STATISTICS IN TRANSITION new series, June 2021
Vol. 22, No. 2, pp. 155–172, DOI 10.21307/stattrans-2021-021
Received – 17.06.2020; accepted – 05.05.2021

Small area estimates of the low work intensity indicator at voivodeship level in Poland

Łukasz Wawrowski[1], **Maciej Beręsewicz**[2]

ABSTRACT

The EU Statistics on Income and Living Conditions (EU-SILC) has provided annual estimates of the number of labour market indicators for EU countries since 2003, with an almost exclusive focus on national rates. However, it is impossible to obtain reliable direct estimates of labour market statistics at low levels based on the EU-SILC survey. In such cases, model-based small area estimation can be used. In this paper, the low work intensity indicator for the spatial domains in Poland between 2005-2012 was estimated. The Rao and You (1994), Fay and Diallo (2012), and Marhuenda, Molina and Morales (2013) models were applied. The bootstrap MSE for the discussed methods was proposed. The results indicate that these models provide more reliable estimates than direct estimation.

Key words: EU-SILC, low work intensity, labour market, small area estimation, temporal models, spatio-temporal models.

1. Introduction

Sample surveys conducted by National Statistical Institutes (NSIs) are in most cases representative at the national or region level (in particular at NUTS 1 level). In more detailed domains, such as states/voivodeships (NUTS 2) or subregions (NUTS 3), a small sample size does not allow for obtaining precise and accurate estimates of socio-economic indicators. Therefore, one needs methods that may provide more reliable estimates. For that purpose small area estimation (SAE) is often used. SAE makes it possible to estimate characteristics even if the sample is small, direct estimation is not reliable or domains are not observed in the sample. The underlying idea of SAE is to account for random effects in studied domains and „borrow strength" from auxiliary variables, over time or in space.

Small area estimation methods are widely used in many statistical domains. Social sciences examples include the labour market (López-Vizcaíno, Lombardía and Morales 2013), poverty (Molina and Rao 2010; Szymkowiak, Młodak and Wawrowski 2017) and business statistics (Chandra, Chambers and Salvati 2012; Dehnel and Wawrowski 2020). Due to limited access to data many applications cover estimation for only one year.

The main goal of the study described in this article was to estimate the low work intensity indicator (LWI) in the domains defined by the level of voivodeships (NUTS 2) in

[1]Computer Science Research Centre, Research Network Łukasiewicz - Institute of Innovative Technologies EMAG. E-mail: lukasz.wawrowski@emag.lukasiewicz.gov.pl. ORCID: https://orcid.org/0000-0002-1201-5344.

[2]Department of Statistics, Poznań University of Economics and Business. Statistical Office in Poznań, Centre for Small Area Estimation. E-mail: maciej.beresewicz@ue.poznan.pl. ORCID: https://orcid.org/0000-0002-8281-4301.

Poland between 2005 and 2012 with acceptable precision measured by coefficient of variation (CV). The LWI indicator, at-risk-of-poverty and material deprivation indicators are required by Eurostat as part of Europe 2020 strategy. The current official information is available only at the national and the NUTS 1 level in Poland and other EU states. At the more detailed domains small sample size results in big variances of obtained estimates. To achieve the main goal we discuss three recent small area models — Rao and You (1994), Fay and Diallo (2012), and Marhuenda, Molina and Morales (2013) — and then apply them. The first two models take into account temporal effects, while the third also incorporates spatial effects.

The article has the following structure. First, we present the variable of interest — the low work intensity indicator. The third section provides the notation for direct and model-based estimation. We also calculate mean square error (MSE) and model diagnostics, and present Generalized Inflation Factors in the context of SAE. The fourth section describes the EU-SILC survey and data from 2005 to 2012. In the fifth section we present the results and model diagnostics. The article ends with a summary.

2. Low work intensity

2.1. EU-SILC survey

The survey to collect EU Statistics on Income and Living Conditions (EU-SILC) was launched in 2003. The main aim of the survey was to deliver comparable data about income, poverty and living conditions of households in EU Member States. EU-SILC data are collected using a questionnaire in face-to-face interviews covering demography, education, health, housing conditions, economic activity, and more importantly, the level and sources of household incomes. EU-SILC is a sample-based, representative survey, in which a household is the basic statistical unit. In addition, every household member above 16 is also surveyed.

Various social cohesion indicators are estimated based on the EU-SILC survey. Several of them are used to monitor Europe 2020 strategy and to calculate the fraction of people living in households with very low work intensity (Statistics Poland 2014).

2.2. Low work intensity indicator

According to Eurostat, "the indicator of persons living in household with low work intensity is defined as the number of persons living in a household where work intensity is below a threshold set at 20%". Intensity of work is defined as the number of months that all working age household members (aged between 18 and 64) worked during the reference year divided by the total number of months that could theoretically be worked within the household. This means that households with low work intensity caused by different factors do not utilize their available time for work. Time spent at work is defined by Eurostat as:

- months in paid employment (full-time or part-time),
- paid internships and trainings,
- self-employment, with or without employees,
- unpaid work in a family business (helping family members).

To calculate the low work intensity indicator the total number of hours worked per week for each respondent is computed. For part-time employees, "the number of months in terms of full time equivalents is estimated on the basis of the number of hours usually worked at the time of the interview" (Mélina and Emilio, 2012). Eurostat set a threshold for the low work intensity at the level of 20%. This value refers to the expected risk of poverty in households with low work intensity. Nevertheless, Ward and Ozdemir (2013) argued that the threshold should be set slightly higher. Equation (1) presents the last stage of calculating the low work intensity indicator for the domains.

$$\mathrm{LWI}_{dt} = \frac{\sum_{i=1}^{n_{dt}} I(WI_{i,dt} < 0.2) d_{i,dt}}{\sum_{i=1}^{n_{dt}} d_{i,dt}}, \tag{1}$$

where: $WI_{i,dt}$ is work intensity of i-th household member in d-th domain at time t, $d_{i,dt}$ is calibrated weight of i-th household member, $I(\bullet)$ is an indicator function with two values $\{0,1\}$.

3. Notation for estimators and diagnostics

The classic Fay and Herriot (1979) area-level model does not take into account temporal nor spatial random effects. Therefore, when a panel survey data are used for estimation, the correlation between years should not be neglected. Thus, for the purpose of estimating LWI we applied two area-level small area estimators that take into account temporal random effects (Rao and You, 1994; Fay and Diallo, 2012) and spatio-temporal random effects (Marhuenda, Molina and Morales, 2013) for NUTS 2 level. The motivation for choosing these estimators is the observed strong temporal effect for NUTS 2 (voivodeships) in Poland. In addition, we would like to verify whether including the spatial effect in the model leads to better estimates.

3.1. Direct estimator

Let $\Omega = \{1,\ldots,N\}$ denote the target population of size N. From this population we draw a sample according to the sampling scheme $s \subseteq \Omega$ of size n. Let Ω_{dt} denote target population in domain (e.g. area), $d = 1,...,D$ denote a domain and $t = 1,...,T$ denote the time when the survey was conducted. Next, π_{dti} denotes the inclusion probability of i-th unit in d-th domain at time t in the corresponding domain sample s_{dt} and $d_{dti} = \pi_{dti}^{-1}$ the corresponding sampling weight. The EU-SILC survey uses the calibration approach proposed by Deville and Särndal (1992) to account for non-response. Thus, $w_{dti} = \lambda_{dti} d_{dti}$ denotes a calibration weight and λ_{dti} is the scaling factor for sampling weights d_{dti}. Let y denote the target variable (low work intensity) defined as follows:

$$y_{dti} = \begin{cases} 1 & \text{if the household suffers from low work intensity,} \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

Therefore, a design-unbiased direct estimator of $\bar{y}_{dt}$ is the Horvitz-Thompson (HT) estimator for the subpopulation Ω_{dt}, given by:

$$\hat{\bar{y}}_{dt} = \sum_{i \in s_{dt}} w_{dti} y_{dti} / \sum_{i \in s_{dt}} w_{dti}. \tag{3}$$

Because NUTS 2 level was used for stratification, the variance of $\hat{\bar{y}}_{dt}$ was estimated using a nonparametric bootstrap method as follows. Separately for each time t according to the sampling scheme, in particular taking into account strata $h = 1, ..., H$, draw a sample with replacement B times. For each sample b calculate the bootstrapped weight defined by the equation (4):

$$w^b_{dti} = w_{dti} \frac{n_{h,dt}}{n_{h,dt} - 1} m^b_{dti}, \tag{4}$$

where $n_{h,dt}$ denotes the number of sampled units in stratum h, domain d, at time t in the original sample and m^b_{dti} denotes the number of times that i-th unit was included in sample b. Finally, the bootstrap estimator of the variance of $\hat{\bar{y}}_{dt}$ for the domain Ω_{dt} is derived by:

$$\hat{V}(\hat{\bar{y}}_{dt}) = \hat{\psi}_{dt} = \frac{1}{B-1} \sum_{b=1}^{B} (\hat{\bar{y}}^b_{dt} - \hat{\bar{y}}_{dt})^2, \tag{5}$$

where $\hat{\bar{y}}^b_{dt} = \sum_{i \in s_{dt}} w^b_{dti} y_{dti} / \sum_{i \in s_{dt}} w^b_{dti}$. For the sake of clarity, we will use ψ_{dt} for the known sampling variance.

3.2. Rao and Yu (1994) model

Rao and You (1994) proposed an extension of Fay and Herriot (1979) model, which accounts for domains defined as time-series and cross-sectional classification. The model assumes two random effects — the domain effect, which is constant in time, and autocorrelation of domain effects in time. The autocorrelation is assumed to be the same between domains.

To enable comparison, we will apply the notation used in Marhuenda, Molina and Morales (2013). Therefore, in the first stage, Rao and You (1994) model assumes the following sampling model:

$$\bar{y}_{dt} = \mu_{dt} + e_{dt} \tag{6}$$

where $e_{dt} \overset{ind.}{\sim} N(0, \psi_{dt})$, where ψ_{dt} is the known sampling variance. The second stage (the linking model) μ_{dt} is assumed to follow a linear mixed model given by:

$$\mu_{dt} = X'_{dt}\beta + u_{1d} + u_{2dt} \tag{7}$$

where X_{dt} is the matrix of auxiliary variables (fixed effects), $u_{1d} \overset{ind.}{\sim} N(0, \sigma^2_1)$ denotes the random effect for domain at time $t = 1$ and constant in time $u_{1d} = u_{1d,t=1} = u_{1d,t=2} = ... = u_{1d,t=T}$. The second random component denoted by u_{2dt} is assumed to follow the autoregressive process $AR(1)$ with σ^2_2 and ρ_2, and is given by:

$$u_{2dt} = \rho_2 u_{2d,t-1} + \varepsilon_{2dt}, \tag{8}$$

where $|\rho_2| < 1$ and $\varepsilon_{2dt} \overset{ind.}{\sim} N(0, \sigma_2^2)$. We use ρ_2 for autocorrelation to be consistent with the number of random effects and for the consistency with the other models. In addition, let $\theta = (\sigma_1^2, \sigma_2^2, \rho_2)'$ be the vector of unknown parameters involved in the covariance structure of the model. Finally, the BLUP estimator of $\bar{y}_{dt}$ obtained by Rao and You (1994) through the method of moments is given by:

$$\mu_{dt} = X'_{dt}\tilde{\beta} + (\sigma_1^2 1'_T + \sigma_2^2 \gamma_T)(\Sigma_d + \sigma_2^2 \Gamma + \sigma_1^2 1_T 1'_T)^{-1}(y_d - X_d \tilde{\beta}), \tag{9}$$

where, for simplicity, we use $u_{1d} = u1$, $u_{2dt} = u2$ and $\rho_2 = \rho$,

- Γ is a symmetric matrix $T \times T$ with elements $\rho^{|i-j|)}/1(-\rho^2)$,
- $V_d = \Sigma_d + \sigma_2^2 \Gamma + \sigma_1^2 1_T 1'_T = Cov(y_d)$,
- $V = diag(V_d) = Cov(y)$,
- $\tilde{\beta} = (X'V^{-1}X)^{-1}X'V^{-1}y$,

When $\hat{\theta} = (\hat{\sigma}_1^2, \hat{\sigma}_2^2, \hat{\rho}_2)$ is known, the EBLUP is given by

$$\hat{\mu}_{dt} = X'_{dt}\tilde{\beta} + (\hat{\sigma}_1^2 1'_T + \hat{\sigma}_2^2 \hat{\gamma}_T)(\hat{\Sigma}_d + \hat{\sigma}_2^2 \hat{\Gamma} + \hat{\sigma}_1^2 1_T 1'_T)^{-1}(y_d - X_d \tilde{\beta}). \tag{10}$$

The notation of EBLUP (10) can be simplified in equation (7) and is given by (11). Moreover, rewriting the equation (10) as (7) enables comparison with Marhuenda, Molina and Morales (2013) and specifies that the model can be estimated using the Henderson (1975) approach.

$$\hat{\mu}_{dt} = X'_{dt}\beta + \hat{u}_{1d} + \hat{u}_{2dt}. \tag{11}$$

3.3. Fay and Diallo (2012) model

Another extension of Fay and Herriot (1979) was proposed by Fay and Diallo (2012) and Fay, Planty and Diallo (2013). Fay and Diallo (2012) proposed a univariate and Fay, Planty and Diallo (2013) a multivariate dynamic small area model that takes into account autocorrelation of random effects for domains. The Fay and Diallo (2012) model also extends Rao and You (1994) by assuming nonstationarity of the domain effect, thus the effect is not constant over time. Fay and Diallo (2012) in the first stage assume a sampling model given by:

$$\bar{y}_{dt} = \mu_{dt} + e_{dt} \tag{12}$$

where $e_{dt} \overset{ind.}{\sim} N(0, \psi_{dt})$, where ψ_{dt} is known sampling variance. The second stage (the linking model) assumes a linear mixed model given by the following equation:

$$\mu_{dt} = X'_{dt}\beta + u_{1dt} + u_{2dt} \tag{13}$$

where $u_{1dt} = \rho_2^{t-1} u_{1d}$ and $u_{1d} \overset{ind.}{\sim} N(0, \sigma_1^2)$ is the random effect for d-th domain at time $t = 1$. The random effect u_{1d} is scaled by ρ_2, which denotes the autocorrelation for the

second random effect u_{2dt}. u_{2dt} is assumed to follow $AR(1)$ process, as does the Rao and You (1994) model, and is defined below:

$$u_{2dt} = \rho_2 u_{2d,t-1} + \varepsilon_{2dt}, \tag{14}$$

where $|\rho_2| < 1$ and $\varepsilon_{2dt} \overset{ind.}{\sim} N(0, \sigma_2^2)$. The main difference between the Fay and Diallo (2012) and Rao and You (1994) approach is that the former does not constrain $|\rho_2| < 1$ and avoids discontinuity at $\rho_2 = 1$. When $\rho_2 > 1$ a divergence between domains is observed. Let $\theta = (\sigma_1^2, \sigma_2^2, \rho_2)'$ be the vector of unknown parameters involved in the covariance structure of the model. The BLUP estimator for μ_{dt} is calculated in the same fashion as (9):

$$\mu_{dt} = x_{dt}'\tilde{\beta} + (\sigma_1^2 \gamma_{T,u1} + \sigma_2^2 \gamma_{T,u2}) V_d^{-1} (y_d - X_d \tilde{\beta}), \tag{15}$$

where the elements are defined as follows (for simplicity $u_1 = u_{1d}$, $u_2 = u_{2dt}$ and $\rho_2 = \rho$ is used):

- Γ_{u1} is a symmetric matrix $T \times T$ where $\Gamma_{u1(1,j)} = 0$ and $\Gamma_{u1(i,j)} = \rho^{(j-i)} \sum_{i'=1}^{i-1} \rho^{(2i'-2)}$ for $1 < i \leq j$,
- Γ_{u2} is a symmetric matrix $T \times T$ of elements ρ^{i+j-2},
- $V_d = \Sigma_d + \sigma_1^2 \Gamma_{u1} + \sigma_2^2 \Gamma_{u2} = Cov(y_d)$,
- $V = diag(V_d) = Cov(y)$,
- $\tilde{\beta} = (X'V^{-1}X)^{-1}X'V^{-1}y$,
- $\gamma_{T,u1}$ is T column of matrix Γ_{u1},
- $\gamma_{T,u2}$ is T column of matrix Γ_{u2}.

Finally, when $\hat{\theta} = (\hat{\sigma}_1^2, \hat{\sigma}_2^2, \hat{\rho}_2)'$ is known, the EBLUP of μ_{dt} is given by

$$\hat{\mu}_{dt} = x_{dt}'\tilde{\beta} + (\hat{\sigma}_1^2 \hat{\gamma}_{T,u1} + \hat{\sigma}_2^2 \hat{\gamma}_{T,u2}) \hat{V}_d^{-1} (y_d - X_d \tilde{\beta}), \tag{16}$$

or, following Henderson (1975) and Marhuenda, Molina and Morales (2013), can also be written as:

$$\hat{\mu}_{dt} = X_{dt}'\beta + \hat{u}_{1dt} + \hat{u}_{2dt}, \tag{17}$$

where $\hat{u}_{1dt} = \hat{\rho}_2^{t-1} \hat{u}_{1d}$. For the proof of (15) and mathematical details of the model (16) refer to Fay and Diallo (2012).

3.4. Marhuenda, Molina and Moralez (2013) model

Finally, in order to verify whether to include the spatial effect, we applied the spatio-temporal model proposed by Marhuenda, Molina and Morales (2013). The model assumes two random effects — spatially correlated and temporally correlated domain effect. As in the previous models, in the first stage it assumes:

$$\bar{y}_{dt} = \mu_{dt} + e_{dt} \tag{18}$$

where $e_{dt} \overset{ind.}{\sim} N(0, \psi_{dt})$, where ψ_{dt} is the known sampling variance. In the second stage (the linking model) a linear mixed model is assumed and is given by:

$$\mu_{dt} = X'_{dt}\beta + u_{1d} + u_{2dt} \tag{19}$$

where u_{1d} denotes a spatial random effect that follows the SAR(1) process with variance σ_1^2, spatial autocorrelation ρ_1 and row-standardized proximity matrix $W = (w_{d,k})$. Such a proximity matrix is created based on neighbours matrix W^0. The matrix W is derived from the matrix W^0 by dividing each row element by the row total (Bivand, Pebesma and Gomez-Rubio, 2013). We assume that the spatial representation of domains does not change over time (borders are the same). The SAR(1) process is given by:

$$u_{1d} = \rho_1 \sum_{d \neq k} w_{d,k} u_{1k} + \varepsilon_{1d}, \tag{20}$$

where $|\rho_1| < 1$, and $\varepsilon_{1d} \overset{ind.}{\sim} N(0, \sigma_1^2)$. The second random effect u_{2dt} is assumed to follow the AR(1) process with σ_2^2 and ρ_2 and is given by the following equation:

$$u_{2dt} = \rho_2 u_{2d,t-1} + \varepsilon_{2dt}, |\rho_2| < 1, \varepsilon_{2dt} \overset{ind.}{\sim} N(0, \sigma_2^2). \tag{21}$$

Let $\theta = (\sigma_1^2, \sigma_2^2, \rho_1, \rho_2)'$ be the vector of unknown parameters involved in the covariance structure of the model. After the estimation of $\hat{\theta} = (\hat{\sigma}_1^2, \hat{\sigma}_2^2, \hat{\rho}_1, \hat{\rho}_2)'$ the EBLUP estimator (19) of $\bar{y}_{dt}$ proposed by Marhuenda, Molina and Morales (2013) is given by:

$$\hat{\mu}_{dt} = X'_{dt}\beta + \hat{u}_{1d} + \hat{u}_{2dt}. \tag{22}$$

In contrast to Rao and You (1994) and Fay and Diallo (2012), Marhuenda, Molina and Morales (2013) estimated the parameters using the Henderson (1975) approach instead of the method of moments. Details about the estimation of the model (19) can be found in Marhuenda, Molina and Morales (2013) and Molina and Marhuenda (2015).

Shortly summarizing the models presented, the following differences can be indicated. Rao and You model assumes stationarity for time series and two uncorrelated random effects. In Fay and Diallo model a time series is non-stationary and random effects are correlation. Marhuenda, Molina and Morales model takes into account SAR(1) process for the first random effect and AR(1) process for the second random effect.

3.5. MSE calculation

Rao and You (1994) and Fay and Diallo (2012) obtained MSE for estimators (16) and (10) by deriving a direct formula using the method of moments based on the Prasad and Rao (1990) approach. In contrast, Marhuenda, Molina and Morales (2013) proposed a parametric bootstrap to estimate MSE of (22). The motivation for such an approach is based on the González-Manteiga et al. (2008) and Molina, Salvati and Pratesi (2009) papers, which discussed estimation of MSE through the parametric bootstrap.

Therefore, to make MSE comparable between the models we applied the parametric bootstrap approach for each model. In the case of Marhuenda, Molina and Morales (2013)

model, the parametric bootstrap was available. For Rao and You (1994) and Fay and Diallo (2012) we developed a procedure to estimate MSE under the parametric bootstrap. The details can be found in Table 1. The notation used in the table is consistent with the Marhuenda, Molina and Morales (2013) article. The steps (3), (5), (7) and (8) are the same for all models.

3.6. Diagnostics measures for models

3.6.1 Model comparison measures

In order to compare and evaluate the models we applied several measures. Firstly, we used cAIC criterion (Greven and Kneib, 2010), pseudo-R^2 and Wald statistic. These measures were used to compare and verify which model is the most suitable for estimation of the low work intensity indicator. In addition, for practical and descriptive reasons, pseudo-R^2 for each model was computed and is given in (23). The inclusion of the pseudo-R^2 measure is motivated by the ease of interpretation as a measure of goodness of fit and end users' experiences with linear models. However, this measure is rarely presented in the context of small area models. For other pseudo-R^2 measures for linear mixed models, see Nakagawa and Schielzeth (2013), and for Wald statistic denoted by W refer to Brown et al. (2001). Calculated information criteria are given in (23):

$$\begin{aligned} cAIC &= -2\times LogLik + 2\times(trace(H)+1), \\ pseudo-R^2 &= \mathrm{Var}(\hat{\mu}_{dt})/\mathrm{Var}(\hat{\bar{y}}_{dt}), \\ W &= \sum(\hat{\bar{y}}_{dt}-\hat{\mu}_{dt})^2/(\mathrm{Var}(\hat{\bar{y}}_{dt})+\mathrm{Var}(\hat{\mu}_{dt})), \end{aligned} \tag{23}$$

where $LogLik$ is the value of log-likelihood estimated through REML estimation of the variance components, p denotes the number of model parameters (fixed and for random effects), n denotes the number of observations, $trace(H)$ trace of hat matrix given by equation (24) and Var denotes simple random sampling variance.

$$\begin{aligned} trace(H) =& trace((X_{dt}'V(\theta)^{-1}X_{dt})^{-1}X_{dt}'V(\theta)^{-1}V_eV(\theta)^{-1}X_{dt}) \\ &+n-trace(V_eV(\theta)^{-1}) \end{aligned} \tag{24}$$

Bias correction of conditional Akaike information criterion is given by equation (24). V_e in this equation denotes variance matrix of random error. Calculation of this term is possible with cAIC4 R package written by Saefken et al. (2018). Conditional Akaike information criterion depends on the structure of the model used so two other metrics in (23) were proposed.

3.6.2 Collinearity diagnostics

To evaluate the models we investigated collinearity measures using generalized variance inflation factors (GVIF) proposed by Fox and Monette (1992). The GVIF measure is limited

Table 1: Calculation of parametric bootstrap MSE in Rao and Yu (1994), Fay and Diallo (2012) and Marhuenda, Molina and Moralez (2013) models

Step	Rao and Yu (1994)	Fay and Diallo (2012)	Marhuenda, Molina and Moralez (2013)
1	Using the available data $\{(\hat{\bar{y}}_{dt}, X_{dt}), t = 1,\ldots,T, d = 1,\ldots,D\}$, fit the Rao and You (1994) model to obtain model parameter estimates $\hat{\beta}$, $\hat{\sigma}_2^1$, $\hat{\sigma}_2^2$ and $\hat{\rho}_2$.	Using the available data $\{(\hat{\bar{y}}_{dt}, X_{dt}), t = 1,\ldots,T, d = 1,\ldots,D\}$, fit the Fay and Diallo (2012) model to obtain model parameter estimates $\hat{\beta}$, $\hat{\sigma}_2^1$, $\hat{\sigma}_2^2$ and $\hat{\rho}_2$.	Using the available data $\{(\hat{\delta}_{dt}^{DIR}, \mathbf{x}_{dt}), t = 1,\ldots,T, d = 1,\ldots,D\}$, fit the Molina, Marhuenda, Molina and Morales (2013) model to obtain model parameter estimates $\hat{\beta}$, $\hat{\sigma}_2^1$, $\hat{\rho}_1$, $\hat{\sigma}_2^2$ and $\hat{\rho}_2$.
2	Generate bootstrap area effects $\{u_{1d}^{*(b)}, d = 1,\ldots,D, t = 1,...,T\}$ using $\hat{\sigma}_1^2$ as true values of parameters σ_1^2 that $\{u_{1d}^{*(b)} = u_{1d,t=1}^{*(b)} = ... = u_{1d,t=T}^{*(b)}\}$.	Generate bootstrap area effects $\{u_{1d}^{*(b)}, d = 1,\ldots,D, t = 1\}$ with known $\hat{\sigma}_1^2$ as true value of parameter σ_1^2. Then, compute $\{u_{1dt}^{*(b)} = \rho_2^{t-1} u_{1d}^{*(b)}, t = 2,...,T\}$ where $\hat{\rho}_2$ is the true value of ρ_2.	Generate bootstrap area effects $\{u_{1d}^{*(b)}, d = 1,\ldots,D, t = 1,...,T\}$, from the SAR(1) process given in (20), using $(\hat{\sigma}_1^2, \hat{\rho}_1)$ as true values of parameters (σ_1^2, ρ_1) and $u_{1d}^{*(b)} = u_{1d,t=1}^{*(b)} = ... = u_{1d,t=T}^{*(b)}$.
3	Independently of $\{u_{1d}^{*(b)}\}$ and independently for each d, generate bootstrap time effects $\{u_{2dt}^{*(b)}, t = 1,\ldots,T\}$, from the AR(1) process given in (8), with $(\hat{\sigma}_2^2, \hat{\rho}_2)$ acting as true values of parameters (σ_2^2, ρ_2).		
4	Calculate true bootstrap quantities, $\mu_{dt}^{*(b)} = X_{dt}'\hat{\beta} + u_{1d}^{*(b)} + u_{2dt}^{*(b)}$.	Calculate true bootstrap quantities, $\mu_{dt}^{*(b)} = X_{dt}'\hat{\beta} + u_{1dt}^{*(b)} + u_{2dt}^{*(b)}$.	Calculate true bootstrap quantities, $\mu_{dt}^{*(b)} = X_{dt}'\hat{\beta} + u_{1d}^{*(b)} + u_{2dt}^{*(b)}$.
5	Generate errors $e_{dt}^{*(b)} \overset{ind.}{\sim} N(0, \psi_{dt})$ and obtain bootstrap data from the sampling model, $\hat{\bar{y}}_{dt}^{*(b)} = \mu_{dt}^{*(b)} + e_{dt}^{*(b)}$		
6	Using the new bootstrap data $\{(\hat{\bar{y}}_{dt}^{*(b)}, X_{dt}), t = 1,\ldots,T, d = 1,\ldots,D\}$, fit the Rao and You (1994) model (7) - (11) to obtain the bootstrap EBLUPs, $\hat{\mu}_{dt}^{*(b)}$	Using the new bootstrap data $\{(\hat{\bar{y}}_{dt}^{*(b)}, X_{dt}), t = 1,\ldots,T, d = 1,\ldots,D\}$, fit Fay and Diallo (2012) model (13) - (17) to obtain the bootstrap EBLUPs, $\hat{\mu}_{dt}^{*(b)}$	Using the new bootstrap data $\{(\hat{\bar{y}}_{dt}^{*(b)}, X_{dt}), t = 1,\ldots,T, d = 1,\ldots,D\}$, fit Marhuenda, Molina and Morales (2013) model (19) - (22) to obtain the bootstrap EBLUPs, $\hat{\mu}_{dt}^{*(b)}$
7	Repeat steps (1)-(6) for $b = 1,\ldots,B$, where B is a large number.		
8	Calculate parametric bootstrap MSE according to the following formula: $MSE(\hat{\mu}_{dt}) = \frac{1}{B}\sum_{b=1}^{B}(\hat{\mu}_{dt}^{*(b)} - \mu_{dt}^{*(b)})^2$		

to fixed effects (X_{dt}) and does not account for the variance structure of random effects. Thus, it overestimates the collinearity between auxiliary variables X_{dt}. Other approaches to

estimate VIF in the context of complex surveys are discussed by Liao and Valliant (2012) and Li and Valliant (2015) assuming a linear model with known sampling variances.

Therefore, we modified GVIF to be conditional on the Fay-Herriot small area model covariance matrix of y given by: $V(\theta) = ZV(\theta)_u Z' + V_e$, where Z is a matrix of random effects, $V(\theta)_u$ denotes block-diagonal covariance structure for random effects and V_e is a diagonal matrix of known sampling variances. Let $\Sigma_x(\theta)$ denote the variance-covariance matrix for the fixed effect X_{dt} defined by the equation (25)

$$\Sigma_x(\theta) = (X'_{dt} V(\theta)^{-1} X_{dt})^{-1}, \tag{25}$$

and the estimator of (25) is given by

$$\hat{\Sigma}_x = \tilde{\Sigma}_x(\hat{\theta}) = (X'_{dt} V(\hat{\theta})^{-1} X_{dt})^{-1}, \tag{26}$$

where $V(\hat{\theta})$ is an estimated covariance structure of the small area model. The $V(\hat{\theta})$ can differ between the models and depends on the assumed underlying structure of random effects. To calculate conditional GVIF $\hat{\Sigma}_x$ need to be transformed into a correlation matrix, which we denote as $R(\theta)$. The estimator of $R(\theta)$ is given by the following transformation of $\hat{\Sigma}_x$

$$R(\hat{\theta}) = D^{-1} \hat{\Sigma}_x D^{-1}, \tag{27}$$

where $D = diag(\sqrt{diag(\hat{\Sigma}_x)})$. Finally, the GVIF conditional on $V(\hat{\theta})$ for each variable of the fixed effect is given by

$$GVIF(x_k | V(\hat{\theta})) = \frac{det(R(\hat{\theta})_{k,k}) \times det(R(\hat{\theta})_{-k,-k})}{det(R(\hat{\theta}))} \tag{28}$$

where x_k denotes k-th variable from the auxiliary matrix X_{dt}, det denotes the determinant of a matrix, $R(\hat{\theta})_{k,k}$ denotes matrix with k-th variable and $R(\hat{\theta})_{-k,-k}$ without k-th variable. According to Chatterjee and Price (1991), it is assumed that values $GVIF(x_k | V(\hat{\theta}))$ exceeding 10 are to be highly correlated with other fixed effects. Thus, a given variable should be removed from the small area model.

4. Data utilized in the study

4.1. EU-SILC data

The study was based on EU-SILC data from 8 years: 2005 to 2012. As mentioned earlier, the EU-SILC survey is conducted to collect information on income, poverty and other aspects of living conditions of households in European countries. The sample size is set to be representative at the national level. However, in Poland the sample size is big enough to publish information about households at the regional level (NUTS 1) as well.

The number of households in the sample varies from 317 (Opolskie Voivodeship in 2009) to 2,212 (Slaskie Voivodeship in 2005). According to the sampling scheme applied, the sample size was distributed proportionally to the domains in the voivodeship. It should

be noticed that the sample of households in the survey decreases from year to year. An average decrease compared to the base year 2005 is 20%. The change is due to several causes. First of all, EU-SILC is a panel and thus requires respondents to participate in the survey multiple times. In addition, non-response is present, which decreases the sample size. Coefficient of variation for direct estimates varies from 5.4% for Slaskie Voivodeship in 2005 to 37.4% for Podlaskie in 2010. For these reasons the sample size in the domains of interest is not acceptable for deriving direct estimates.

4.2. Auxiliary variables

Small area estimation at area-level requires auxiliary information about study domains. Rao and Molina (2015) recommend using register or census data that are free from sampling errors. Therefore, to estimate models, we collected socio-economic data from the Local Data Bank maintained by Statistics Poland. The main criteria for the choice of variables were availability at NUTS 2 level for the years 2005-2012 and the source of data, in particular registers. Several variables were considered and finally the following ones were chosen: registered unemployment rate, working and post-working age people and the number of people in NUTS 2 regions.

The registered unemployment rate is calculated as the ratio of the number of registered unemployed persons to the economically active civilian population. Working and post-working age was used to create two ratios. First, the number of people of working age (aged 15-64) divided by the number of people of post-working age (65 and over). This measure can be interpreted as describing how many independent workers have to provide for one pensioner. The second ratio has the same numerator but the denominator is the number of people without additional criterion (the whole population).

5. Estimation of low work intensity indicator at voivodeship level

In this section we describe the results and provide diagnostics for each model. All the calculations were done in R using the following packages: `sae` (Molina and Marhuenda, 2015), `sae2` (Fay and Diallo, 2015), `metafor` (Viechtbauer, 2010). For the sake of simplicity, we will use `RY` for Rao and You (1994), `FD` for Fay and Diallo (2012) and `MMM` for Marhuenda, Molina and Morales (2013) model.

5.1. Comparison of models

Table 2 contains a comparison of the parameters and statistics for all the models. `RY` and `FD` had 7 parameters, while `MMM` had a total of 8 parameters. The fixed effects in all the models are significant and have expected signs. Slight differences can be observed in the level of the intercept in `FD`. In all the models registered unemployment rate is positively correlated with the LWI indicator: a rise in the level of registered unemployment is associated with higher LWI. When the ratio of the post-working age to working age population rises, the LWI also rises and the ratio of the working-age population to the whole population has the expected sign: if the ratio grows, the LWI decreases. Therefore, we can conclude that the

auxiliary variables are good predictors for the LWI indicator and do not differ between the models.

The second group of parameters are variances of random effects. For the sake of simplicity, standard deviations σ_{u*} are used instead of variances σ^2_{u*}. In `RY` the AR(1) effect dominates the domain effect and is responsible for almost all the variance of random effects. In contrast, in `FD` and `MMM` the domain effect is higher than the AR(1) process of random effects. In the case of the `FD` model, this means that the domain effect is not constant over time (is nonstationary) and is higher in the first year of the survey but decreases over time by 0.9407^{t-1}. On the other hand, in the `MMM` model the domain effect is spatially correlated and the variance of this random effect is higher than the AR(1) effect.

In all of the models the AR(1) effect has a strong autocorrelation (ρ_2), which means that the effects within domains depend strongly on what happened in the previous year. The `RY` and `FD` models indicate that this autocorrelation is over 0.9 while, in the case of `MMM`, we can observe a slightly smaller value. In the case of the `MMM` model, this is due to the second autocorrelation parameter (ρ_1), which is associated with the spatial effect (SAR(1)). The value of $\rho_1 = 0.4866$ indicates that a moderate spatial effect between NUTS 2 is observed, which is smaller than the AR(1) autocorrelation.

If we compare the model statistics concerning information criteria and R^2 we can observe slight differences between the models. All the models explained almost 85% of the variance of the direct estimator. The `RY` and `FD` models have similar information criteria, while the `MMM` model differs slightly. However, the differences between the model statistics do not clearly indicate which model should be recommended. Nonetheless, if we compare all the statistics in Table 2 the model proposed by Marhuenda, Molina and Morales (2013) seems to be the most reasonable due to the significant spatial effect.

The comparison of EBLUPs for the `RY`, `FD` and `MMM` models with the direct estimator indicates that the model-based estimation is coherent with direct estimation. Pearson correlation coefficients for all estimates are above 0.9. EBLUPs obtained for the models do not differ significantly; however, compared to direct estimates, we can observe differences between estimates.

The differences between model-based and direct estimates are visible in Figure 1. LWI decreases over time from over 15% to below 10%. The solid line indicates direct estimates and dashed lines represent model-based estimates. In general, we can observe a similar trend in all NUTS 2 regions in Poland, but at different levels of intensity. In addition, model-based estimates are more stable over time than direct estimates. In some voivodeships (Lubuskie, Podlaskie or Zachodniopomorskie) there is a clearly visible rise in LWI after 2008, which can be associated with the start of the 2008 crisis.

The biggest differences in the LWI indicator can be observed for Lubuskie and Opolskie Voivodeships. Direct estimates for Lubuskie indicate that from 2008 to 2010 LWI increased, while model-based estimates indicate that the increase was smaller and was only present between 2009 and 2010. These differences, however, may be due to the sampling error, which is higher at NUTS 2 level. It is possible that in the case of Lubuskie specific units were included in the sample in 2008 and took part in the EU-SILC survey until 2010. In Opolskie Voivodeship, there was a considerable increase in direct estimates between 2009 and 2010, followed by a decrease. These differences may also be due to the sampling error,

Table 2: Summary of the estimated model parameters and statistics. Standard deviations are given in parentheses after the mean values.

Parameters/Models	RY	FD	MMM
Model parameters – fixed effects			
Intercept	1.28 (0.27)	1.47 (0.30)	1.28 (0.29)
Register Unemployment Rate	0.37 (0.07)	0.35 (0.07)	0.38 (0.07)
Working / Post-Working Ratio	0.09 (0.01)	0.09 (0.01)	0.09 (0.01)
Working age / Population Ratio	-2.49 (0.46)	-2.80 (0.49)	-2.49 (0.46)
Model parameters – random domain effects variances			
σ_1 Domain effect	0.00 (0.11)	0.04 (0.03)	0.03 (0.02)
σ_2 AR(1)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)
Model parameters – random domain effects autocorrelation			
ρ_1 SAR(1)	-	-	0.47 (0.00)
ρ_2 AR(1)	0.98 (0.26)	0.94 (0.02)	0.88 (0.00)
Model statistics			
REML LL	327.72	329.13	336.69
cAIC	-582.90	-591.53	-581.96
pseudo R^2	83.97	84.57	84.43
$W(\chi^2_{0.05} = 155.40)$	43.76	47.49	43.45
Degrees of freedom	7	7	8

especially given that the region of Opolskie is characterized by the highest level of the standard error of direct estimates.

According to Brown et al (2001) the difference between direct estimates and model-based estimates should be not significant. Figure 1 shows that these differences are rather small. Pearson correlation coefficient for direct and `RY` model estimates vary from 0.3856 to 0.9886 with average equal to 0.9226. For `FD` model correlations are in the range [0.3600;0.9873] (average 0.9192) while for estimates derived from `MMM` the model minimum value is equal to 0.3906, maximum to 0.9894 and average to 0.9239. In all cases the lowest values were observed in Lubuskie voivodeship and the highest in Śląskie Voivodeship. The highest similarity of estimates measured by average correlation coefficient was obtained for Marhuenda, Molina and Moralez (2013) model. These results show consistency of direct and small area estimates.

5.2. Comparison of coefficient of variations of estimates

The distribution of the CV is given in Table 3. An increase in CV over time was observed, which is due to increasing non-response and the respondent burden in the EU-SILC survey. On average, CV for direct estimates is equal to 15.77%.

In comparison to model-based estimation, the CV for direct estimation increases more rapidly, while the CV for `RY`, `FD` and `MMM` models increase more steadily. Moreover, CVs differ depending on the NUTS 2 unit. For example, in Opolskie and Podlaskie CV is significantly higher in comparison to other NUTS 2 units in Poland, mainly owing to smaller sample sizes. Therefore, especially for these regions, the direct estimator is not reliable.

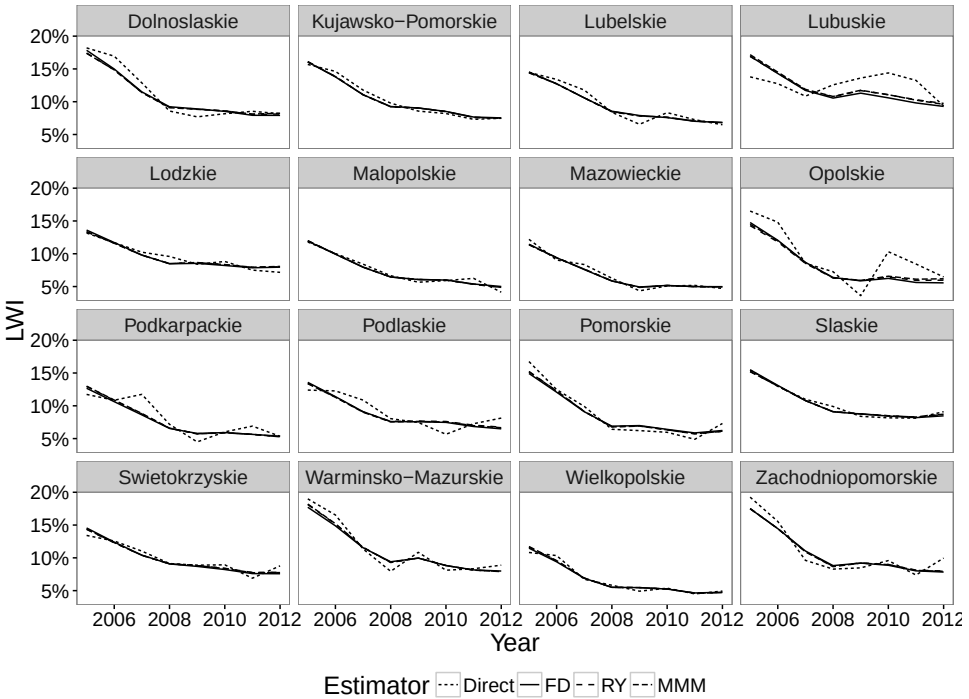

Figure 1: Comparison of direct, Rao and Yu (1994), Fay and Diallo (2012) and Marhuenda, Molina and Moralez (2013) models estimates

Table 3: Comparison of direct, Rao and Yu (1994), Fay and Diallo (2012) and Marhuenda, Molina and Moralez (2013) models coefficient of variations

Model	Min	Q1	Median	Mean	Q3	Max
Direct	5.37	11.62	14.76	15.77	18.55	37.38
RY	4.01	6.91	8.38	8.99	10.39	20.26
FD	3.99	6.92	8.36	9.04	10.33	21.74
MMM	4.19	6.84	8.55	9.09	10.45	20.57

CVs for all the models of interest are lower in comparison to the direct estimator. On average, the CV for each model is approximately 9%, which indicates that the model-based approach provides more reliable estimates. However, as was the case with model diagnostics, models `RY`, `FD` and `MMM` provide similar level of precision and, on average, the `RY` model is slightly better in comparison to the other models. The lowest CV can be observed for Slaskie and Mazowieckie Voivodeships and the highest for Podlaskie and Opolskie. What is worth noticing is model-based estimation provides more reliable estimates over time even if the non-response increases.

5.3. Diagnostics of the models

Table 4 contains information about GVIF calculated using formula (28). The first three columns refer to the model in question and the last one, denoted by WOLS, refers to weighted ordinary least square regression, where weights are the inverse of sampling errors. Results indicate that GVIF values for all variables in the `RY`, `FD` and `MMM` models are close to 1.3, which is lower than the threshold of 4. Moreover, as expected, the values are smaller than those observed in weighted OLS. The inclusion of the estimated covariance matrix accounts for the uncertainty.

Table 4: Generalized variance inflation measures for auxiliary variables used in Rao and Yu (1994), Fay and Diallo (2012) and Marhuenda, Molina and Moralez (2013) models

Variable	RY	FD	MMM	WOLS
Register Unemployment Rate	1.30	1.31	1.29	1.42
Working / Post-Working Ratio	1.31	1.33	1.29	2.05
Working age / Population Ratio	1.44	1.51	1.37	1.54

6. Conclusions

Application of three proposed models — Rao and You (1994), Fay and Diallo (2012), and Marhuenda, Molina and Morales (2013), allows to obtain more reliable (in the sense of CV) estimates in previously unpublished domains. All models take into account auxiliary variables, temporal effect, however Marhuenda, Molina and Morales (2013) also deal with spatial information. The registered unemployment rate showed the strongest relation with the indicator. Based on the results and strong spatial autocorrelation, we choose Marhuenda, Molina and Morales (2013) model as the most suitable for the estimation of the low work intensity indicator. The final results are presented in Figure 2.

Based on Figure 2, we noticed spatial regimes in the low work intensity in the West (Zachodniopomorskie, Lubuskie and Dolnoslaskie Voivodeships) and Central (Lodzkie, Swietokrzyskie and Slaskie Voivodeships) Poland between 2005 and 2012. Mazowieckie (with Warsaw) and Wielkopolskie (with Poznań) regions are characterized by the lowest level of the indicator.

Future works will focus on estimation of Europa 2020 indicators at more detailed levels of spatial aggregation, i.e. NUTS 3 or LAU 1. Local authorities demand such information to conduct adequate social policy. However, due to sample sizes at such low level as LAU 1 (380 areas in Poland) area-level models might not be adequate. Instead, unit-level models might be useful, but require access to population unit-level data, e.g. from registers or census.

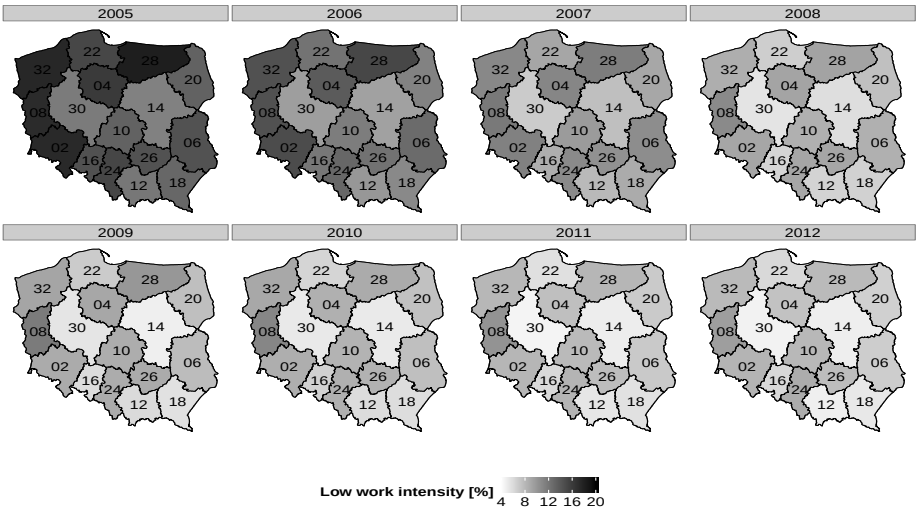

Figure 2: Spatial distribution of low work intensity indicator in Poland between 2005 and 2012 estimated using Marhuenda, Molina and Moralez (2013) model.
Codes on map: 02 – Dolnoslaskie, 04 – Kujawsko-Pomorskie, 06 – Lubelskie, 08 – Lubuskie, 10 – Lodzkie, 12 – Malopolskie, 14 – Mazowieckie, 16 – Opolskie, 18 – Podkarpackie, 20 – Podlaskie, 22 – Pomorskie, 24 – Slaskie, 26 – Swietokrzyskie, 28 – Warminsko-Mazurskie, 30 – Wielkopolskie, 32 – Zachodniopomorskie.

References

Brown, G., Chambers, R., Heady, P., Heasman, D., (2001). Evaluation of small area estimation methods – an application to unemployment estimates from the UK LFS. *Proceedings of Statistics Canada Symposium*.

Bivand, R., Pebesma, E., Gomez-Rubio, V., (2013). *Applied spatial data analysis with R, Second edition.* Springer. New York.

Chandra, H., Chambers, R., Salvati, N., (2012). Small area estimation of proportions in business surveys. *Journal of Statistical Computation and Simulation*, 82(6), pp. 783–795.

Chatterjee, S., Price, B., (1991). *Regression Diagnostics*, New York: John Wiley.

Dehnel, G., Wawrowski, Ł., (2020). Robust estimation of wages in small enterprises: the application to Poland's districts. *Statistics in Transition New Series*, 21(1), pp. 137–157.

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

Fay, R. E., Diallo, M. S., (2012). Small Area Estimation Alternatives for the National Crime Victimization Survey. *Proceedings of the Section on Survey Research Methods. American Statistical Association*, pp. 3742–3756.

Fay, R. E., Diallo, M. S., (2015). *sae2: Small Area Estimation: Time-series Models*. R package.

Fay, R. E., Herriot, R. A., (1979). Estimates of income for small places: an application of James-Stein procedures to census data. *Journal of the American Statistical Association*, 74, pp. 269–277.

Fay, R. E., Planty, M., Diallo, M. S., (2013). Small area estimates from the national crime victimization survey. *Proceedings of the Section on Survey Research Methods. American Statistical Association*, pp. 1544–1557.

Fox, J., Monette, G., (1992). Generalized collinearity diagnostics. *Journal of the American Statistical Association*, 87, pp. 178–183.

GonzáLez-Manteiga, W., LombardíA, M., Molina, I., Morales, D., SantamaríA, L., (2008). Bootstrap mean squared error of a small-area EBLUP. *Journal of Statistical Computation and Simulation*, 78, pp. 443–462.

Greven, S., Kneib, T., (2010). On the behaviour of marginal and conditional AIC in linear mixed models. *Biometrika*, 97(4), pp. 773–789.

Henderson, C. R., (1975). Best linear unbiased estimation and prediction under a selection model. *Biometrics* 31, pp. 423–447.

Li, J., Valliant, R., (2015). Linear Regression Diagnostics in Cluster Samples. *Survey Methodology*, 31, pp. 61–75.

Liao, D., Valliant, R., (2012)., Variance inflation factors in the analysis of complex survey data. *Survey Methodology*, 38, pp. 53–62.

López-Vizcaíno, E., Lombardía, M. J., Morales, D., (2013). Multinomial-based small area estimation of labour force indicators. Statistical Modelling,13(2), pp. 153–178, doi:10.1177/1471082X13478873.

Marhuenda, Y., Molina, I., Morales, D., (2013). Small area estimation with spatio-temporal Fay–Herriot models. *Computational Statistics and Data Analysis*, 58, pp. 308–325.

Mélina, A., Emilio, D. M., (2012). 23% of EU citizens were at risk of poverty or social exclusion in 2010. *Statistics in Focus*, 9, pp. 1–7.

Molina, I., Marhuenda, Y., (2015). sae: An R package for small area estimation. *The R Journal*, 7(1), pp. 81–98.

Molina, I., Rao, J. N. K., (2010). Small area estimation of poverty indicators. *The Canadian Journal of Statistics*, 38(3), pp. 369–385.

Molina, I., Salvati, N., Pratesi, M., (2009). Bootstrap for estimating the MSE of the Spatial EBLUP. *Computational Statistics*, 24, pp. 441–458.

Nakagawa, S., Schielzeth, H., (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, 4, pp. 133–142.

Prasad, N., Rao, J. N. K., (1990). The estimation of the mean squared error of small-area estimators. *Journal of the American Statistical Association*, 85, pp. 163–171.

Rao, J. N. K., Molina, I., (2015). *Small Area Estimation*. Wiley and Sons.

Rao, J. N. K., Yu, M., (1994). Small-area estimation by combining time-series and cross-sectional data. *Canadian Journal of Statistics*, 22, pp. 511–528.

Saefken, B., Ruegamer, D., Kneib, T., Greven, S., (2018). Conditional Model Selection in Mixed-Effects Models with cAIC4, *ArXiv e-prints 1803.05664*.

Statistics Poland, (2014). *Incomes and living conditions of the population of Poland —- report from the EU-SILC survey of 2012*. Statistical Publishing Establishment. Warsaw.

Szymkowiak, M., Młodak, A., Wawrowski, Ł., (2017). Mapping poverty at the level of subregions in Poland using indirect estimation. *Statistics in Transition new series*, 18(4), pp. 609–635.

Viechtbauer, W., (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, 36(3), pp. 1–48.

Ward, T., Ozdemir, E., (2013). Measuring low work intensity — an analysis of the indicator. *ImPRovE Discussion Paper*, 13, pp. 1–37.