{d^{*t^*}}} Y_{id^{*t^*}} + \tilde{\mathbf{x}}_{rd^{*t^*}} \hat{\boldsymbol{\beta}}_{d^*} + \boldsymbol{\gamma}_{rd^*}^T \left(\sigma_u^2 \mathbf{Z}_{rd^*} \mathbf{C}_d^{-1} \mathbf{Z}_{sd^*}^T + \text{diag}_{1 \leq i \leq N_{rd^*}}(\Sigma_{rs id^*}) \right) \mathbf{V}_{ss d^*}^{-1} \left(\mathbf{Y}_{sd^*} - \mathbf{X}_{sd^*} \hat{\boldsymbol{\beta}}_{d^*} \right), \quad (6)$$

where $\tilde{\mathbf{x}}_{rd^{*t^*}}$ is a $1 \times p$ vector of totals of auxiliary variables in $\Omega_{rd^{*t^*}}$,

$$\hat{\boldsymbol{\beta}}_{d^*} = \left(\mathbf{X}_{sd^*}^T \mathbf{V}_{ss d^*}^{-1} \mathbf{X}_{sd^*} \right)^{-1} \mathbf{X}_{sd^*}^T \mathbf{V}_{ss d^*}^{-1} \mathbf{Y}_{sd^*},$$

$$\mathbf{V}_{ss d^*}^{-1} = \left(\sigma_u^2 \mathbf{Z}_{sd^*} \mathbf{C}_d^{-1} \mathbf{Z}_{sd^*}^T + \text{diag}_{1 \leq i \leq N_{d^*}}(\Sigma_{ss id^*}) \right)^{-1},$$

\mathbf{X}_{sd^*} is known $\sum_{i=1}^{n_{d^*}} m_{id^*} \times p$ matrix of auxiliary variables, \mathbf{Y}_{sd^*} is a $\sum_{i=1}^{n_{d^*}} m_{id^*} \times 1$ vector of random variables Y_{idj} , $\boldsymbol{\gamma}_{rd^*}$ is a $\sum_{i=1}^{n_{d^*}} M_{rid^*} \times 1$ vector of one's for observations in period t^* (in $\Omega_{rd^*t^*}$) and zero otherwise, \mathbf{Z}_{sd} and \mathbf{Z}_{rd} are submatrices of \mathbf{Z}_d obtained by deleting rows for unsampled and sampled elements respectively, $\boldsymbol{\Sigma}_{ssid}$ is a submatrix obtained from $\boldsymbol{\Sigma}_{id}$ by deleting rows and columns for unsampled observations, where $\boldsymbol{\Sigma}_{rsid}$ is a submatrix obtained from $\boldsymbol{\Sigma}_{id}$ by deleting rows for sampled observations and columns for unsampled observations.

Let the unknown variance parameters in (6) be replaced by restricted maximum likelihood (REML) estimates under normality. Hence, the two-stage predictor (denoted by $\hat{\theta}_{EBLU}(\hat{\boldsymbol{\delta}})$) called EBLUP is obtained. It remains unbiased under some weak assumptions (inter alia symmetric but not necessarily normal distribution of random components for the model assumed for the whole population). The proof is presented by Żądło (2004). Using the theorem presented by Żądło (2009) the MSE estimator of the EBLUP based on Taylor's expansion is given by (the result was obtained by Żądło (2012)):

$$M\hat{S}E_{\xi}^{Taylor}(\hat{\theta}_{EBLU}(\hat{\boldsymbol{\delta}})) = g_1(\hat{\boldsymbol{\delta}}) + g_2(\hat{\boldsymbol{\delta}}) + 2g_3^*(\hat{\boldsymbol{\delta}}) \quad (7)$$

where

$$g_1(\boldsymbol{\delta}) = \boldsymbol{\gamma}_{rd^*}^T \left(\sigma_u^2 \mathbf{Z}_{rd^*} \mathbf{C}_d^{-1} \mathbf{Z}_{rd^*}^T + \text{diag}_{1 \leq i \leq N_{rd^*}}(\boldsymbol{\Sigma}_{rrid^*}) \right) \boldsymbol{\gamma}_{rd^*} + \\ - \boldsymbol{\gamma}_{rd^*}^T \left(\sigma_u^2 \mathbf{Z}_{rd^*} \mathbf{C}_d^{-1} \mathbf{Z}_{sd^*}^T + \text{diag}_{1 \leq i \leq N_{rd^*}}(\boldsymbol{\Sigma}_{rsid^*}) \right) \quad (8)$$

$$\mathbf{V}_{ssd^*}^{-1} \left(\sigma_u^2 \mathbf{Z}_{rd^*} \mathbf{C}_d^{-1} \mathbf{Z}_{sd^*}^T + \text{diag}_{1 \leq i \leq N_{rd^*}}(\boldsymbol{\Sigma}_{rsid^*}) \right)^T \boldsymbol{\gamma}_{rd^*}$$

$$g_2(\boldsymbol{\delta}) = \boldsymbol{\gamma}_{rd^*}^T \left(\mathbf{X}_{rd^*} - \left(\sigma_u^2 \mathbf{Z}_{rd^*} \mathbf{C}_d^{-1} \mathbf{Z}_{sd^*}^T + \text{diag}_{1 \leq i \leq N_{rd^*}}(\boldsymbol{\Sigma}_{rsid^*}) \right) \mathbf{V}_{ssd^*}^{-1} \mathbf{X}_{sd^*} \right) \\ \times \left(\mathbf{X}_{sd^*}^T \mathbf{V}_{ssd^*}^{-1} \mathbf{X}_{sd^*} \right)^{-1} \left(\mathbf{X}_{rd^*} - \left(\sigma_u^2 \mathbf{Z}_{rd^*} \mathbf{C}_d^{-1} \mathbf{Z}_{sd^*}^T + \text{diag}_{1 \leq i \leq N_{rd^*}}(\boldsymbol{\Sigma}_{rsid^*}) \right) \mathbf{V}_{ssd^*}^{-1} \mathbf{X}_{sd^*} \right)^T \boldsymbol{\gamma}_{rd^*} \quad (9)$$

$$g_3^*(\boldsymbol{\delta}) = tr \left(\frac{\partial \mathbf{c}_{\mathbf{d}^*}^T}{\partial \boldsymbol{\delta}} \mathbf{V}_{\mathbf{ss d}^*} \left(\frac{\partial \mathbf{c}_{\mathbf{d}^*}^T}{\partial \boldsymbol{\delta}} \right)^T \mathbf{I}_{\boldsymbol{\delta}}^{-1} \right), \quad (10)$$

where

$$\frac{\partial \mathbf{c}_{\mathbf{d}^*}^T}{\partial \boldsymbol{\delta}} = col_{1 \leq k \leq q} \frac{\partial \mathbf{c}_{\mathbf{d}^*}^T}{\partial \delta_k} = col_{1 \leq k \leq q} \frac{\partial \boldsymbol{\gamma}_{\mathbf{rd}^*}^T \mathbf{V}_{\mathbf{rs d}^*} \mathbf{V}_{\mathbf{ss d}^*}^{-1}}{\partial \delta_k}, \quad (11)$$

$$\mathbf{c}_{\mathbf{d}^*}^T = \boldsymbol{\gamma}_{\mathbf{rd}^*}^T \mathbf{V}_{\mathbf{rs d}^*} \mathbf{V}_{\mathbf{ss d}^*}^{-1} = \boldsymbol{\gamma}_{\mathbf{rd}^*}^T \left(\sigma_u^2 \mathbf{Z}_{\mathbf{rd}^*} \mathbf{C}_{\mathbf{d}^*}^{-1} \mathbf{Z}_{\mathbf{sd}^*}^T + diag_{1 \leq i \leq N_{rd^*}} (\boldsymbol{\Sigma}_{\mathbf{rs id}^*}) \right) \left(\sigma_u^2 \mathbf{Z}_{\mathbf{sd}^*} \mathbf{C}_{\mathbf{d}^*}^{-1} \mathbf{Z}_{\mathbf{sd}^*}^T + diag_{1 \leq i \leq n_{d^*}} (\boldsymbol{\Sigma}_{\mathbf{ss id}^*}) \right)^{-1} \quad (12)$$

Elements of (11) are given by

$$\frac{\partial \mathbf{c}_{\mathbf{d}^*}^T}{\partial \sigma_u^2} = \boldsymbol{\gamma}_{\mathbf{rd}^*}^T \left[\mathbf{Z}_{\mathbf{rd}^*} - \left(\sigma_u^2 \mathbf{Z}_{\mathbf{rd}^*} \mathbf{C}_{\mathbf{d}^*}^{-1} \mathbf{Z}_{\mathbf{sd}^*}^T + diag_{1 \leq i \leq N_{rd^*}} (\boldsymbol{\Sigma}_{\mathbf{rs id}^*}) \right) \mathbf{V}_{\mathbf{ss d}^*}^{-1} \mathbf{Z}_{\mathbf{sd}^*} \right] \mathbf{C}_{\mathbf{d}^*}^{-1} \mathbf{Z}_{\mathbf{sd}^*}^T \mathbf{V}_{\mathbf{ss d}^*}^{-1}, \quad (13)$$

$$\begin{aligned} \frac{\partial \mathbf{c}_{\mathbf{d}^*}^T}{\partial \sigma_\varepsilon^2} &= \boldsymbol{\gamma}_{\mathbf{rd}^*}^T \left(\sigma_\varepsilon^{-2} diag_{1 \leq i \leq N_{rd^*}} (\boldsymbol{\Sigma}_{\mathbf{rs id}^*}) \right) \mathbf{V}_{\mathbf{ss d}^*}^{-1} \\ &- \boldsymbol{\gamma}_{\mathbf{rd}^*}^T \left(\sigma_u^2 \mathbf{Z}_{\mathbf{rd}^*} \mathbf{C}_{\mathbf{d}^*}^{-1} \mathbf{Z}_{\mathbf{sd}^*}^T + diag_{1 \leq i \leq N_{rd^*}} (\boldsymbol{\Sigma}_{\mathbf{rs id}^*}) \right) \mathbf{V}_{\mathbf{ss d}^*}^{-1} \times \\ &\quad \times \sigma_\varepsilon^{-2} diag_{1 \leq i \leq n_{d^*}} (\boldsymbol{\Sigma}_{\mathbf{ss id}^*}) \mathbf{V}_{\mathbf{ss d}^*}^{-1}, \end{aligned} \quad (14)$$

$$\begin{aligned} \frac{\partial \mathbf{c}_{\mathbf{d}^*}^T}{\partial \rho_{(sp)}} &= \boldsymbol{\gamma}_{\mathbf{rd}^*}^T \left(\sigma_u^2 \mathbf{Z}_{\mathbf{rd}^*} \mathbf{A}_{\mathbf{d}^*} \mathbf{Z}_{\mathbf{sd}^*}^T \right) \mathbf{V}_{\mathbf{ss d}^*}^{-1} \\ &- \boldsymbol{\gamma}_{\mathbf{rd}^*}^T \left(\sigma_u^2 \mathbf{Z}_{\mathbf{rd}^*} \mathbf{C}_{\mathbf{d}^*}^{-1} \mathbf{Z}_{\mathbf{sd}^*}^T + diag_{1 \leq i \leq N_{rd^*}} (\boldsymbol{\Sigma}_{\mathbf{rs id}^*}) \right) \mathbf{V}_{\mathbf{ss d}^*}^{-1} \times \\ &\quad \times \left(\sigma_u^2 \mathbf{Z}_{\mathbf{sd}^*} \mathbf{A}_{\mathbf{d}^*} \mathbf{Z}_{\mathbf{sd}^*}^T \right) \mathbf{V}_{\mathbf{ss d}^*}^{-1}, \end{aligned} \quad (15)$$

$$\begin{aligned} \frac{\partial \mathbf{c}_{\mathbf{d}^*}^T}{\partial \rho_{(t)}} &= \boldsymbol{\gamma}_{\mathbf{rd}^*}^T \left(diag_{1 \leq i \leq N_{rd^*}} (\mathbf{B}_{\mathbf{rs id}^*}) \right) \mathbf{V}_{\mathbf{ss d}^*}^{-1} - \\ &\boldsymbol{\gamma}_{\mathbf{rd}^*}^T \left(\sigma_u^2 \mathbf{Z}_{\mathbf{rd}^*} \mathbf{C}_{\mathbf{d}^*}^{-1} \mathbf{Z}_{\mathbf{sd}^*}^T + diag_{1 \leq i \leq N_{rd^*}} (\boldsymbol{\Sigma}_{\mathbf{rs id}^*}) \right) \mathbf{V}_{\mathbf{ss d}^*}^{-1} \\ &\quad \times diag_{1 \leq i \leq n_{d^*}} (\mathbf{B}_{\mathbf{ss id}^*}) \mathbf{V}_{\mathbf{ss d}^*}^{-1}. \end{aligned} \quad (16)$$

The kl -th element of \mathbf{I}_δ is given by:

$$I_{kl}(\delta) = \frac{1}{2} \text{tr} \left(\mathbf{V}_{ss}^{-1} \frac{\partial \mathbf{V}_{ss}}{\partial \delta_k} \mathbf{V}_{ss}^{-1} \frac{\partial \mathbf{V}_{ss}}{\partial \delta_l} \right) \quad (17)$$

where

$$\begin{aligned} \frac{\partial \mathbf{V}_{ss}}{\partial \sigma_u^2} &= \text{diag}_{1 \leq d \leq D} \left(\mathbf{Z}_{sd} \mathbf{C}_d^{-1} \mathbf{Z}_{sd}^T \right), \quad \frac{\partial \mathbf{V}_{ss}}{\partial \sigma_\varepsilon^2} = \text{diag}_{1 \leq d \leq D} \left(\sigma_\varepsilon^{-2} \text{diag}_{1 \leq i \leq n_d} (\boldsymbol{\Sigma}_{ss \text{ id}}) \right), \\ \frac{\partial \mathbf{V}_{ss}}{\partial \rho_{(t)}} &= \text{diag}_{1 \leq d \leq D} \left(\text{diag}_{1 \leq i \leq n_d} (\mathbf{B}_{ss \text{ id}}) \right), \quad \frac{\partial \mathbf{V}_{ss}}{\partial \rho_{(sp)}} = \text{diag}_{1 \leq d \leq D} \left(\sigma_u^2 \mathbf{Z}_{sd} \mathbf{A}_d \mathbf{Z}_{sd}^T \right). \end{aligned}$$

What is important, if the assumptions presented by Żądło (2009) are met, that the order of the bias of the estimator (7) is $o(D^{-1})$. Żądło (2009) assumed inter alia that variance-covariance matrix may be decomposed into linear combination of variance parameters (similarly to the regularity condition e) presented by Datta and Lahiri (2000)). In the case of the proposed model (1) this assumption is not met. This means that the MSE estimator (7) is not proved to be approximately unbiased for the considered model. In the simulation study Żądło (2012) showed that although the performance of the estimator (7) is quite good, in some cases it can significantly overestimate the MSE (see results presented in Table 1). This is the reason, that new MSE estimator should be studied.

3. Jackknife MSE estimator

In this section we propose the Jiang, Lahiri, Wan (2002) type MSE estimator of the proposed EBLUP. This is given by:

$$\begin{aligned} & \widehat{MSE}_\xi^{jack}(\hat{\theta}_{d^{**}}^{EBLU}) \\ &= b_{d^{**}}(\hat{\boldsymbol{\delta}}) - \frac{D-1}{D} \sum_{d=1}^D (b_{d^{**}}(\hat{\boldsymbol{\delta}}_{-d}) - b_{d^{**}}(\hat{\boldsymbol{\delta}})) + \frac{D-1}{D} \sum_{d=1}^D (\hat{\theta}_{d^{**}}^{EBLU}(\hat{\boldsymbol{\delta}}_{-d}) - \hat{\theta}_{d^{**}}^{EBLU}(\hat{\boldsymbol{\delta}}))^2 \end{aligned} \quad (18)$$

where $\hat{\boldsymbol{\delta}}_{-d}$ is the estimator given by the same formula as $\hat{\boldsymbol{\delta}}$ but based on data without the d -th domain,

$$b_{d^{**}}(\hat{\boldsymbol{\delta}}) = g_1(\hat{\boldsymbol{\delta}}) + g_2(\hat{\boldsymbol{\delta}}), \quad (19)$$

$g_1(\hat{\delta}), g_2(\hat{\delta})$ are given by (8) and (9) respectively, where δ is replaced by $\hat{\delta}$, $b_{d^{**}}(\hat{\delta}_{-d})$ is given by (19), where $\hat{\delta}$ is replaced by $\hat{\delta}_{-d}$, $\hat{\theta}_{d^{**}}^{EBLU}(\hat{\delta}_{-d})$ is given by (6) where $\hat{\delta}$ is replaced by $\hat{\delta}_{-d}$.

Importantly, Jiang, Lahiri, Wan (2002) do not assume that the variance-covariance matrix may be decomposed into linear combination of variance parameters, as is assumed in the case of estimator (7). The properties of MSE estimators (7) and (18) will be studied in the simulation study.

4. Simulation study

We conduct a limited model-based simulation study prepared using R (R Development Core Team, 2012) similarly to Żądło (2012). It is based on artificial data. A population of size $N = 200$ elements is divided into $D = 10$ domains of sizes $\{15, 15, 15, 20, 20, 20, 20, 25, 25, 25\}$. Number of periods $M = 3$ is assumed and a balanced panel sample is studied – in each period the same $n_d = 5$ elements from each domain are observed in the sample (overall sample size in each period is $n = 50$). The purpose of the study is to predict $D = 10$ domain totals for the last period.

Data are generated based on the model (1) where $\forall_{idj} x_{idj} = 1$, $\forall_{idj} z_{idj} = 1$, $\forall_d \beta_d = \beta$ and for arbitrary chosen values of parameters $\beta = 100$, $\sigma_\varepsilon^2 = 1$, $\sigma_u^2 = 1$. Matrix \mathbf{W}_d is row-standardized neighborhood matrix (each element has two neighbours). In the simulation the following values of $\rho_{(sp)}$ and $\rho_{(t)}$ are considered: 0.8; 0.3; -0.3 and -0.8 which gives sixteen pairs of these correlation coefficients (as presented in Table 1). Realizations of random components are generated using multivariate normal distribution.

Żądło (2012) received values of relative biases of the MSE estimator based on the Taylor expansion which are summarized in Table 1.

Although the assumptions of the approximate unbiasedness of the estimator (7) are not met for the studied superpopulation model, the estimator performs quite well in the simulation (as presented in Table 1). On average it overestimates the unknown MSE, which is preferable to an underestimation. In most of the cases (out of the studied 16) the overestimation is not high, but there are two cases where the means over domains of relative biases of the MSE estimator based on the Taylor expansion are higher than 10% (i.e. 26.49% and 18.78%).

Table 1. Values of means over domains of relative biases of the MSE estimator based on the Taylor expansion (in %)

| ρ_{sp} | ρ_t | Values of means over domains of relative biases of the MSE estimator based on the Taylor expansion (in %) |
|-------------|-------------|---|
| 0.8 | 0.8 | 7.17 |
| 0.8 | 0.3 | 1.88 |
| 0.8 | -0.3 | 3.86 |
| 0.8 | -0.8 | 1.26 |
| 0.3 | 0.8 | 26.49 |
| 0.3 | 0.3 | 9.75 |
| 0.3 | -0.3 | 3.31 |
| 0.3 | -0.8 | 1.11 |
| -0.3 | 0.8 | 18.78 |
| -0.3 | 0.3 | 5.37 |
| -0.3 | -0.3 | 2.79 |
| -0.3 | -0.8 | -0.33 |
| -0.8 | 0.8 | 1.37 |
| -0.8 | 0.3 | 4.10 |
| -0.8 | -0.3 | 1.58 |
| -0.8 | -0.8 | 0.60 |

Source: own computations based on results presented by Żądło (2012).

Three of the cases in Table 1 are presented in bold, where the following values of mean relative biases of the MSE estimator based on the Taylor expansion were obtained:

- the maximum value (for $\rho_{(t)} = 0.3$ and $\rho_{(sp)} = 0.8$),
- the minimum value (for $\rho_{(t)} = -0.3$ and $\rho_{(sp)} = -0.8$),
- the mean value (for $\rho_{(t)} = -0.8$ and $\rho_{(sp)} = -0.3$).

In this paper we compare values of the biases of the MSE estimator based on the Taylor expansion (presented in Table 1) with values of biases of jackknife MSE estimator. Because in each iteration the parameters of the assumed model must be estimated D -times (residual maximum likelihood method is used) due to the jackknife procedure, the Monte Carlo simulation analysis is very time consuming. In the studied cases which are described below, the time needed for 5000 iterations of one Monte Carlo simulation study was between c.a. 40-80 hours for different assumed values of model

parameters. This is the reason that only 3 out of 16 cases (3 out of 16 pairs of values of $\rho_{(sp)}$ and $\rho_{(t)}$) were chosen for further analysis – the 3 cases presented in bold in Table 1 will be studied below.

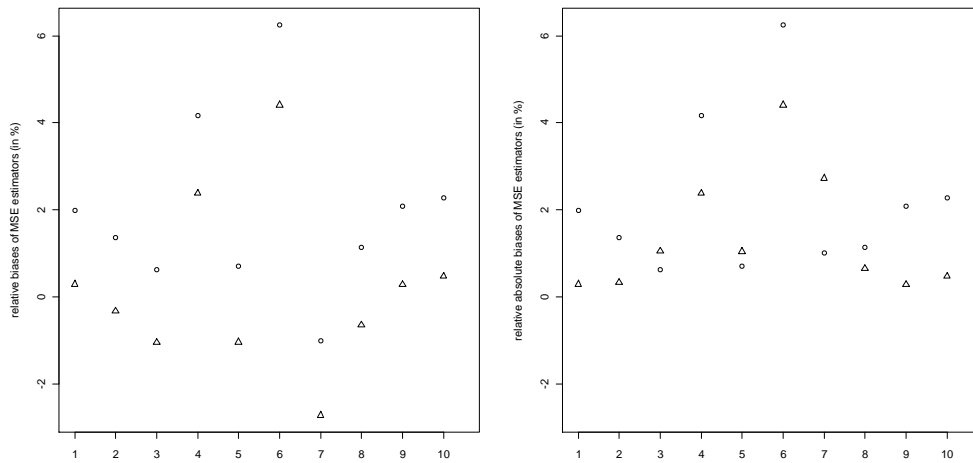


Fig. 1. Values of the jackknife and the Taylor expansion based MSE estimators for 10 domains ($d = 1, \dots, 10$) for $\rho_{(t)} = -0.8$ and $\rho_{(sp)} = -0.3$ (the mean case)

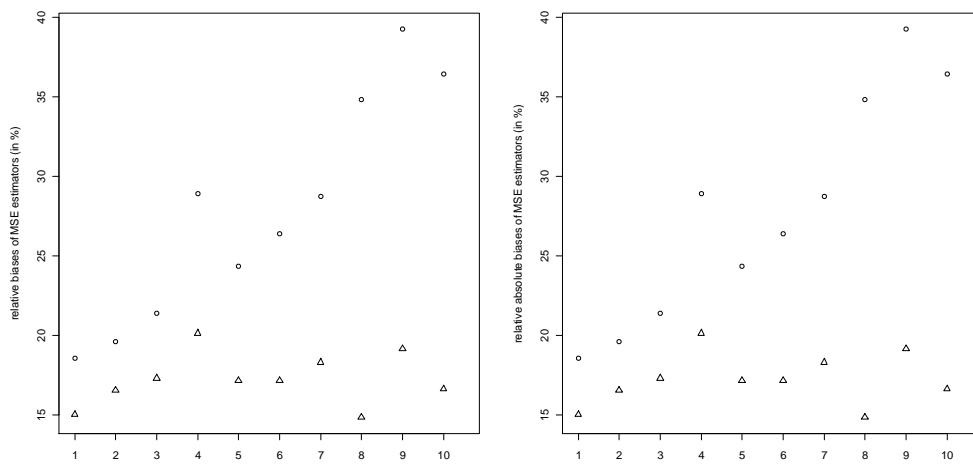


Fig. 2. Values of the jackknife and the Taylor based MSE estimators for 10 domains ($d = 1, \dots, 10$) for $\rho_{(t)} = 0.3$ and $\rho_{(sp)} = 0.8$ (the maximum case)

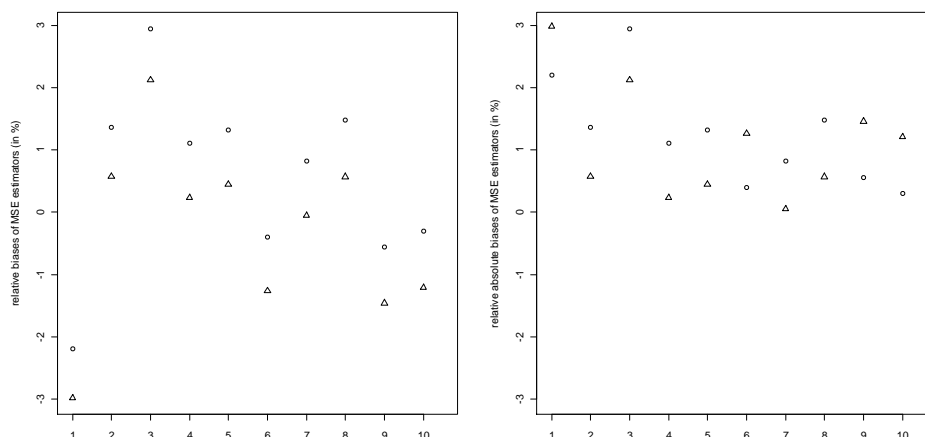


Fig. 3. Values of the jackknife and the Taylor based MSE estimators for 10 domains ($d = 1, \dots, 10$) for $\rho_{(t)} = -0.3$ and $\rho_{(sp)} = -0.8$ (the minimum case)

In Figures 1-3, symbol “ \circ ” denotes the value of the MSE estimator based on the Taylor expansion and symbol “ Δ ” denotes the value of the jackknife MSE estimator. On y-axis, the values of relative (or relative absolute) biases of the estimators are presented. On x-axis, the number of the domain under study is presented ($d = 1, \dots, 10$).

In the studied three cases in the simulation the relative biases of the jackknife MSE estimator are smaller for all of the domains than the relative biases of the MSE estimator based on the Taylor expansion. This is not true for relative absolute biases, but if the absolute relative biases of the jackknife MSE estimator is higher in the simulation than the absolute relative biases of the MSE estimator based on the Taylor expansion, the difference is small. The values of the relative biases of the MSE estimator based on the Taylor expansion presented in Figure 2 are quite high, but using the jackknife MSE estimator in some cases allows to reduce the bias significantly. The advantage of the jackknife MSE estimator over the MSE estimator based on the Taylor expansion is especially visible for cases of the highest bias of the MSE estimator based on the Taylor expansion (see the results in Figure 2 for the domains 7, 8, 9 and 10). Summarizing, the jackknife MSE estimator has a simpler form than the MSE estimator based on the Taylor expansion, which is very important in survey sampling practice, and its absolute biases – for different cases – are similar or smaller than absolute biases of the MSE estimator based on the Taylor expansion.

5. Summary

In the paper, two MSE estimators of some EBLUP of the domain total are compared in the simulation study. The first MSE estimator was derived by Żądło (2012) based on the Taylor expansion. The second MSE estimator is proposed based on the results of Jiang, Lahiri, Wan (2002). In the simulation study it is shown that the jackknife MSE estimator may be preferred to the MSE estimator based on the Taylor expansion.

Acknowledgement

The research was supported by National Science Centre grant 2011/03/B/HS4/00954.

References

- Datta G.S., Lahiri P. (2000). *A unified measure of uncertainty of estimated best linear unbiased predictors in small area estimation problems*. Statistica Sinica 10. Pp. 613-627.
- Jiang J., Lahiri P., Wan S.-M. (2002). *A unified jackknife theory for empirical best prediction with M-estimation*. The Annals of Statistics. Vol. 30. No 6. Pp. 1782-1810.
- R Development Core Team (2012). *A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna.
- Royall R.M. (1976). *The linear least squares prediction approach to two-stage sampling*. Journal of the American Statistical Association 71. Pp. 657-473.
- Żądło T (2004). *On unbiasedness of some EBLU predictor*. In: J. Antoch (ed). *Proceedings in Computational Statistics 2004*. Heidelberg-New York. Physica-Verlag. Pp. 2019-2026.
- Żądło T. (2009). *On MSE of EBLUP*. Statistical Papers 50. Pp. 101-118.
- Żądło T. (2014). *On the prediction of the subpopulation total based on spatially correlated longitudinal data*. Mathematical Population Studies 21. Pp. 30-44.