Available online at www.isas.org.in/jisas
JOURNAL OF THE INDIAN SOCIETY OF
AGRICULTURAL STATISTICS 66(1) 2012 125-133

Small Area Poverty Estimation by Model Calibration

Risto Lehtonen[1*] and Ari Veijanen[2]
[1]*University of Helsinki, Finland, and* [2] *Statistics Finland*

Received 02 May 2011; Revised 25 September 2011; Accepted 25 September 2011

SUMMARY

Calibration techniques using auxiliary data offer efficient tools for design-based estimation of population totals and means. In linear or model-free calibration, the weights are calibrated to reproduce the known population totals of the auxiliary variables. A key property of model calibration is that the weights are calibrated to the population total of the predictions derived via a specified model. We introduce model calibration methods for estimation of poverty rate for domains and small areas and present some new semi-direct and semi-indirect calibration estimators. They benefit from spatial correlations of variables in a hierarchy of regions or spatial neighbourhoods. Our study variable is binary and we use logistic mixed models under unequal probability sampling. The properties (design bias and accuracy) of the estimators are compared with generalized regression estimators and Horvitz-Thompson type estimators by using simulation experiments with unit-level register data of Statistics Finland.

Keywords : Small area estimation, Poverty rate, Spatial statistics, Mixed models.

1. INTRODUCTION

A calibration estimator is a weighted sum of observed values of the survey variable, with the weights constructed by using information about auxiliary variables (Huang and Fuller 1978, Deville and Särndal 1992, Kott 2006, 2009, Särndal 2007). A calibration equation is imposed: the weighted sample sums of auxiliary variables reproduce the corresponding known population sums. In other words, if the estimator is applied to an auxiliary variable, the known marginal total is obtained. This property is important in the production of official statistics. Calibration estimators are often considered nearly design unbiased (the design bias is, under mild conditions, an asymptotically insignificant contribution to the estimator's mean squared error, Särndal, 2007, p. 99). Accuracy improvement can be expected if there is association between the study variable and the auxiliary variables. Classical calibration methods are often called model-free (Särndal 2007), that is, it is not necessary to specify an assisting model, whereas an explicit model is imposed in model calibration.

In domain estimation, estimates are required for subgroups of population called domains. Examples of domains are regions and demographic subdivisions within regions. Estimation for small domains, commonly known as Small Area Estimation (SAE), is discussed for example in Rao (2003) and Datta (2009). Lehtonen and Veijanen (2009) review model-free calibration and generalized regression estimators in the context of domain estimation.

In *model calibration* introduced by Wu and Sitter (2001) and Wu (2003), a model is first fitted to the sample. Calibration weights are determined using the fitted values instead of the original auxiliary variables: the weighted sample sum of fitted values must equal the population sum of predictions. Chandra and Chambers (2011) discuss model calibration for skewed data. Montanari and Ranalli (2005a) introduce model

**Corresponding author* : Risto Lehtonen
E-mail address : risto.lehtonen@helsinki.fi

calibration estimators that use nonparametric regression methods and Montanari and Ranalli (2005b) discuss multiple model calibration. Lehtonen *et al.* (2009) introduce model calibration for domain estimation. Our calibration equations specify that the weighted sum of fitted values over a subgroup of the sample equals the sum of predictions over the corresponding population subgroup.

A model calibration procedure for domain estimation consists of two phases. In the *modelling phase*, a model is specified and predictions are calculated for population elements by using estimated model parameters and known values of auxiliary variables. There is much flexibility in the model choice. We have chosen a mixed model formulation involving components that account for spatial heterogeneity in the population (Lehtonen *et al.* 2005). The predictions are used in the *calibration phase* when constructing calibration equation and a calibrated domain estimator. There are different options also in this phase. Calibration can be defined at the population level, at the domain level or at an intermediate level, for example at a regional level (neighbourhood) that contains the domain of interest. Further, in the construction of the calibrated domain estimator, a "semi-direct" approach involves using observations only from the domain of interest, whereas in a "semi-indirect" approach, also observations outside the domain of interest are included.

In constructing a calibration estimator for a domain of interest, we consider the option of "borrowing strength" by using information on study variable from other domains deemed similar. If the values of the study variable in neighbouring domains are positively correlated, accuracy may be improved by incorporating observations from related domains into the estimator. Such estimators are called *indirect* in contrast with *direct* estimators that only contain sample values of study variable from the domain of interest (Federal Committee on Statistical Methodology, 1993). Estevao and Särndal (2004) have argued that in a general class of design-based model-free calibration estimators, the optimal estimator uses data on study variable solely from the domain of interest, without an attempt to borrow strength from the other domains. We introduce here calibration estimators that do not belong to the class of estimators discussed in Estevao and Särndal (2004). We discuss new indirect model calibration estimators that aim to borrow strength from neighbouring domains or larger regions with a property of imposing small calibration weights outside the domain of interest. Model-based indirect estimators that borrow strength in terms of responses are not studied.

As an application of model calibration, we consider the estimation of poverty rate for regions. Poverty rate is defined as the proportion of poor people, with income below or at a threshold called poverty line. Domain estimation of poverty indicators has been recently studied by D'Alo *et al.* (2006), Fabrizi *et al.* (2005, 2007a, 2007b), Tzavidis *et al.* (2007), Giusti *et al.* (2009), Molina and Morales (2009), Srivastava (2009), Haslett *et al.* (2010) and Molina and Rao (2010). The estimation of poverty indicators for population subgroups has been investigated extensively in certain international research projects funded under the European Commission Framework Programmes. Recent examples are the AMELI project (Advanced Methodology for European Laeken Indicators, Lehtonen *et al.* 2011) and the SAMPLE project (Small Area Methods for Poverty and Living Condition Estimates); EURAREA (Enhancing small area estimation techniques to meet European needs) is a related former EU funded project on small area estimation of poverty related indicators.

2. NOTATION AND MODELS

The fixed and finite population of interest is denoted $U = \{1, 2,..., k ,..., N\}$, where k refers to the label of population element. A *domain* is a subset of population U such as a regional population. The number of units in the domain is denoted by N_d. In sample s, the corresponding subset is defined as $s_d = U_d \cap s$; it has n_d observations. A small area is a domain whose realized sample size is small (even zero). The design weights a_k are inverses of inclusion probabilities π_k of the sampling design ($a_k = 1/\pi_k$). The domain structure we are considering is of an unplanned type (Lehtonen and Veijanen, 2009, p. 222).

The domain total of a study variable y is defined by

$$t_d = \sum_{k \in U_d} y_k , \tag{1}$$

where y_k denotes the value of the study variable for element k.

In order to account for possible differences between regions, a mixed model incorporates domain-specific random effects $u_d \sim N(0, \sigma_u^2)$ for domain U_d, or regional random effects $u_r \sim N(0, \sigma_u^2)$ for region U_r, where $U_d \subset U_r$. For domain-specific random intercepts, a linear mixed model is given by

$$Y_k = \mathbf{x}_k'\boldsymbol{\beta} + u_d + \varepsilon_k,\ k \in U_d,\ \varepsilon_k \sim N(0, \sigma^2).$$

For a binary y-variable, a logistic mixed model is of the form

$$E_m(y_k \mid u_d) = P\{y_k = 1 \mid u_d; \boldsymbol{\beta}\} = \frac{\exp(\mathbf{x}_k'\boldsymbol{\beta} + u_d)}{1+\exp(\mathbf{x}_k'\boldsymbol{\beta} + u_d)},$$

where $\mathbf{x}_k$ is a known vector value for every $k \in U$ and $\boldsymbol{\beta}$ is a vector of fixed effects common for all domains. The parameters $\boldsymbol{\beta}$, σ_u^2 and σ^2 are first estimated from the data, and the values of the random effects are then predicted. Predictions $\hat{y}_k = P\{y_k = 1 \mid \hat{u}_d; \hat{\boldsymbol{\beta}}\}$ are calculated for every $k \in U$. Lehtonen *et al.* (2005) give several special cases of the model. Jiang and Lahiri (2006) discuss predictors based on mixed models.

3. DESIGN-BASED ESTIMATORS

Horvitz-Thompson (HT) estimator (also Narain 1951) of domain total is a weighted sum of values in the sample:

$$\hat{t}_d = \sum_{k \in s_d} a_k y_k. \tag{2}$$

This is a direct estimator as it only involves observations from the domain of interest. The estimator is design unbiased but it can have large variance, especially for small domains. HT does not incorporate any auxiliary data.

Generalized regression (GREG) estimators (Särndal *et al.* 1992; Lehtonen and Veijanen 2009) are assisted by a model fitted to the sample. By choosing different models we obtain a family of GREG estimators with same form but different predicted values (Lehtonen *et al.* 2003, 2005).

Ordinary GREG estimator

$$\hat{t}_{d;GREG} = \sum_{k \in U_d} \hat{y}_k + \sum_{k \in s_d} a_k (y_k - \hat{y}_k) \tag{3}$$

incorporating a linear fixed-effects regression model is often used to estimate domain totals (1) of a continuous study variable. For a binary or polytomous response variable, a linear model formulation will not necessarily fit the data well. A logistic model formulation might be a more realistic choice. LGREG (logistic GREG; Lehtonen and Veijanen 1998) estimates the frequency f_d of a class C in each domain. A logistic regression model is fitted to indicators $v_k = I\{y_k \in C\}$, $k \in s$, using the design weights. The fitted model yields estimated probabilities $\hat{p}_k = P\{v_k = 1; \mathbf{x}_k, \hat{\boldsymbol{\beta}}\}$. The LGREG estimator of the class frequency in U_d is

$$\hat{f}_{d;LGREG} = \sum_{k \in U_d} \hat{p}_k + \sum_{k \in s_d} a_k (v_k - \hat{p}_k). \tag{4}$$

Here the first term is the sum of predicted values over the population domain. The calculation of $\hat{p}_k$ for all $k \in U$ requires access to unit-level population data on auxiliary variables. The last component of (4), i.e. an HT estimator of the residual total, aims at correcting the possible bias of the first part (synthetic estimator of (1)).

In the MLGREG estimator (Lehtonen and Veijanen 1999; Lehtonen *et al.* 2005; Torabi and Rao 2008), we use an alternative logistic mixed model for (4) involving fitted values $\hat{p}_k = P\{v_k = 1 \mid \hat{\mathbf{u}}_d; \mathbf{x}_k, \hat{\boldsymbol{\beta}}\}$ instead of the fixed-effects logistic model. The random effects are associated with domains U_d or with larger regions U_r.

4. MODEL CALIBRATION

In population level calibration (Wu and Sitter 2001), the weights must satisfy calibration equation

$$\sum_{i \in s} w_i z_i = \sum_{i \in U} z_i = \left(N, \sum_{i \in U} \hat{y}_i\right), \tag{5}$$

where $z = (1, \hat{y}_i)$. Using the technique of Lagrange multiplier (λ), we minimize

$$\sum_{k \in s} \frac{(w_k - a_k)^2}{a_k} - \lambda'\left(\sum_{i \in s} w_i z_i - \sum_{i \in U} z_i\right)$$

subject to the conditions (5). The first part of the equation is the distance between the weights w_k and the known design weights a_k. The latter part corresponds to the constraints (5). The equation is minimized by weights

$$w_k(\lambda) = a_k(1 + \lambda' z_k), \tag{6}$$

where

$$\lambda = \left(\sum_{i \in U} z_i - \sum_{i \in s} a_i z_i\right)' \left(\sum_{i \in s} a_i z_i z_i'\right)^{-1}.$$

In domain estimation, these weights are applied over a domain: the estimator is

$$\hat{f}_{d;pop} = \sum_{k \in s_d} w_k y_k. \tag{7}$$

A straightforward generalization of the population-level calibration equation is a domain-level calibration equation

$$\sum_{i \in s_d} w_{di} z_i = \sum_{i \in U_d} z_i = \left(N_d, \sum_{i \in U_d} \hat{y}_i\right), \tag{8}$$

where the weights w_{di} are specific to the domain. From (8) we see that the domain sizes must be known. We minimize

$$\sum_{i \in s_d} \frac{(w_{dk} - a_k)^2}{a_k} - \lambda_d' \left(\sum_{i \in s_d} w_{di} z_i - \sum_{i \in U_d} z_i\right)$$

subject to the calibration equations (8). The solution is $w_{dk} = w_k(\lambda_d)$, defined by (6) for

$$\lambda_d = \left(\sum_{i \in U_d} z_i - \sum_{i \in s_d} a_i z_i\right)' \left(\sum_{i \in s_d} a_i z_i z_i'\right)^{-1}.$$

The domain estimator is then a weighted domain sum

$$\hat{f}_{d;s} = \sum_{k \in s_d} w_{dk} y_k. \tag{9}$$

We call this estimator "semi-direct", as the sum only contains y-observations from the domain of interest. It is not a direct estimator, however, as the weights are determined by a model that is fitted to the whole sample. Sum of estimators (9) over domains is not necessarily equal to the model calibration estimator for the population; this property holds for estimators (7).

The domain totals of auxiliary variables are not always known in calibration. This is the main motivation of Estevao and Särndal (2004) to define calibration using marginal totals known over calibration groups, such as larger regions. In the context of model calibration, the calibration equation is defined over domain-dependent sets $C_d \supset U_d$ and $r_d = C_d \cap s$:

$$\sum_{i \in r_d} w_{di} z_i = \sum_{i \in C_d} z_i. \tag{10}$$

The weights are $w_{Ck} = w_k(\lambda_{Cd})$ defined by (6) for

$$\lambda_{Cd} = \left(\sum_{i \in C_d} z_i - \sum_{i \in r_d} a_i z_i\right)' \left(\sum_{i \in r_d} a_i z_i z_i'\right)^{-1}.$$

The semi-direct estimator is then a weighted domain sum

$$\hat{y}_{Cd} = \sum_{k \in s_d} w_{Ck} y_k. \tag{11}$$

We introduce next various new "semi-indirect" estimators. They are weighted sums over a set that is larger than the domain of interest. Our goal is to "borrow strength" from other domains, in an attempt to reduce mean squared error. A semi-indirect domain estimator incorporates whole sample, an enclosing aggregate of regions in a hierarchy of regions or the set of neighbouring domains, including the domain itself. A neighbourhood of a region comprises regions that share a common border with the specified region or regions with centre closer than a given distance threshold. Strictly speaking, the neighbour relation does not have to be symmetric. Neighbourhoods may utilize spatial correlations better than a more rigid hierarchy of regions.

In a semi-indirect estimator, we use, as in (10), supersets $C_d \supset U_d$ of domains with corresponding sample subsets $r_d = C_d \cap s$. They could contain genuine subsets of other domains, but in our simulations they are composed of domains. In contrast with (11), we define the domain estimator as a weighted sum of all observations in r_d:

$$\hat{f}_{d;r} = \sum_{k \in r_d} w_{dk} y_k \tag{12}$$

The calibration equation is

$$\sum_{i \in r_d} w_{di} z_i = \sum_{i \in U_d} z_i. \tag{13}$$

Note that the sum on the left side of (13) extends over r_d which corresponds to population subset C_d, a larger set than U_d on the right side of the equation. To satisfy (13), the weights must be smaller than in (11). In order to reduce design bias, the weights should be larger in the domain than outside it. We have required that the weights w_{dk} are close to weights a_k in the domain and close to zero outside the domain. In other words, the weights should be close to $I\{k \in s_d\}a_k = I_{dk}a_k$ ($I_{dk} = I\{k \in s_d\}$). The weights minimize

$$\sum_{k \in r_d} \frac{(w_{dk} - I_{dk}a_k)^2}{a_k}$$

subject to the calibration equations (13) when

$$w_{dk} = I_{dk}a_k + \lambda_d' a_k z_k;$$

$$\lambda_d = \left(\sum_{i \in U_d} z_i - \sum_{i \in r_d} I_{di}a_i z_i\right)' \left(\sum_{i \in r_d} a_i z_i z_i'\right)^{-1}.$$

This idea is similar to ideas presented by Singh and Mian (1995) in the context of calibration based on auxiliary variables.

5. AT-RISK-OF POVERTY RATE

Poverty rate is defined in terms of equivalized income, a household's total disposable income divided by its "equivalent size". It is attributed to each household member, including children (European Commission, 2006). Equivalized size of a household is a sum of weights of its members. The OECD modified scale assigns weight 1.0 for the first adult, 0.5 for every additional person aged 14 or over, and 0.3 for every child under 14.

At-risk-of-poverty threshold, or poverty line, is 60% of the median equivalized income of persons in the whole population. People whose income is below or at the poverty line are here referred to as "poor". To estimate the *reference median income M*, we first present the HT estimator of the distribution function of equivalized income in the whole population. The distribution function of y in U is

$$F_U(t) = \frac{1}{N} \sum_{k \in U} I\{y_k \le t\}.$$

This is estimated by HT:

$$\hat{F}_U(t) = \frac{1}{\hat{N}} \sum_{k \in s} a_k I\{y_k \le t\},$$

where the estimated population size is $\hat{N} = \sum_{k \in s} a_k$.

$\hat{M}$ is obtained from $\hat{F}_U$ as the smallest y_k ($k \in s$) for which $\hat{F}_U(y_k) > 0.5$. In the special case $\hat{F}_U(y_{(k)}) = 0.5$ for kth observation in sorted y, the median is the average of $y_{(k)}$ and $y_{(k+1)}$.

At-risk-of-poverty rate is the proportion of poor people in a domain with equivalized income at or below the poverty line. Our goal is to estimate

$$R_d = \frac{1}{N_d} \sum_{k \in U_d} I\,\{y_k \le 0.6M\}.$$

HT-CDF estimator of poverty rate is based on the HT estimator of the distribution function. The distribution function defined in domain U_d is estimated by

$$\hat{F}_d(t) = \frac{1}{\hat{N}_d} \sum_{k \in s_d} a_k I\{y_k \le t\},$$

where $\hat{N}_d = \sum_{k \in s_d} a_k$. The poverty rate is then estimated by

$$\hat{r}_{d;HT} = \hat{F}_d(0.6\hat{M}). \quad (14)$$

To estimate domain poverty rate by LGREG, MLGREG or model calibration, we first estimate the domain total of a *poverty indicator* $v_k = I\{y_k \le 0.6\ \hat{M}\}$, which equals 1 for persons with income below or at the poverty line and 0 for others. The estimate of the domain total t_d is then divided by the known domain size N_d (or, its estimate $\hat{N}_d$). For example, the MLGREG estimator of the poverty rate is

$$\hat{r}_{d;MLGREG} = \frac{\hat{f}_{d;MLGREG}}{N_d}. \quad (15)$$

6. MONTE CARLO EXPERIMENTS

Design bias and accuracy of estimators of poverty rate were examined by design-based simulation techniques. An artificial population of one million persons was constructed from real income data of seven NUTS (European Union's Nomenclature of Territorial Units for Statistics) level 3 regions in Western Finland. Household attributes such as demographic composition and equivalized income were obtained from registers maintained by Statistics Finland (the attributes of the household head were obtained from a sample but some auxiliary information of other household members had to be imputed). The population was still realistic enough for a simulation study.

In the simulations, $K = 1000$ samples of $n = 5000$ persons were drawn with without-replacement probability proportional to size (PPS) sampling from the unit-level population. For PPS, an artificial size variable was generated as a function of the socio-economic status of household head. People with low income appear in samples with larger probability than people with large income.

Our models incorporated the following auxiliary variables: age class (0-15, 16-24, 25-49, 50-64, 65-years), gender with interactions with age class, socio-economic status of the household head (wage and salary earners, farmers, other entrepreneurs, pensioners, and others), and labour force status (employed, unemployed, and not in workforce). We created indicators for each class of a qualitative variable. The models were fitted by R function glmer (package lme4), with design weights incorporated into the fitting procedure. As domains we used the 36 NUTS4 regions. The NUTS classification is hierarchical: each NUTS4 region is contained within a larger NUTS3 region.

From each sample, the following quality indicators were calculated for each domain estimator: absolute relative bias

$$ARB = \frac{\left| \frac{1}{K} \sum_{k=1}^{K} (\hat{\theta}_{dk} - \theta_d) \right|}{\theta_d}$$

and relative root mean squared error

$$RRMSE = \frac{\sqrt{\frac{1}{K} \sum_{k=1}^{K} (\hat{\theta}_{dk} - \theta_d)^2}}{\theta_d} .$$

We present the averages of RRMSE over domain classes defined by expected domain sample size: Minor (0-50 units), Medium-sized (50-100) and Major (100-) domains.

We compare three different models (Table 1). In the common logistic fixed-effects model (a), there are no domain-specific terms. The logistic mixed model (b) contains regional random intercepts associated with NUTS3 regions. The mixed model (c) contains domain-level (NUTS4) random intercepts. Design weights were incorporated into the fitting of each model.

In Table 1, we use the following labels for the methods:

Default	Direct HT-based estimator (14)
LGREG	Indirect GREG estimator (15) assisted by a logistic fixed-effects model
MLGREG	Indirect GREG estimator (15) assisted by a logistic mixed model
SD-population	Semi-direct estimator (7) based on the original method of Wu and Sitter (2001)
SD-domain	Semi-direct estimator (9) incorporating calibration at domain level
SD-regional	Semi-direct estimator (11) incorporating calibration (10) over NUTS3 regions
SD-spatial	Semi-direct estimator (11) based on calibration over a neighbourhood
SI-population	Semi-indirect estimator (12) defined as a weighted sum over the whole sample
SI-regional	Semi-indirect estimator (12) over enclosing NUTS3 region
SI-spatial	Semi-indirect estimator (12) over a neighbourhood containing regions sharing a common border

Table 1. Mean relative root mean squared error (RRMSE) (%) of poverty rate estimators over domain size classes, under three different logistic model formulations.

Estimator	Expected domain sample size			All
	Minor	Medium	Major	
(a) Logistic fixed-effects (common) model				
Default	41.1	28.9	18.0	26.7
LGREG	39.6	28.7	17.8	26.3
Model calibration estimators				
Semi-direct estimators				
SD-population	42.7	29.4	18.5	27.3
SD-domain	40.4	28.5	17.8	26.3
SD-regional	42.6	29.3	18.2	27.1
SD-spatial	42.5	29.4	18.3	27.2
Semi-indirect estimators				
SI-population	39.7	28.7	17.8	26.3
SI-regional	39.7	28.6	17.8	26.2
SI-spatial	39.7	28.6	17.8	26.2
(b) Logistic mixed model (NUTS3 level)				
MLGREG	39.6	28.6	17.8	26.3
Model calibration estimators				
Semi-direct estimators				
SD-population	42.7	37.9	26.4	34.4
SD-domain	40.4	28.5	17.8	26.3
SD-regional	42.6	29.3	18.2	27.1
SD-spatial	42.5	29.4	18.3	27.2
Semi-indirect estimators				
SI-population	39.7	28.6	17.8	26.3
SI-regional	39.7	28.6	17.8	26.2
SI-spatial	39.7	28.6	17.8	26.2
(c) Logistic mixed model (NUTS4 level)				
MLGREG	39.6	28.6	17.8	26.2
Model calibration estimators				
Semi-direct estimators				
SD-population	42.8	43.3	33.4	39.7
SD-domain	40.4	28.6	17.8	26.3
SD-regional	42.6	29.2	18.2	27.1
SD-spatial	42.4	29.3	18.3	27.2
Semi-indirect estimators				
SI-population	39.7	28.6	17.8	26.2
SI-regional	39.7	28.5	17.8	26.2
SI-spatial	39.7	28.6	17.8	26.2

Calibration at domain level (8) results in a better estimator than calibration at population level (5). All the methods were nearly design unbiased: maximum ARB over all model choices and domain size classes was below 1.65% (results of ARB not shown). The semi-direct estimators calibrated at population level (5), regional or neighbourhood level (10) did not perform better than the default method. The semi-indirect methods were usually more successful.

The choice of the model did not have much effect on most estimators. In semi-direct method, population-level estimators showed bad performance when using mixed model with regional random intercepts. The model calibration estimators did not outperform the generalized regression estimators.

7. CONCLUSION

At least in the smallest areas, the semi-indirect model calibration estimators involving calibration at the level of regions or neighbourhoods (13) had smaller RRMSE than the semi-direct estimators (9) calibrated at domain level. The semi-indirect methods were also better than corresponding semi-direct methods based on calibration equation (10). Thus, borrowing strength by our approach may be a reasonable strategy, given that the domain sums of predictions are known. However, when the domain sums are not known, the methods of Estevao and Särndal (2004) are still applicable.

The differences between the semi-indirect methods are small. If there is a lot of spatial variation, calibration equation based on neighbourhoods might be preferred. The best model calibration methods were not significantly better than the generalized regression estimator assisted by logistic mixed model. A technique using spatial mixed models with correlated random effects (e.g. Chandra 2009) would provide an alternative to the model calibration methods discussed in this paper.

ACKNOWLEDGEMENT

The work of the second author was supported by grant of European Commission under the AMELI project. Constructive comments of a referee were helpful in finalizing the paper.

REFERENCES

AMELI (Advanced Methodology for European Laeken Indicators). http://www.ameli.surveystatistics.net

Chandra, H. (2009). Estimation of small area proportions under unit level spatial models. *J. Ind. Soc. Agril. Statist.,* **63(3)**, 267-276.

Chandra, H. and Chambers, R. (2011). Small area estimation under transformation to linearity. *Survey Methodology*, **37(1)**, 39-51.

D'Alo, M., Di Consiglio, L., Falorsi, S. and Solari, F. (2006). Small area estimation of the Italian poverty rate. *Statist. Trans.,* **7**, 771-784.

Datta, G.S. (2009). Model-based approach to small area estimation. In: C. R. Rao and D. Pfeffermann (eds.), *Handbook of Statistics,* **29(B)***. Sample Surveys: Theory, Methods and Inference.* Elsevier, Amsterdam.

Deville, J.C. and Särndal, C.E. (1992). Calibration estimators in survey sampling. *J. Amer. Statist. Assoc.*, **87**, 376-382.

Estevao, V.M. and Särndal, C.E. (2004). Borrowing strength is not the best technique within a wide class of design-consistent domain estimators. *J. Off. Statist.*, **20**, 645-669.

EURAREA (Enhancing small area estimation techniques to meet European needs). http://www.statistics.gov.uk/eurarea/

Fabrizi, E., Ferrante, M.R. and Pacei, S. (2005). Estimation of poverty indicators at sub-national level using multivariate small area models. *Statist. Trans.*, **7**, 587-608.

Fabrizi, E., Ferrante, M.R. and Pacei, S. (2007a). Small area estimation of average household income based on unit level models for panel data. *Survey Methodology*, **33**, 187-198.

Fabrizi, E., Ferrante, M.R. and Pacei, S. (2007b). Comparing alternative distributional assumptions in mixed models used for small area estimation of income parameters. *Statist. Trans.*, **8**, 423-439.

Federal Committee on Statistical Methodology (1993). Indirect Estimators in Federal Programs. U.S. Office of Management and Budget, Statistical Policy Working Paper 21.

Giusti, C., Pratesi, M. and Salvati, N. (2009). Small area methods in the estimation of poverty indicators: The case of Tuscany. *Politica Economica*, 369-381.

Haslett, S.J., Isidro, M.C. and Jones, G. (2010). Comparison of survey regression techniques in the context of small area estimation of poverty. *Survey Methodology*, **36**, 157-170.

Huang, E.T. and Fuller, W.A. (1978). Nonnegative regression estimation for sample survey data. In: *Proc. Social Statist. Sect., Am. Statist. Assoc.*, 300-305. Washington, DC.

Jiang, J. and Lahiri, P. (2006). Mixed model prediction and small area estimation. *Sociedad de Estadistica e Investigacion Operative Test*, **15**, 1-96.

Kott, P.S. (2006). Using calibration weighting to adjust for nonresponse and coverage errors. *Survey Methodology*, **32**, 133-142.

Kott, P.S. (2009). Calibration weighting: Combining probability samples and linear prediction models. In: C.R. Rao and D. Pfeffermann (eds.), *Handbook of statistics,* **29(B)***. Sample Surveys: Theory, Methods and Inference.* Elsevier, Amsterdam.

Lehtonen, R., Särndal, C.-E., and Veijanen, A. (2003). The effect of model choice in estimation for domains, including small domains. *Survey Methodology,* **29**, 33-44.

Lehtonen, R., Särndal, C.-E., and Veijanen, A. (2005). Does the model matter? Comparing model-assisted and model-dependent estimators of class frequencies for domains. *Statist. Trans.,* **7**, 649-673.

Lehtonen, R., Särndal, C.-E. and Veijanen, A. (2009). Model calibration and generalized regression estimation for domains and small areas. Invited paper, SAE2009 Conference on Small Area Estimation, 29 June – 1 July, 2009, Elche, Spain.

Lehtonen, R. and Veijanen, A. (1998). Logistic generalized regression estimators. *Survey Methodology,* **24**, 51-55.

Lehtonen, R. and Veijanen, A. (1999). Domain estimation with logistic generalized regression and related estimators. IASS Satellite Conference on Small Area Estimation, Riga: Latvian Council of Science, 121-128.

Lehtonen, R. and Veijanen, A. (2009). Design-based methods of estimation for domains and small areas. In: C.R. Rao and D. Pfeffermann (eds.), *Handbook of Statistics,* **29(B)**. *Sample Surveys: Theory, Methods and Inference.* Elsevier, Amsterdam.

Lehtonen, R., Veijanen, A., Myrskylä, M. and Valaste, M. (2011). Small Area Estimation of Indicators on Poverty and Social Exclusion. AMELI WP2 Deliverable 2.2. Available at: http://www.uni-trier.de/index.php?id=24676&L=2

Molina, I. and Morales, D. (2009). Small area estimation of poverty indicators. *Boletín de Estadística e Investigación Operativa* BEIO, **25(3)**, 218-225.

Molina, I. and Rao, J.N.K. (2010). Small area estimation of poverty indicators. *Canad. J. Statist.,* **38**, 369-385.

Montanari, G.E. and Ranalli, M.G. (2005a). Nonparametric model calibration estimation in survey sampling. *J. Amer. Statist. Assoc.,* **100**, 1429–1442.

Montanari, G.E. and Ranalli, M.G. (2005b). Multiple model calibration in survey sampling. In: *Metodi Statistici per l'Integrazione di Dati da Fonti Diverse, Franco Angeli,* 273-298.

Narain, R.D. (1951). On sampling without replacement with varying probabilities. *J. Ind. Soc. Agril. Statist.,* **3**, 169-175.

Rao, J.N.K. (2003). *Small Area Estimation*. John Wiley & Sons, New York.

SAMPLE (Small Area Methods for Poverty and Living Condition Estimates). http://www.sample-project.eu/

Singh, A.C. and Mian, I.U.H. (1995). Generalized sample size dependent estimators for small areas. *Proc. Annual Res. Conf., U.S. Bureau of the Census*, 687-701.

Srivastava, A.K. (2009). Some aspects of estimating poverty at small area level. *J. Ind. Soc. Agril. Statist.,* **63(1)**, 1-23.

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

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

Torabi, M. and Rao, J.N.K. (2008). Small area estimation under a two-level model. *Survey Methodology*, **34**, 11-17.

Tzavidis, N., Salvati, N., Pratesi, M. and Chambers, R. (2007). M-quantile models for poverty mapping. *Statist. Methods Appl.*, **17**, 393-411.

Wu, C. (2003). Optimal calibration estimators in survey sampling. *Biometrika*, **90**, 937-951.

Wu, C. and Sitter, R.R. (2001). A model-calibration approach to using complete auxiliary information from survey data. *J. Amer. Statist. Assoc.*, **96**, 185-193.