Journal of the Royal Statistical Society
Statistics in Society
Series A

J. R. Statist. Soc. A (2017)
180, *Part* 4, *pp.* 1111–1136

Poverty mapping in small areas under a twofold nested error regression model

Yolanda Marhuenda,

Universidad Miguel Hernández de Elche, Spain

Isabel Molina,

Universidad Carlos III de Madrid, Spain

Domingo Morales

Universidad Miguel Hernández de Elche, Spain

and J. N. K. Rao

Carleton University, Ottawa, Canada

[Received April 2016. Final revision June 2017]

Summary. Poverty maps at local level might be misleading when based on direct (or area-specific) estimators obtained from a survey that does not cover adequately all the local areas of interest. In this case, small area estimation procedures based on assuming common models for all the areas typically provide much more reliable poverty estimates. These models include area effects to account for the unexplained between-area heterogeneity. When poverty figures are sought at two different aggregation levels, domains and subdomains, it is reasonable to assume a twofold nested error model including random effects explaining the heterogeneity at the two levels of aggregation. The paper introduces the empirical best (EB) method for poverty mapping or, more generally, for estimation of additive parameters in small areas, under a twofold model. Under this model, analytical expressions for the EB estimators of poverty incidences and gaps in domains or subdomains are given. For more complex additive parameters, a Monte Carlo algorithm is used to approximate the EB estimators. The EB estimates obtained of the totals for all the subdomains in a given domain add up to the EB estimate of the domain total. We develop a bootstrap estimator of the mean-squared error of EB estimators and study the effect on the mean-squared error of a misspecification of the area effects. In simulations, we compare the estimators obtained under the twofold model with those obtained under models with only domain effects or only subdomain effects, when all subdomains are sampled or when there are unsampled subdomains. The methodology is applied to poverty mapping in counties of the Spanish region of Valencia by gender. Results show great variation in the poverty incidence and gap across the counties from this region, with more counties affected by extreme poverty when restricting ourselves to women.

Keywords: Empirical best estimator; Nested error model; Poverty mapping; Small area estimation; Unit level models

1. Introduction

Maps showing the distribution of poverty at local level are used by many government institutions

Address for correspondence: Domingo Morales, Centro de Investigación Operativa, Universidad Miguel Hernández de Elche, Avenida de la Universidad, Elche 03202, Spain.
E-mail: d.morales@umh.es

 0964–1998/17/1801111

and international organizations to target in a more efficient way social development policies and to allocate funds. For example, the World Bank produces poverty and inequality maps for many countries. The US Census Bureau, through the 'Small area income and poverty estimates' programme, produces estimates of poor school-age children for counties (National Research Council, 2000), and these estimates are used by the US Department of Education to allocate annually several billion dollars of general funds to school districts. European Union (EU) regional policies also rely increasingly on detailed regional statistics; in particular, poverty maps. In Mexico, the National Council for the Evaluation of the Social Development Policy has a mandate to produce poverty and inequality estimates at the municipal level. The problem is that the surveys that are used to obtain data on income- and poverty-related issues, e.g. the survey on income and living conditions (SILC) that is carried out in all EU countries, typically have sample sizes that do not allow efficient estimation at local level. For areas that are not well covered by the survey, direct estimates, based only on scarce area-specific data, are not reliable. Therefore, poverty maps based on these estimates can be misleading. Small area estimation techniques link the data from the entire sample to 'borrow strength' from all the areas in the population, leading to much more reliable estimates. For an updated monograph on small area estimation, see Rao and Molina (2015).

When auxiliary information is available, a popular small area estimation approach is the use of regression models that represent the common factors affecting the variable of interest (e.g. income) across the different areas in the population. Auxiliary information might be available in different forms. Sometimes only domain aggregates of auxiliary variables are available. In these cases, small area estimation models have been established at the domain level (see for example Fay and Herriot (1979)). Other times, the values of auxiliary variables are available for all the units in the population in the form of a census or a register. In agriculture or environmental applications, unit level auxiliary information might come from satellite or laser scanner images (see for example Battese *et al.* (1988)). Small area estimators based on unit level models typically achieve very large reductions in mean-squared error (MSE) compared with direct estimation.

Many income-based poverty indicators are complex non-linear functions of the incomes for the individuals. The first method based on a unit level model that was designed to estimate general non-linear indicators in small areas is the traditional method that is used by the World Bank for poverty mapping, due to Elbers *et al.* (2003). Later, Molina and Rao (2010) proposed the empirical best (EB) method, giving the approximately optimal estimator in the sense of minimum MSE under a unit level model with random effects for the areas of interest. Molina *et al.* (2015) proposed a hierarchical Bayes approach for that problem; see Guadarrama *et al.* (2014) for a comparison of poverty mapping methods based on unit level models for income. Other approaches are based on unit level logit mixed models (Hobza and Morales, 2016), on M-quantile regression models (Tzavidis *et al.*, 2008; Chambers *et al.*, 2012, 2016) or temporal and spatiotemporal area level models (Esteban *et al.*, 2012a, b; Marhuenda *et al.*, 2013; Morales *et al.*, 2015). For further references, see Pratesi (2016).

The EB method of Molina and Rao (2010) considers that a one-to-one transformation of income follows the nested error model of Battese *et al.* (1988) with random effects for the domains of interest. In contrast, since many household surveys use two-stage sampling, the Elbers–Lanjouw–Lanjouw (ELL) method (Elbers *et al.*, 2003) considers a nested error model with random effects for the primary sampling units (called clusters in the ELL method), which are typically nested in the domains of estimation. An example of a household survey with two-stage sampling is the SILC that is carried out in Spain, where primary sampling units (clusters) are census tracks, whereas the domains of interest can be the Spanish provinces or counties. In US household surveys, where estimation is desired at the county level (domains of interest), the primary sampling units are often blocks of dwellings (Ghosh and Lahiri, 1988). In India, there is considerable interest in producing crop yield estimates at groups of villages with common administration (called '*tehsil*') and two- or three-stage sampling is used within each *tehsil* (Stukel and Rao, 1999). In these cases, data might present between-cluster variation beyond the domain variation that is already incorporated in the EB method. To account for both types of variation, we consider a twofold nested error model that includes cluster effects nested within the usual domain effects. This model was used by Stukel and Rao (1997) for the estimation of linear parameters such as domain means. Here we extend the EB method of Molina and Rao (2010) to estimate general non-linear parameters under the twofold model.

In fact, the available software for small area estimation of non-linear parameters includes only estimators based on models with random effects at one of the levels (clusters or domains) but not both; see the `ebBHF()` function from R package `sae` (Molina and Marhuenda, 2015). Practitioners might be willing to use this software or to consider a onefold model just for simplicity. Then, at which level should we include the random effects? Is a model with cluster effects alone sufficiently good for estimation at the domain level as in the ELL method, or should we always include random effects for the domains of interest? Is a model with only random effects for the domains sufficiently good, or should we also include the cluster effects? To answer these questions, we compare, in terms of bias and efficiency, EB estimators that are obtained under the twofold model with those obtained under onefold models with random effects either at the cluster level or at the domain level of interest. We also compare with the traditional ELL method, based on a onefold nested error model with random effects for the sampling clusters but not for the domains of interest. A bootstrap procedure is used for MSE estimation and the problem of underestimation of the MSE when the random effects are misspecified is studied for the estimation of domain means. The method proposed is used to estimate income and poverty indicators in counties and provinces from the Spanish region of Valencia by gender.

2. Poverty mapping in Valencia region

The main source of data that is used in the EU countries for a systematic production of statistics on household income, poverty and social exclusion is the SILC. The statistics that are obtained from this survey are comparable within Spain and also across the EU countries. Poverty statistics at local level are especially important for regional policies and allocation of EU funds. The government from the autonomous community of Valencia is especially interested in obtaining income and poverty indicators at the finest geographical level. These statistics are used to inform citizens and to support local policies.

The autonomous community of Valencia is divided into the provinces of Alicante, Castellón and Valencia, which are further divided into '*comarcas*' (which are similar to counties). There are nine *comarcas* in Alicante, eight in Castellón and 17 in the province of Valencia. In this work we deal with the estimation of average income and basic poverty indicators for *comarcas* and provinces from the autonomous community of Valencia. Estimates are desired also by gender for obvious sociological reasons. For each gender, estimates of totals for the *comarcas* from each province should add up to the corresponding estimate of the province total. Thus, in this application, the target areas are the $D=3\times 2=6$ provinces by gender (domains) and the $M=(9+8+17)\times 2=68$ *comarcas* by gender (subdomains).

The available data come from the 2012 SILC for the autonomous community of Valencia. This survey provides information regarding the household income that is received during the year before that of the interview, including income from work for others, benefits or losses from self-employed work, social benefits, income from private pension schemes that are not related to work, capital and property income, transfers between other households, income received by minors and the result of the income tax statement. To account for scale economies in households, the total household income is then divided by the number of consumption units according to the modified scale of the Organisation for Economic Co-operation and Development, which assigns weight 1 to the first adult, 0.5 to the remaining adults and 0.3 to children under 14 years old. The resulting income per household consumption unit, called equivalent personal income, is assigned to each of the household members.

We denote by N_d the population size of domain d, $d=1,\ldots,D$, and $N=\Sigma_{d=1}^{D}N_d$ the total population size. Domain d contains M_d subdomains (*comarcas* by gender), of respective sizes $N_{d1},\ldots,N_{dM_d}$, with $N_d=\Sigma_{t=1}^{M_d}N_{dt}$. Let E_{dtj} denote the equivalent personal income of individual j from subdomain t within domain d, $j=1,\ldots,N_{dt}$, $t=1,\ldots,M_d$ and $d=1,\ldots,D$. The target parameters are, for each subdomain t and domain d, average equivalent personal income given by

$$E_{dt}=\frac{1}{N_{dt}}\sum_{j=1}^{N_{dt}}E_{dtj}, \qquad E_d=\frac{1}{N_d}\sum_{t=1}^{M_d}\sum_{j=1}^{N_{dt}}E_{dtj}, \qquad t=1,\ldots,M_d, \quad d=1,\ldots,D, \tag{1}$$

and poverty incidences and gaps. These two poverty indicators are members of the Foster–Greer–Thorbecke (FGT) family that was introduced by Foster *et al.* (1984) and defined for subdomain t within domain d as

$$F_{\alpha,dt}=\frac{1}{N_{dt}}\sum_{j=1}^{N_{dt}}F_{\alpha,dtj}, \qquad F_{\alpha,dtj}=\left(\frac{z-E_{dtj}}{z}\right)^{\alpha}I(E_{dtj}<z), \qquad \alpha\geqslant 0, \tag{2}$$

where z is the poverty line and $I(E_{dtj}<z)$ equals 1 if $E_{dtj}<z$ and 0 otherwise. For $\alpha=0$, we obtain the poverty incidence, which measures the share of people in subdomain t whose welfare is below the poverty line z. For $\alpha=1$, we obtain the poverty gap, measuring the depth of poverty of the people in that subdomain. For a domain d, the FGT poverty indicator of order α is defined analogously and denoted $F_{\alpha,d}$. Obviously, the totals at the subdomains from the same domain add up to the corresponding estimated domain total, i.e.

$$E_d=\frac{1}{N_d}\sum_{t=1}^{M_d}N_{dt}E_{dt},$$

$$F_{\alpha,d}=\frac{1}{N_d}\sum_{t=1}^{M_d}N_{dt}F_{\alpha,dt}.$$

Estimates for subdomains are required to be consistent with the domain estimates in the sense of satisfying the above restrictions. All estimates should be supplemented with their estimated MSEs.

In the SILC, the poverty line z is calculated for each year by using the distribution of income in the previous year. Following the criteria that are recommended by Eurostat, z is set at 60% of the median equivalent personal income. In the autonomous community of Valencia, the 2012 poverty line is $z=€6840$.

The total sample size of the 2012 SILC for the autonomous community of Valencia is $n=2678$. With this sample size, not all the *comarcas* are covered. From the nine, eight and 17 *comarcas* of Alicante, Castellón and Valencia, eight, four and 14 are sampled respectively in the SILC. Tables 1 and 2 report the SILC sample sizes n_{dt} for the subdomains (*comarcas* × gender) that are sampled within each province. The 'Total' columns report the sample sizes n_d for the domains (provinces × gender). For the subdomains with small sample sizes, direct estimators, obtained by using only the data from the target subdomain, are highly inefficient. In fact, they cannot be

Table 1. SILC sample sizes for the *comarcas* in Alicante and Castellón by gender

	Results for Alicante province								*Total*	*Results for Castellón province*				*Total*
Comarcas	27	28	29	30	31	32	33	34		3	5	6	7	
Men	42	42	44	77	20	144	74	34	477	28	53	34	7	122
Women	50	40	41	70	17	154	80	32	484	28	64	54	10	156
Total	92	82	85	147	37	298	154	66	961	56	117	88	17	278

Table 2. SILC sample sizes for the *comarcas* in Valencia province by gender

	Results for Valencia province														*Total*
Comarcas	11	12	13	14	15	16	17	18	20	21	22	23	24	25	
Men	49	7	60	100	214	46	6	21	62	25	9	18	36	40	693
Women	60	11	65	113	227	63	6	16	62	22	9	27	30	35	746
Total	109	18	125	213	441	109	12	37	124	47	18	45	66	75	1439

calculated for the unsampled *comarcas*. For this reason, we need to resort to models that link the data from all the subdomains to obtain estimators of acceptable efficiency. Section 3 proposes a suitable small area estimation model for this problem with a nested structure of domains.

3. Twofold nested error model

In populations with the given nested structure of domains or where two-stage sampling is carried out, it makes sense to consider a model representing both types of variation of the variable of interest (or a transformation of it), namely the variation across domains and the variation across subdomains within each domain. In fact, since income is heavily right skewed, it is common practice to take a transformation of it as response variable in a regression model. Let $Y_{dtj} = T(E_{dtj})$ be the transformed income for individual j, where $T(\cdot)$ is one to one. We assume that Y_{dtj} follows the twofold nested error model given by

$$Y_{dtj} = \mathbf{x}'_{dtj}\boldsymbol{\beta} + u_{1,d} + u_{2,dt} + w_{dtj}^{-1/2} e_{dtj}, \qquad j=1,\ldots,N_{dt}, \quad t=1,\ldots,M_d, \quad d=1,\ldots,D, \tag{3}$$

where $\mathbf{x}_{dtj}$ is a $p \times 1$ vector with the values of auxiliary variables for the individuals (the factors that affect income), $\boldsymbol{\beta}$ is the $p \times 1$ vector of regression coefficients (which is common for all domains and subdomains), $u_{1,d}$ is the effect of domain d, $u_{2,dt}$ is the effect for subdomain t within domain d, e_{dtj} is the individual model error and w_{dtj} is a known heteroscedasticity weight. Domain and subdomain effects, and individual model errors represent respectively the unexplained variation of the transformed incomes across domains, subdomains and individuals. They are all assumed to be mutually independent, satisfying

$$\left.\begin{aligned} u_{1,d} &\overset{\text{IID}}{\sim} N(0,\sigma_1^2), \\ u_{2,dt} &\overset{\text{IID}}{\sim} N(0,\sigma_2^2), \\ e_{dtj} &\overset{\text{IID}}{\sim} N(0,\sigma_0^2). \end{aligned}\right\} \tag{4}$$

Heteroscedasticity weights w_{dtj} can be obtained by, first, fitting a model with homoscedasticity ($w_{dtj}=1$, $\forall\ d,t,j$) and finding whether there is any departure from this assumption by looking at residual plots; in particular, plotting model residuals against predicted values or against covariates that may be related to the error variances. If these plots indicate changes in the variability of residuals that are related to some covariate, then the functional relationship between the variance and the potential covariates can be estimated by using the model residuals obtained. The resulting fitted variances are used as constants w_{dtj}^{-1}.

We denote by $\boldsymbol{\theta}=(\boldsymbol{\beta}',\sigma_0^2,\sigma_1^2,\sigma_2^2)'$ the vector of unknown model parameters. Setting $\sigma_2^2=0$ but letting $\sigma_1^2>0$ in model (3), we obtain the onefold nested error model with only domain effects $u_{1,d}$. Setting instead $\sigma_1^2=0$ but keeping $\sigma_2^2>0$, we obtain a onefold nested error model with subdomain effects $u_{2,dt}$ alone.

We wish to estimate domain and subdomain parameters that might be non-linear but are additive in the individual values Y_{dtj}. Specifically, for a subdomain t within a given domain d, we consider general parameters δ_{dt} satisfying

$$\delta_{dt}=\frac{1}{N_{dt}}\sum_{j=1}^{N_{dt}}h(Y_{dtj}), \tag{5}$$

where $h(\cdot)$ is a known real-valued function. For a domain d, we consider additive parameters δ_d, defined analogously as

$$\delta_d=\frac{1}{N_d}\sum_{t=1}^{M_d}\sum_{j=1}^{N_{dt}}h(Y_{dtj})=\frac{1}{N_d}\sum_{t=1}^{M_d}N_{dt}\delta_{dt}. \tag{6}$$

Average income (1) and FGT poverty indicators (2) are special cases of additive parameters in the above sense.

4. Empirical best predictors

We wish to estimate the target domain and subdomain parameters (5) and (6) by using data coming from a sample s drawn from the population U. Let U_d and U_{dt} be the population from the dth domain and tth subdomain within domain d, and let $s_d\subset U_d$ and $s_{dt}\subset U_{dt}$ be the domain and the subdomain subsamples of sizes $n_d\leqslant N_d$ and $n_{dt}\leqslant N_{dt}$ respectively, $d=1,\ldots,D$, $t=1,\ldots,M_d$. Some domains or subdomains might have zero sample sizes, i.e. $n_d=0$ or $n_{dt}=0$. We assume without loss of generality that, for each sampled domain U_d, the first m_d subdomains are those represented in the sample, whereas the last M_d-m_d subdomains are the non-sampled subdomains. We also assume that there is no sample selection bias, which means that the sample observations $\mathbf{y}_s=\{Y_{dtj},\ j\in s_{dt},\ t=1,\ldots,m_d,\ d=1,\ldots,D\}$ obey exactly the same population model (3)–(4).

According to Molina and Rao (2010), for a sampled subdomain t, the EB estimator of δ_{dt} is obtained by preserving the $h(Y_{dtj})$ values corresponding to the sample units and predicting those corresponding to the non-sampled units, i.e.

$$\hat{\delta}_{dt}^{\text{EB}}=\frac{1}{N_{dt}}\Big\{\sum_{j\in s_{dt}}h(Y_{dtj})+\sum_{j\in U_{dt}-s_{dt}}E[h(Y_{dtj})|\mathbf{y}_s;\hat{\boldsymbol{\theta}}]\Big\},\qquad t=1,\ldots,m_d, \tag{7}$$

where the expected value in the second term is with respect to the distribution of $Y_{dtj}|\mathbf{y}_s$, with the unknown $\boldsymbol{\theta}$ replaced by a consistent estimator $\hat{\boldsymbol{\theta}}=(\hat{\boldsymbol{\beta}}',\hat{\sigma}_0^2,\hat{\sigma}_1^2,\hat{\sigma}_2^2)'$. For more details on EB estimation, see the on-line supplementary material. For a subdomain t that is not represented in the sample, its corresponding subsample is $s_{dt}=\emptyset$ and the EB estimator of δ_{dt} is obtained by predicting the $h(Y_{dtj})$ values for all subdomain units, i.e.

$$\hat{\delta}_{dt}^{\text{EB}} = \frac{1}{N_{dt}} \sum_{j=1}^{N_{dt}} E[h(Y_{dtj})|\mathbf{y}_s; \hat{\boldsymbol{\theta}}], \qquad t = m_d + 1, \ldots, M_d. \tag{8}$$

To estimate a parameter δ_d in a given domain d, independently of whether it is sampled or not, the EB estimator can be obtained by adding the EB estimators of the totals for the subdomains that it contains, i.e.

$$\hat{\delta}_{d}^{\text{EB}} = \frac{1}{N_{d}} \sum_{t=1}^{M_{d}} N_{dt} \hat{\delta}_{dt}^{\text{EB}}, \qquad d = 1, \ldots, D. \tag{9}$$

According to expression (9), the EB estimators of the totals for subdomains $N_{dt}\hat{\delta}_{dt}^{\text{EB}}$ automatically aggregate to the EB estimator of the corresponding domain total $N_d\hat{\delta}_d^{\text{EB}}$. In fact, since our target parameters are additive, the EB estimator for a subdomain or a domain under a onefold model with only domain or subdomain effects (i.e. with $\sigma_1^2 = 0$ or $\sigma_2^2 = 0$) can be obtained by adding the sample values with predicted values similarly to expression (7). Thus, for additive parameters, the estimates of the subdomain totals that are obtained under onefold models also add up to the corresponding estimated domain total.

National statistical institutes are interested in estimates that are benchmarked at the national level, i.e. which satisfy the additional restriction $N\delta = \Sigma_{d=1}^{D} N_d \delta_d$, where

$$\delta = N^{-1} \sum_{d=1}^{D} \sum_{t=1}^{M_d} \sum_{j=1}^{N_{dt}} h(Y_{dtj}).$$

If the national estimator $\hat{\delta}^{\text{EB}}$ is also an EB predictor under the same model (3)–(4), then the benchmarking property $N\hat{\delta}^{\text{EB}} = \Sigma_{d=1}^{D} N_d \hat{\delta}_d^{\text{EB}}$ is also automatically fulfilled. Otherwise, benchmarking adjustments (e.g. ratios or differences) can be applied to both domain and subdomain estimators.

When the covariates are continuous, the EB estimator $\hat{\delta}_{dt}^{\text{EB}}$ of the subdomain parameter δ_{dt} given in expression (7) requires that we have a census or a register with the values of the auxiliary variables for all the population units. It also requires us to identify the sample units in the census or register of the auxiliary variables. This information is not always available. However, since typically the sample is small compared with the population, instead of preserving the values for the sample units, we can predict the values for all units similarly to expression (8) by treating the sample values as if they were not in $\mathbf{y}_s$. This estimator is called census EB and its performance does not differ significantly from the original EB estimator as long as the sample fractions in the subdomains are negligible; see Guadarrama *et al.* (2014).

When the covariates considered are categorical or take a finite number of values, as in our application to poverty mapping, neither the values of the covariates for each individual nor the identification of the sample units in the census or register are required. Suppose that the vector of covariates $\mathbf{x}_{dtj}$ can take only values in the finite set $\{\mathbf{z}_1, \ldots, \mathbf{z}_K\}$. In this case, the EB estimator (7) reduces to

$$\hat{\delta}_{dt}^{\text{EB}} = \frac{1}{N_{dt}} \left\{ \sum_{j \in s_{dt}} h(Y_{dtj}) + \sum_{k=1}^{K} (N_{dtk} - n_{dtk}) \hat{E}_{dt}^{(k)} \right\}, \qquad t = 1, \ldots, m_d, \tag{10}$$

where $\hat{E}_{dt}^{(k)} = E[h(Y_{dtj})|\mathbf{y}_s; \hat{\boldsymbol{\theta}}]$ for $\mathbf{x}_{dtj} = \mathbf{z}_k$, $N_{dtk} = \#\{j \in U_{dt} : \mathbf{x}_{dtj} = \mathbf{z}_k\}$ is the number of population units in the subdomain with covariate vector equal to $\mathbf{z}_k$ and $n_{dtk} = \#\{j \in s_{dt} : \mathbf{x}_{dtj} = \mathbf{z}_k\}$ is the corresponding number of sample units in the subdomain. In this case, in addition to the survey data, we require only the counts N_{dtk} of people in each covariate class within each subdomain, which might be obtained from external sources of data providing aggregated auxiliary information.

Since the elements Y_{dtj} are normally distributed, the conditionals are also normally distributed. Concretely, we have $Y_{dtj}|\mathbf{y}_s \sim N\{\mu_{dtj|s}(\boldsymbol{\theta}), \sigma^2_{dtj|s}(\boldsymbol{\theta})\}$ where, for a sampled subdomain t within sampled domain d, the conditional mean is

$$\mu_{dtj|s}(\boldsymbol{\theta}) = \mathbf{x}'_{dtj}\boldsymbol{\beta} + \gamma_{dt}\left\{ \bar{y}_{dt} - \bar{\mathbf{x}}'_{dt}\boldsymbol{\beta} + \left(\frac{\sigma_0^2}{\sigma_2^2}\right)^2 \frac{\varphi_d}{w_{dt\cdot}} \sum_{l=1}^{m_d} \gamma_{dl}(\bar{y}_{dl} - \bar{\mathbf{x}}'_{dl}\boldsymbol{\beta}) \right\}, \qquad t = 1, \ldots, m_d. \quad (11)$$

In this expression, $w_{dt\cdot} = \Sigma_{j\in s_{dt}} w_{dtj}$ is the total of heteroscedasticity weights in subdomain t, $\bar{y}_{dt} = w_{dt\cdot}^{-1} \Sigma_{j\in s_{dt}} w_{dtj} Y_{dtj}$ and $\bar{\mathbf{x}}_{dt} = w_{dt\cdot}^{-1} \Sigma_{j\in s_{dt}} w_{dtj} \mathbf{x}_{dtj}$ are the weighted sample means of the response variable and the vector of auxiliary variables, $\gamma_{dt} = \sigma_2^2/(\sigma_2^2 + \sigma_0^2/w_{dt\cdot})$ and $\varphi_d = \sigma_1^2\{\sigma_0^2 + \sigma_1^2 \Sigma_{t=1}^{m_d}(1-\gamma_{dt}) w_{dt\cdot}\}^{-1}$. For details on the derivation, see the on-line supplementary material.

Observe that the conditional mean (11) is the sum of the regression part $\mathbf{x}'_{dtj}\boldsymbol{\beta}$, the estimated effect of the corresponding subdomain $\gamma_{dt}(\bar{y}_{dt} - \bar{\mathbf{x}}'_{dt}\boldsymbol{\beta})$ and the estimated effect of the corresponding domain, which is composed of a weighted sum of estimated effects of all the sampled subdomains belonging to that domain. Note that the regression errors $\bar{y}_{dt} - \bar{\mathbf{x}}'_{dt}\boldsymbol{\beta}$ defining subdomain effects correct the regression term $\mathbf{x}'_{dtj}\boldsymbol{\beta}$ in the case that covariates are not strongly powerful. This is in fact the advantage of EB estimators compared with the ELL method, which is basically synthetic and therefore does not correct the regression.

For a non-sampled subdomain t contained in a sampled domain d, the conditional mean takes advantage of the domain effect, which can be estimated with the observations in the domain, i.e.

$$\mu_{dtj|s}(\boldsymbol{\theta}) = \mathbf{x}'_{dtj}\boldsymbol{\beta} + \frac{\sigma_0^2}{\sigma_2^2}\varphi_d \sum_{l=1}^{m_d} \gamma_{dl}(\bar{y}_{dl} - \bar{\mathbf{x}}'_{dl}\boldsymbol{\beta}), \qquad t = m_d + 1, \ldots, M_d. \quad (12)$$

Now, if domain d is not sampled at all, it is not possible to account for model error and the conditional mean reduces to the same as in the ELL method: $\mu_{dtj|s}(\boldsymbol{\theta}) = \mathbf{x}'_{dtj}\boldsymbol{\beta}$. For the conditional variance, in the supplementary material we obtain

$$\sigma^2_{dtj|s}(\boldsymbol{\theta}) = \begin{cases} \sigma_0^2[w_{dtj}^{-1} + \varphi_d\{1 + \gamma_{dt}(\gamma_{dt} - 2)\}] + \sigma_2^2(1-\gamma_{dt}), & t = 1, \ldots, m_d, \\ \sigma_0^2(w_{dtj}^{-1} + \varphi_d) + \sigma_2^2, & t = m_d + 1, \ldots, M_d, \\ w_{dtj}^{-1}\sigma_0^2 + \sigma_1^2 + \sigma_2^2, & n_d = 0. \end{cases} \quad (13)$$

The conditional means (11) and (12) and the conditional variance (13) depend on the unknown $\boldsymbol{\theta}$. A consistent estimator $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\beta}}', \hat{\sigma}_0^2, \hat{\sigma}_1^2, \hat{\sigma}_2^2)'$ of $\boldsymbol{\theta}$ under the twofold model (3) can be obtained by usual fitting methods for linear mixed models, such as maximum likelihood, restricted maximum likelihood or the method of moments. The onefold nested error models that are obtained when $\sigma_1^2 = 0$ or $\sigma_2^2 = 0$ can be fitted similarly.

FGT poverty indicators (2) can be expressed in terms of model responses Y_{dtj} as

$$F_{\alpha,dtj} = \left\{ \frac{z - T^{-1}(Y_{dtj})}{z} \right\}^{\alpha} I\{T^{-1}(Y_{dtj}) < z\} \triangleq h_\alpha(Y_{dtj}).$$

As we have already noted, these parameters are special cases of additive parameters (5). Then, applying expression (7) to $\delta_{dt} = F_{\alpha,dt}$, we obtain the EB estimator

$$\hat{F}^{\mathrm{EB}}_{\alpha,dt} = \frac{1}{N_{dt}} \left(\sum_{j\in s_{dt}} F_{\alpha,dtj} + \sum_{j\in U_{dt} - s_{dt}} \hat{F}^{\mathrm{EB}}_{\alpha,dtj} \right),$$

where $\hat{F}^{\mathrm{EB}}_{\alpha,dtj} =^{\triangle} E[h_\alpha(Y_{dtj})|\mathbf{y}_s; \hat{\boldsymbol{\theta}}]$. For a non-sampled subdomain ($t = m_d + 1, \ldots, M_d$), we have $s_{dt} = \emptyset$ in this estimator. For the special cases of poverty incidence and poverty gap ($\alpha = 0, 1$) and for certain transformations $T(\cdot)$, the expectation defining $\hat{F}^{\mathrm{EB}}_{\alpha,dtj}$ can be calculated analytically. Concretely, for the poverty incidence ($\alpha=0$), if $T(\cdot)$ is monotone non-decreasing, we obtain

$$\hat{F}^{\mathrm{EB}}_{0,dtj}=\Phi(\hat{\alpha}_{dtj}), \qquad \hat{\alpha}_{dtj}=\frac{T(z)-\hat{\mu}_{dtj|s}}{\hat{\sigma}_{dtj|s}}, \tag{14}$$

where $\hat{\mu}_{dtj|s}=\mu_{dtj|s}(\hat{\boldsymbol{\theta}})$, $\hat{\sigma}^2_{dtj|s}=\sigma^2_{dtj|s}(\hat{\boldsymbol{\theta}})$ and $\Phi(\cdot)$ denotes the cumulative distribution function of a standard normal random variable. To derive a closed form expression for the poverty gap ($\alpha=1$), we need the form of the specific transformation of income $T(E_{dij})$. A typical transformation that stabilizes the variance and achieves approximate normality is $T(E_{dtj})=\log(E_{dtj}+c)$, for a constant $c>0$. For this transformation, we obtain

$$\hat{F}^{\mathrm{EB}}_{1,dtj}=\Phi(\hat{\alpha}_{dtj})-\frac{1}{z}\left\{\exp\left(\hat{\mu}_{dtj|s}+\frac{\hat{\sigma}^2_{dtj|s}}{2}\right)\Phi(\hat{\alpha}_{dtj}-\hat{\sigma}_{dtj|s})-c\Phi(\hat{\alpha}_{dtj})\right\}. \tag{15}$$

For more complex parameters δ_{dt} defined as in equation (5) with non-linear $h(\cdot)$, the expected value $E[h(Y_{dtj})|\mathbf{y}_s;\hat{\boldsymbol{\theta}}]$ that is involved in the EB estimator (7) might not be analytically tractable. In such cases, similarly to Molina and Rao (2010), a Monte Carlo (MC) procedure based on repeated generation of random values from the distribution of $Y_{dtj}|\mathbf{y}_s$ can be applied to approximate the EB estimator. Note that, in the case of additive parameters satisfying equation (5), the MC approximation of the above expectation requires generation of only univariate normal random variables, which makes this MC method computationally feasible for large populations. Under the twofold model (3)–(4), the MC procedure for approximation of the EB estimator of additive subdomain and domain parameters δ_{dt} and δ_d is as follows.

(a) Fit the twofold model (3)–(4) to the sample data $\mathbf{y}_s$, obtaining an estimator $\hat{\boldsymbol{\theta}}=(\hat{\boldsymbol{\beta}}',\hat{\sigma}^2_0,\hat{\sigma}^2_1,\hat{\sigma}^2_2)'$ of $\boldsymbol{\theta}=(\boldsymbol{\beta}',\sigma^2_0,\sigma^2_1,\sigma^2_2)'$.

(b) Calculate $\hat{\mu}_{dtj|s}$ and $\hat{\sigma}^2_{dtj|s}$ by replacing the estimator $\hat{\boldsymbol{\theta}}=(\hat{\boldsymbol{\beta}}',\hat{\sigma}^2_0,\hat{\sigma}^2_1,\hat{\sigma}^2_2)'$ that is obtained in (a) for $\boldsymbol{\theta}=(\boldsymbol{\beta}',\sigma^2_0,\sigma^2_1,\sigma^2_2)'$ in expressions (11)–(13). Then generate L independent copies of Y_{dtj}, for each non-sample unit j, as

$$Y^{(l)}_{dtj}\sim N(\hat{\mu}_{dtj|s},\hat{\sigma}^2_{dtj|s}), \qquad j\in U_{dt}-s_{dt}, \quad t=1,\ldots,M_d, \quad d=1,\ldots,D, \quad l=1,\ldots,L.$$

(c) For each non-sample unit $j\in U_{dt}-s_{dt}$ from subdomain t from domain d, an MC approximation of the expected value required is then

$$E[h(Y_{dtj})|\mathbf{y}_s;\hat{\boldsymbol{\theta}}]\approx\frac{1}{L}\sum_{l=1}^{L}h(Y^{(l)}_{dtj}). \tag{16}$$

An MC approximation to the EB estimator of the subdomain parameter δ_{dt} is then obtained by plugging approximation (16) in expression (7).

As to the MSE of the proposed EB estimators, for complex parameters such as the FGT poverty indicators, analytical approximations with good properties are difficult to derive. However, resampling techniques can be readily applied to obtain MSE estimators. In the on-line supplementary material we describe an extension of the parametric bootstrap method for finite populations that was introduced by González-Manteiga *et al.* (2007, 2008) to the twofold nested error model (3)–(4).

5. Mean-squared error under misspecified random effects

The ELL method that is used traditionally by the World Bank assumes a onefold nested error model including only cluster effects, where clusters are typically nested in the domains of interest (subdomains). If the model assumed contains only subdomain random effects but auxiliary variables do not fully explain the between-domain variation of the response variable, there is a model misspecification. In this case, the corresponding MSE estimators that are derived under the model assumed can be misleading. Here we study the effect on MSE estimation of this model misspecification, when we estimate the domain means of the response variables by using the ELL and EB methods. A similar study can be done for the alternative misspecified model with only domain effects, but we omit it for brevity and because comparing with the traditional ELL approach might have greater interest.

Thus, in this section, the target parameters are the domain means

$$\bar{Y}_d = \frac{1}{N_d} \sum_{t=1}^{M_d} \sum_{j=1}^{N_{dt}} Y_{dtj}, \qquad d=1,\ldots,D.$$

We assume that the true (but unknown) model is the twofold nested error model (3)–(4) that includes domain as well as subdomain random effects. However, we put ourselves in the role of a practitioner who ignores the domain effects and considers as correct model the onefold model containing subdomain effects only:

$$Y_{dtj} = \mathbf{x}'_{dtj}\boldsymbol{\beta} + u_{2,dt} + w_{dtj}^{-1/2} e_{dtj}, \qquad j=1,\ldots,N_{dt}, \quad t=1,\ldots,M_d, \quad d=1,\ldots,D, \tag{17}$$

similarly to traditional ELL applications. We compare the MSE that is obtained under the assumed onefold model with the MSE that is derived under the true twofold model.

For simplicity of exposition, we take $w_{dtj}=1$ for all j, t and d, and balanced population and sample sizes across all subdomains within the same domain, i.e. $N_{dt}=N_{d0}$ and $n_{dt}=n_{d0}$, $t=1,\ldots,M_d$, $d=1,\ldots,D$. In this case, $\gamma_{dt}=\sigma_2^2/(\sigma_2^2+\sigma_0^2/n_{d0}) \stackrel{\triangle}{=} \gamma_{d0}$ is also constant across subdomains $t=1,\ldots,M_d$. Moreover, we consider that the model parameters $\boldsymbol{\theta}=(\boldsymbol{\beta}',\sigma_0^2,\sigma_1^2,\sigma_2^2)'$ are known.

In the ELL method, the estimator of the target parameter and its MSE are obtained by using a bootstrap procedure under the model assumed. In particular, the ELL estimator of $\bar{Y}_d$ is a bootstrap approximation to the marginal expectation of $\bar{Y}_d$. Thus, for large numbers of bootstrap replicates, the ELL estimator of $\bar{Y}_d$ is given by $\hat{\bar{Y}}_d^{\mathrm{ELL(A)}} \approx \bar{\mathbf{X}}'_d\boldsymbol{\beta}$, where 'A' stands for 'assumed model'. Note that $\hat{\bar{Y}}_d^{\mathrm{ELL(A)}}$ is synthetic, in the sense that it does not account for domain or subdomain effects. In fact, the ELL estimator under the true twofold model is the same.

As MSE estimator of $\hat{\bar{Y}}_d^{\mathrm{ELL(A)}}$, the ELL method actually delivers a bootstrap approximation to the variance of $\bar{Y}_d$ under the assumed model (17). Averaging across the domain units in model (17), we obtain $\bar{Y}_d=\bar{\mathbf{X}}'_d\boldsymbol{\beta}+\bar{U}_{d,2}+\bar{E}_d$, for $\bar{U}_{d,2}=M_d^{-1}\Sigma_{t=1}^{M_d}u_{2,dt}$ with $V_{\mathrm{A}}(\bar{U}_{d,2})=\sigma_2^2/M_d$ and $\bar{E}_d=N_d^{-1}\Sigma_{t=1}^{M_d}\Sigma_{j=1}^{N_{dt}}e_{dtj}$ with $V_{\mathrm{A}}(\bar{E}_d)$ negligible for large N_d. Hence, for large numbers of bootstrap replicates and large N_d, the ELL estimator of the MSE is given by

$$\mathrm{mse}_{\mathrm{ELL(A)}}(\hat{\bar{Y}}_d^{\mathrm{ELL(A)}}) \approx V_{\mathrm{A}}(\bar{Y}_d) \approx \sigma_2^2/M_d.$$

However, by similar calculations, the correct MSE under the true twofold model (3) for N_d large is given by

$$\mathrm{MSE}_{\mathrm{T}}(\hat{\bar{Y}}_d^{\mathrm{ELL(A)}}) = V_{\mathrm{T}}(\bar{Y}_d) \approx \sigma_1^2+\sigma_2^2/M_d.$$

The difference between the true and estimated MSE by the ELL method for N_d large is then σ_1^2, which is strictly positive as long as domain effects are significant. Hence, the ELL approach based on the model with subdomain effects underestimates the true MSE of the ELL estimator of the domain mean $\bar{Y}_d$ by only approximately σ_1^2 when the available unit or domain level regressors do not explain completely the between-domain variation of the variable of interest ($\sigma_1^2 > 0$).

In contrast, note that, if the domain sampling fraction is negligible ($n_d/N_d \approx 0$), the EB estimator of $\bar{Y}_d$ under the assumed onefold model (17) is

$$\hat{\bar{Y}}_d^{\mathrm{EB(A)}} \approx \bar{\mathbf{X}}_d' \boldsymbol{\beta} + \hat{\bar{U}}_{2,d}^{\mathrm{EB}}, \tag{18}$$

for $\hat{\bar{U}}_{2,d}^{\mathrm{EB}} = M_d^{-1} \Sigma_{t=1}^{M_d} \hat{u}_{2,dt}^{\mathrm{EB}}$, where $\hat{u}_{2,dt}^{\mathrm{EB}} = \gamma_{d0}(\bar{y}_{dt} - \bar{\mathbf{x}}_{dt}'\boldsymbol{\beta})$ is the estimated subdomain effect under the model assumed. The EB estimator (18) of the domain mean $\bar{Y}_d$ includes the average of estimated subdomain effects, in contrast with the ELL estimator $\hat{\bar{Y}}_d^{\mathrm{ELL(A)}}$, which is synthetic. For the EB estimator, the difference between the true MSE and the MSE under the misspecified model is

$$\mathrm{MSE_T}(\hat{\bar{Y}}_d^{\mathrm{EB(A)}}) - \mathrm{MSE_A}(\hat{\bar{Y}}_d^{\mathrm{EB(A)}}) \approx \gamma_{d0}^2 \sigma_1^2 / M_d + \sigma_1^2 (1 - 2\gamma_{d0}); \tag{19}$$

see the on-line supplementary material. The first term in approximation (19) is strictly positive. For small numbers of subdomains within the domain, M_d, the MSE difference might be positive, leading to underestimation of the true MSE. However, when M_d is large, the first term will be small and the second term will determine the sign of the MSE difference. The second term will be strictly positive if $\sigma_1^2 > 0$ and $\sigma_2^2 < \sigma_0^2/n_{d0}$. In this case, if domain effects are significant but we consider the model with only subdomain effects as correct, the MSE of the EB estimator is underestimated if the subdomain variability is smaller than σ_0^2/n_{d0}, but it might be overestimated otherwise.

The number of subdomains within each domain M_d is large in many applications, like when considering the sampling clusters as subdomains, but of course this depends on the choice of domains and subdomains in the particular problem. In Spain, if we take autonomous communities as domains and *comarcas* as subdomains, we can find autonomous communities with more that 100 *comarcas*. In our actual application restricted to the autonomous community of Valencia, the estimate of σ_1^2 turns out to be close to 0 and therefore the MSE difference is around 0 regardless of M_d.

6. Model-based simulation experiments

This section describes model-based simulation experiments designed mainly to compare EB estimators derived under the twofold model that contains both domain and subdomain effects with EB estimators based on a onefold model containing either only subdomain effects or only domain effects. We shall also compare with direct estimators and with ELL estimators based on the onefold model including only subdomain effects (the usual set-up of the ELL method).

In these simulation experiments, populations are generated similarly to Molina and Rao (2010), but including an indicator of subdomains within domains. Concretely, we consider a population of size $N = 20000$ partitioned into $D = 40$ domains, with domain population sizes $N_d = 500$, $d = 1, \ldots, D$. In line with the average number of *comarcas* in our application to poverty mapping in the region of Valencia, we consider that each domain d is composed of $M_d = 10$ subdomains, $d = 1, \ldots, D$. Subdomain population sizes are $N_{dt} = 50$, $t = 1, \ldots, M_d$, $d = 1, \ldots, D$. Also, similarly to the application, where only categorical variables are available, we consider as auxiliary variables two dummy indicators, which could correspond to two categories of a factor affecting income. The values of these indicators for the units are generated as $x_{kdtj} \overset{\mathrm{ind.}}{\sim} \mathrm{Bin}(1, p_{kdt})$, $k = 1, 2$, with probabilities $p_{1dt} = 0.2 + 0.4d/D + 0.4t/M_d$ and $p_{2dt} = 0.2$, $t = 1, \ldots, M_d$, $d = 1, \ldots, D$.

Domain effects, subdomain effects and individual errors are generated independently as $u_{1,d} \sim^{\text{IID}} N(0,\sigma_1^2)$, $u_{2,dt} \sim^{\text{IID}} N(0,\sigma_2^2)$ and $e_{dtj} \sim^{\text{IID}} N(0,\sigma_0^2)$. Simulations will be repeated for various combinations of random-effects variances σ_1^2 and σ_2^2, keeping the individual error variance fixed at $\sigma_0^2 = 0.25$. To imitate our real application, we consider as target variable E_{dtj} but, as dependent variable in the model, we take $Y_{dtj} = \log(E_{dtj})$. Thus, E_{dtj} can be interpreted as income. The population of Y-values is then generated from the twofold model (3)–(4) with $\mathbf{x}_{dtj} = (1, x_{1dtj}, x_{2dtj})'$, $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)' = (3, 0.03, -0.04)'$ and $w_{dtj} = 1$ for all j, t and d. As poverty line, we consider the value that is used in official statistics, taking $z = 0.6\,\text{median}(E_{dtj})$ for a population generated as above, with $E_{dtj} = \exp(Y_{dtj})$.

Concerning the sample, we consider two cases. In case I, all subdomains are sampled ($m_d = M_d = 10$). A simple random sample s_{dt} of size $n_{dt} = 10 \leqslant N_{dt} = 50$ is drawn without replacement from each subdomain U_{dt}. Hence, we have $(m_d, n_{dt}) = (10, 10)$, $t = 1, \ldots, m_d$, $d = 1, \ldots, D$. In case II, half of the subdomains are sampled ($m_d = 5 < 10 = M_d$), and a simple random sample of size $n_{dt} = 20$ is drawn from each sampled subdomain, i.e. $(m_d, n_{dt}) = (5, 20)$, $t = 1, \ldots, m_d$, $d = 1, \ldots, D$. Domain sample sizes in both cases are equal to $n_d = \Sigma_{t=1}^{m_d} n_{dt} = 100$, $d = 1, \ldots, D$.

In MC replicate i, for $i = 1, \ldots, I$, for $I = 1000$, we generate a population of Y-values as described above and obtain the true values of the domain and subdomain poverty incidences and gaps, denoted $F_{\alpha,d}^{(i)}$ and $F_{\alpha,dt}^{(i)}$, $\alpha = 0, 1$, $t = 1, \ldots, M_d$, $d = 1, \ldots, D$. Then, we draw a sample from each MC population and, with the sample observations, we fit the three models (twofold model and onefold models with only domain or subdomain effects) and compute direct estimates (domain and subdomain means of sampled $F_{\alpha,dtj}$-values), EB estimators based on these models and ELL estimates based on the model with only subdomain effects. Let $\hat{F}_{\alpha,d}^{(i)}$ and $\hat{F}_{\alpha,dt}^{(i)}$ be one of these estimators. We compute empirical biases and empirical MSEs of the estimators for domains and subdomains as follows:

$$\text{BIAS}_{\alpha,d} = \frac{1}{I} \sum_{i=1}^{I} (\hat{F}_{\alpha,d}^{(i)} - F_{\alpha,d}^{(i)}),$$

$$\text{BIAS}_{\alpha,dt} = \frac{1}{I} \sum_{i=1}^{I} (\hat{F}_{\alpha,dt}^{(i)} - F_{\alpha,dt}^{(i)}),$$

$$\text{MSE}_{\alpha,d} = \frac{1}{I} \sum_{i=1}^{I} (\hat{F}_{\alpha,\,d}^{(i)} - F_{\alpha,d}^{(i)})^2,$$

$$\text{MSE}_{\alpha,dt} = \frac{1}{I} \sum_{i=1}^{I} (\hat{F}_{\alpha,\,dt}^{(i)} - F_{\alpha,dt}^{(i)})^2.$$

After that, we average across domains or subdomains the empirical MSEs as

$$\overline{\text{MSE}}_{\alpha} = \frac{1}{D} \sum_{d=1}^{D} \text{MSE}_{\alpha,d},$$

$$\overline{\overline{\text{MSE}}}_{\alpha} = \frac{1}{M} \sum_{d=1}^{D} \sum_{t=1}^{M_d} \text{MSE}_{\alpha,dt}.$$

In our application to poverty mapping in the autonomous community of Valencia, the variation across subdomains (*comarcas*) is clearly greater than that across domains (provinces). However, if we do not restrict ourselves to a region and estimate in the whole of Spain, we expect the between-province variation to increase because of the marked socio-economic differences between northern and southern provinces. To represent the various scenarios that can be found in real cases, in these simulation experiments we repeat the experiments for various values of σ_1 and σ_2. Moreover, since what actually matters is their relative values, we first keep $\sigma_2 = 0.10$ fixed and vary $\sigma_1 \in \{0, 0.025, 0.05, 0.1, 0.15, 0.2\}$, and then we keep $\sigma_1 = 0.05$ fixed and move σ_2 instead.

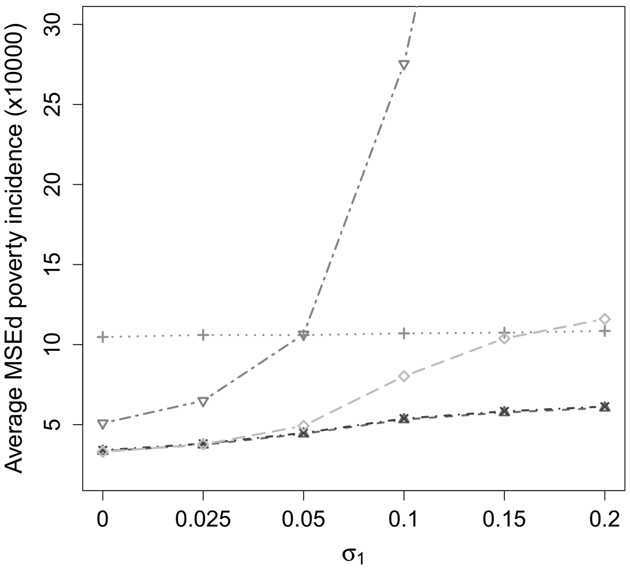
The models considered are fitted by restricted maximum likelihood using the function `lme()` of R package `nlme` (Pinheiro *et al.*, 2017). When fitting the twofold model to the data coming from a model with $\sigma_1 = 0$ or $\sigma_2 = 0$, the fitting algorithm might end up with a negative estimate for the corresponding variance, but in those cases the function truncates the estimates to 0.0001. We take as variance estimates those delivered by the output of the R function.

We first comment on the results for case I with $(m_d, n_{dt}) = (10, 10)\ \forall\, t, d$ (all subdomains are sampled). Fig. 1(a) shows the average MSE for the estimators of poverty incidences at the domain level $\overline{\mathrm{MSE}}_0$, for $\sigma_2 = 0.10$ with σ_1 varying (x-axis). Fig. 1(b) shows the average MSE for the estimators of poverty incidences at the subdomain level $\overline{\mathrm{MSE}}_0$, for $\sigma_1 = 0.05$ and σ_2 varying. When $\sigma_1 = 0$, the model with only subdomain effects is correct. Similarly, when $\sigma_2 = 0$, the model with only domain effects is correct. As Fig. 1(a) shows, direct and ELL estimators perform considerably worse than EB estimators. Direct estimators in Fig. 1(b) exceed the upper limit of the y-axis and are not shown because they would not allow a comparison with the other estimators. From the EB estimators, in both plots we can see that $\mathrm{EB}_{\mathrm{two}}$ based on the twofold model performs well for all the combinations of σ_1 and σ_2 even when the model with only subdomain effects is correct ($\sigma_1 = 0$) and when the model with only domain effects is correct ($\sigma_2 = 0$). In these two cases, the differences from the corresponding EB estimators under the correct model are quite small. This suggests that assuming the twofold model always does not practically entail loss in efficiency even if the random effects at a certain level are not needed. Note that the ELL estimator performs worse than the three types of EB estimator even when the ELL model is correct ($\sigma_1 = 0$).

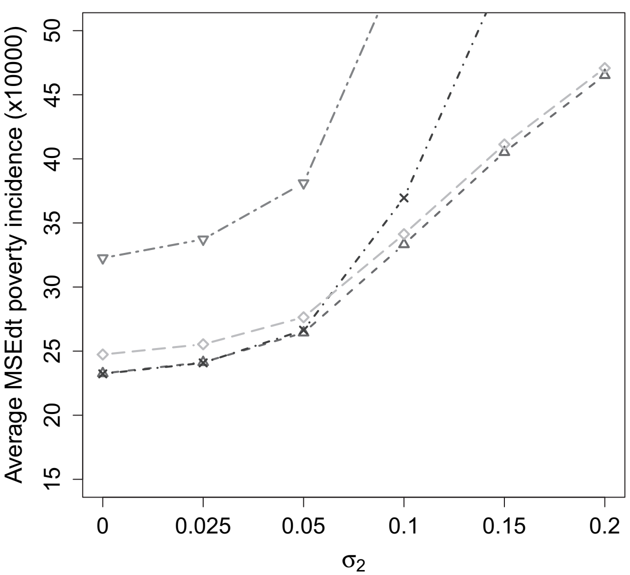
As the domain variability σ_1 decreases compared with the subdomain variability, the true model grows closer to the onefold model with subdomain effects. Consequently, as we can see in Fig. 1(a), $\mathrm{EB}_{\mathrm{sdo}}$ performs virtually the same as the EB estimator based on the twofold model. In contrast, as long as the domain variability σ_1 increases compared with the subdomain variability σ_2, the true model grows closer to the onefold model with domain effects, and $\mathrm{EB}_{\mathrm{sdo}}$ performs worse, becoming very inefficient for large σ_1. Hence, the traditional ELL approach of assuming a model with subdomain effects when estimating domain parameters is only acceptable when σ_1 is small. Note that, when estimating at the domain level, the estimator $\mathrm{EB}_{\mathrm{dom}}$ based on the model with only domain effects performs very similarly to the EB estimator under the twofold model for all the values of σ_1.

Similarly, Fig. 1(b) shows that, when estimating in subdomains, the $\mathrm{EB}_{\mathrm{dom}}$ estimator is very inefficient compared with the other EB estimators if σ_2 is large. These facts suggest that, if we need to estimate at only one of the levels and we wish to use a onefold model for simplicity or because of software availability, we should consider a model with random effects at the level where we want to estimate, provided that the target domains are sampled. This conclusion is confirmed by additional simulation results presented in Tables 1 and 2 and Figs 2 and 3 of the on-line supplementary material. Fig. 1 there also shows that EB estimators track true values better than direct estimators and seem to have smaller model bias.

In the application to poverty mapping in the region of Valencia, the resulting estimated standard deviation for the subdomain effects is around $\sigma_2 = 0.10$ and for the domain effects is in between $\sigma_1 = 0$ and $\sigma_1 = 0.025$; see Section 8. According to Fig. 1(a), for these combinations of variances, the ELL estimator performs worse than EB estimators, although the differences are much smaller than in the case of large σ_1. However, if we estimate in the whole of Spain, we expect σ_1 to be larger and the differences between ELL and EB estimators will be marked in that case.



(a)



(b)

Fig. 1. (a) Average $\mathrm{MSE}_{0,d} \times 10^4$ across domains for the estimators of domain parameters, with (a) $\sigma_2 = 0.10$ ($m_d = 10$; $n_{dt} = 10$ (domain)) and (b) average $\mathrm{MSE}_{0,dt} \times 10^4$ across subdomains for the estimators of subdomain parameters, with $\sigma_1 = 0.05$ ($m_d = 10$; $n_{dt} = 10$ (subdomain)): $+$, direct; $\times$, $\mathrm{EB}_{\mathrm{dom}}$; $\diamond$, $\mathrm{EB}_{\mathrm{sdo}}$; $\triangle$, $\mathrm{EB}_{\mathrm{two}}$; ∇, ELL

Results for case II, $(m_d, n_{dt}) = (5, 20)$, $\forall t, d$, where only half of the subdomains from each domain are sampled but the total sample size is the same as in case I, are shown in the on-line supplementary material. These simulations illustrate that, even though when estimating in subdomains $\mathrm{EB}_{\mathrm{sdo}}$ is supposed to perform well according to previous results, in the case of non-sampled subdomains this estimator does not perform well because of the zero subdomain sample size, whereas $\mathrm{EB}_{\mathrm{dom}}$, based on the onefold model including only domain effects, performs well because it is taking advantage of the domain effect to increase the efficiency (there are observations from other subdomains in the same domain).

We also analysed the precision of the MC approximation (16) in the EB estimator (7) and obtained that, for estimation of poverty incidences and gaps, the MC approximation with $L \geqslant 50$ performs nearly the same as the exact analytical expressions (14) and (15) when comparing the point estimates and also their empirical MSEs. Results also indicate good performance of the bootstrap MSE estimator.

7. Design-based simulation experiments

In this section we describe a design-based simulation study, where the 2012 SILC data play the role of a fixed population and samples are drawn from it. Since the actual data do not necessarily follow the model, this type of simulation experiment allows us to account for model error apart from sampling error. From this fixed population, we draw $I = 1000$ simple random samples of sampling fraction $f_{dt} = \frac{1}{3}$, independently within each subdomain t. In the original SILC sample there are several subdomains with zero sample size. For the sampled subdomains, the minimum, quartiles and maximum sample sizes are respectively 2, 7, 14, 21 and 76.

The target parameters are the poverty incidences for domains and subdomains. The true poverty incidences in this simulation are indeed direct estimators based on the original sample. These 'true' values turn out to be quite small (close to 0) for few subdomains. These extremely small values are unlikely to be correct in the real population and we believe that they are a consequence of the instability of direct estimators for small sample sizes. Since the model cannot reproduce unrealistic true values, in this experiment we estimate only in the subdomains for which the true poverty incidence is at least 0.1 (no extremely large poverty incidences appeared).

Fig. 2(a) shows boxplots of the empirical design biases across domains for direct, EB and ELL estimators of domain poverty incidences. EB estimators are computed by using the exact analytical formula given in expression (14) and the ELL method using a similar formula. Fig. 2(b) shows boxplots of the corresponding design MSEs. As expected, Fig. 2(a) shows that direct estimators are nearly design unbiased, in contrast with model-based estimators, which are all biased under the design. However, as we can see in Fig. 2(b), the MSE of direct estimators can take very large values compared with the MSEs of EB estimators. This is a typical situation encountered in small area estimation, where large MSE reductions are achieved at the price of increasing the design bias. However the design bias of EB estimators is kept roughly within 10% for all the domains; see Fig. 6 of the on-line supplementary material, displaying relative bias and relative root MSE (RRMSE) of these estimators. The ELL estimators perform the worst in terms of design bias and MSE.

Analogous results for the estimation in subdomains are shown in Figs 7 and 8 of the on-line supplementary material. Conclusions are similar to the previous conclusions although biases and MSEs increase because of the substantially smaller subdomain sample sizes. Note that the actual sample sizes in our real application are three times larger than those considered in this simulation experiment, so actual biases and MSEs are expected to be much smaller when estimating with the actual SILC data in subdomains. The performance of the parametric bootstrap MSE estimator, originally designed for estimation of the MSE under the model, is also analysed in the on-line supplementary material as an estimation of the design MSE. Results show that the bootstrap MSE estimates follow the tendency of design MSEs except for

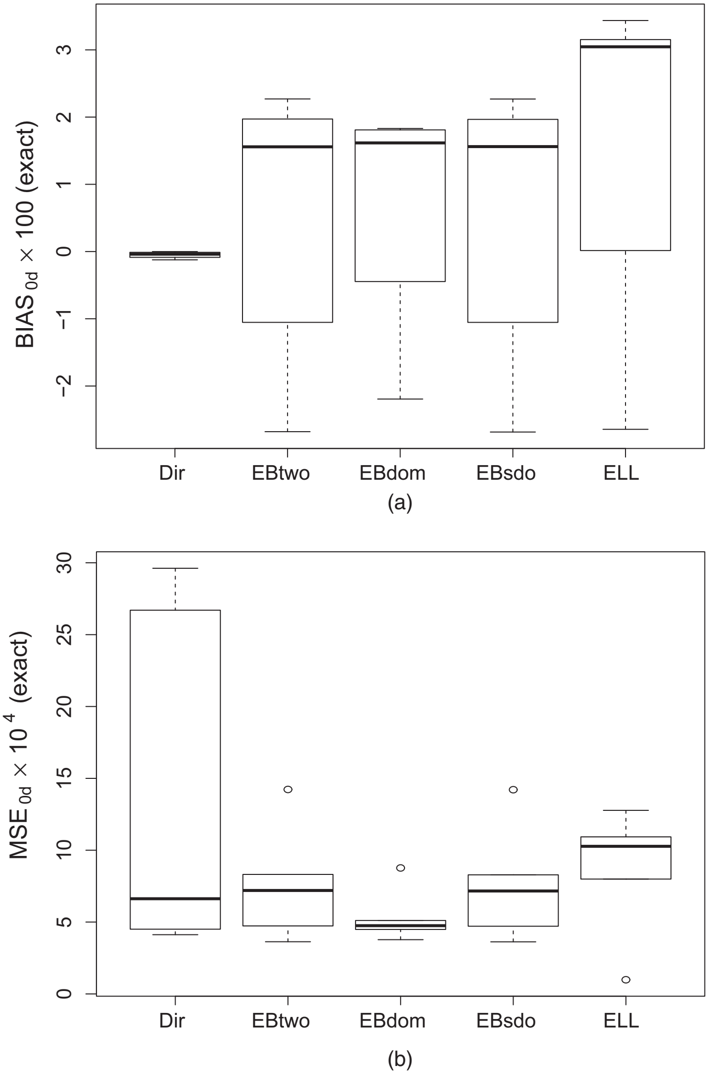


Fig. 2. Boxplots of (a) empirical biases $\text{BIAS}_{0,d} \times 10^2$ and (b) MSEs $\text{MSE}_{0,d} \times 10^4$ of EB estimators of domain poverty incidences in design-based simulation

the subdomains with extremely small sample sizes ($n_{dt} < 5$). In fact, in these subdomains, design MSEs take unrealistically small values whereas the bootstrap MSE estimates increase as we would expect, so they seem to be more realistic uncertainty measures.

8. Poverty mapping in Valencia region: results

This section describes the results for the estimation of average income, poverty incidence and poverty gap in the *comarcas* and provinces of the autonomous community of Valencia in Spain by gender, as defined in equations (1) and (2) for $\alpha = 0, 1$, using the 2012 SILC data. We use EB estimators based on the twofold model (3)–(4) for (transformed) equivalent personal income and compare with results obtained with simpler models.

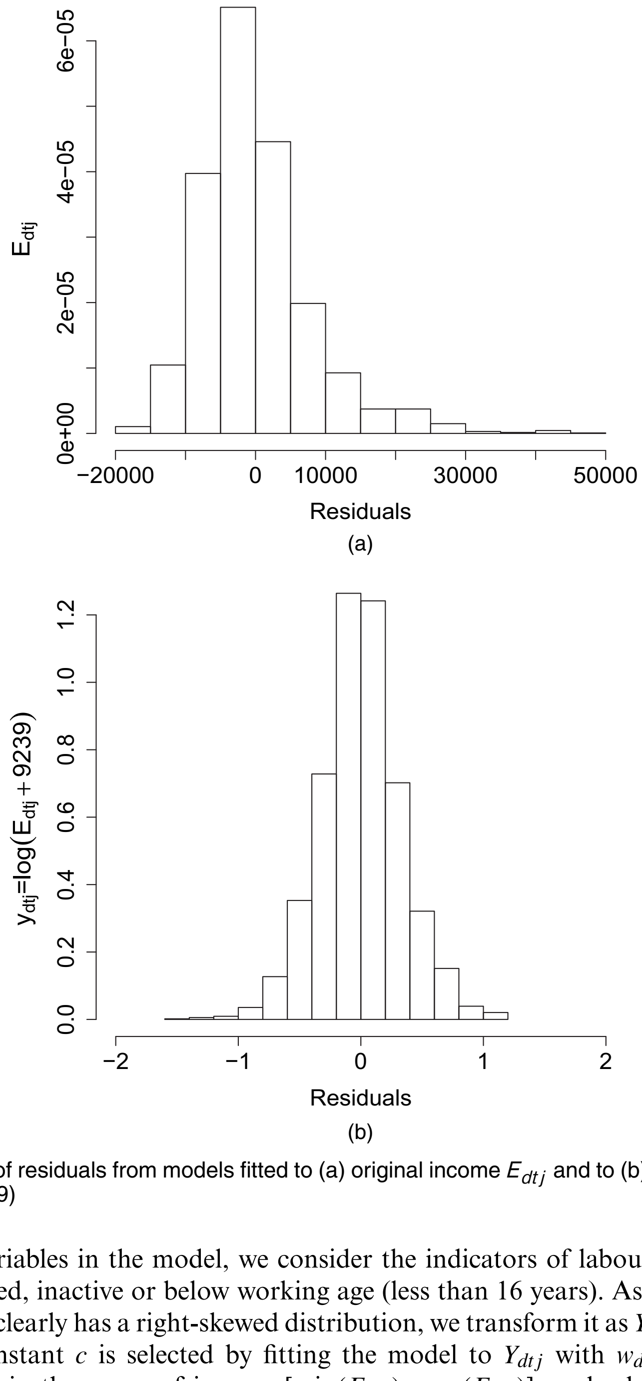


Fig. 3. Histogram of residuals from models fitted to (a) original income E_{dtj} and to (b) transformed income $Y_{dtj} = \log(E_{dtj} + 9239)$

As auxiliary variables in the model, we consider the indicators of labour force status: employed, unemployed, inactive or below working age (less than 16 years). As response variable, since income E_{dtj} clearly has a right-skewed distribution, we transform it as $Y_{dtj} = \log(E_{dtj} + c)$, for $c > 0$. The constant c is selected by fitting the model to Y_{dtj} with $w_{dtj} = 1$, $\forall j, t, d$, for a grid of c-values in the range of income $[\min(E_{dtj}), \max(E_{dtj})]$, and selecting the value of c that minimizes the Fisher skewness of model residuals. The resulting value is $c = 9239$. Fig. 3 clearly shows the skewness of the distribution of model residuals when fitting the model to the original income data (Fig. 3(a)) and the approximate symmetry after transformation (Fig. 3(b)).

Table 3. Model fitting results

Variable	*Coefficient*	*Estimate*	*Standard error*	*Degrees of freedom*	*t-value*	*p-value*
Age < 16	β_0	9.9184	0.0209	2623	473.867	0.0000
Employed	β_1	0.1494	0.0194	2623	7.715	0.0000
Unemployed	β_2	−0.1289	0.0227	2623	−5.674	0.0000
Inactive	β_3	0.0086	0.0193	2623	0.443	0.6576

As mentioned in Section 4, even if we are estimating non-linear parameters, when only categorical covariates are considered, calculation of EB estimators as given in expression (7) does not actually require availability of the values of the auxiliary variables for each population unit. In this application, the covariate vector $\mathbf{x}_{dtj}$ for individual j from subdomain t within domain d can be equal to $\mathbf{z}_1 = (1, 1, 0, 0)$, $\mathbf{z}_2 = (1, 0, 1, 0)$, $\mathbf{z}_3 = (1, 0, 0, 1)$ or $\mathbf{z}_4 = (1, 0, 0, 0)$ depending on whether the individual is employed, unemployed, inactive or below working age respectively. According to expression (10), we need only the counts N_{dtk} of people in each covariate class $k = 1, 2, 3, 4$ for each subdomain. These counts have been taken from the 2012 Spanish Labour Force Survey, which has a much larger sample size than the SILC, using the 2012 Valencian administrative register of labour market indicators at the subprovincial level to fill some Spanish Labour Force Survey missing data.

We start fitting model (3)–(4) with homoscedasticity ($w_{dtj} = 1, \forall d, t, j$) to the SILC data and checking whether model residuals give any indication against homoscedasticity. The fitted regression coefficients are reported in Table 3. Three of the categories of labour force status considered are strongly significant and the signs of the estimated coefficients are sensible, indicating that being employed corresponds to an increase in the average income whereas being unemployed is reflected in a decrease in the average income. The estimated standard deviations of errors, domain and subdomain effects turn out to be $\hat{\sigma}_0 = 0.3459$, $\hat{\sigma}_1 = 0.000\,023$ and $\hat{\sigma}_2 = 0.0919$, indicating low between-domain variation when subdomain variation is already explained and when taking into account labour force status.

The economy of Valencia region is largely based on the tourism sector, with coastal *comarcas* receiving plenty of tourists. Several main cities and industries are also close to the seaside. Therefore, their incomes are expected to be greater than those in the interior *comarcas*, independently of the province. For this reason, it is sensible that the income variation across counties is greater than across provinces as our results suggest. Remember that simulation results indicated no efficiency loss due to considering the twofold model in this case. In fact, we also fitted the model with only subdomain effects and the resulting estimated regression coefficients and residual plots are virtually identical to those obtained under the twofold model. Additionally, domain effects may provide greater efficiency for the estimators in the non-sampled subdomains. For these reasons, we retained the twofold model.

Let us now check model assumptions. Fig. 4 shows residuals against predicted values (Fig. 4(a)) and boxplots of residuals for each labour force status category (Fig. 4(b)). No clear heteroscedasticity pattern can be discerned from Fig. 4(a) and boxplots show similar quartiles across the four categories of labour force status, indicating not much difference in variability, with median residuals around 0 in all cases. These plots show no severe departure from the homoscedasticity assumption.

Next we analyse whether there is any sample selection bias, by checking whether the model residuals are somehow correlated with the sampling weights. Fig. 5 shows residuals against

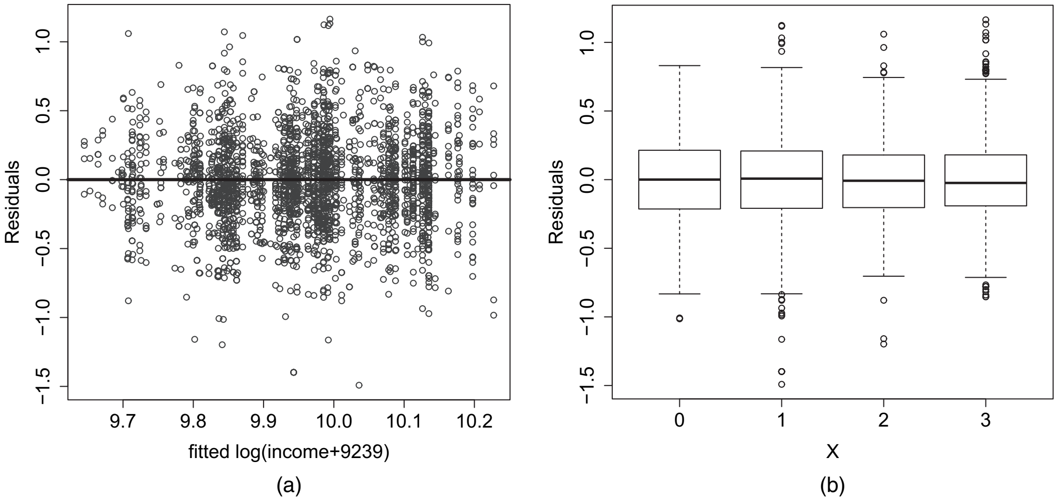


Fig. 4. (a) Residuals against predicted values and (b) boxplots of residuals for each labour force status category

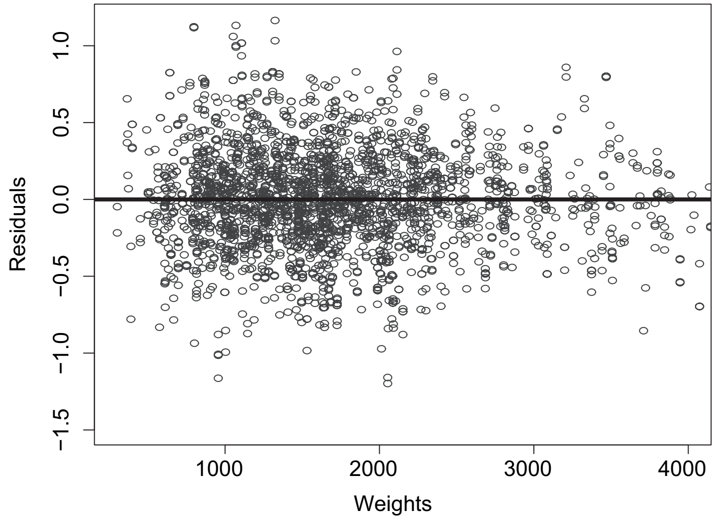


Fig. 5. Residuals against sampling weights

sampling weights. The sampling weights have a skewed distribution and, apart from the smaller number of large weights, we do not observe any clear dependence pattern in this plot.

Let us now check whether the normality assumptions hold. Fig. 6 shows normal Q–Q-plots of unit level residuals (Fig. 6(a)) and of the estimated subdomain effects (Fig. 6(b)). These figures plot theoretical $N(0,1)$ quantiles (on the x-axis) against empirical quantiles (on the y-axis). The plots indicate only mild deviations from normality at the tail ends. In real life applications, an exact fit to a distribution is barely met. In this application, we consider that the model fits reasonably well to the available data.

Since in this case domain sample sizes are not so small, it makes sense to consider also a model with fixed (instead of random) effects for the domains. However, when fitting the model

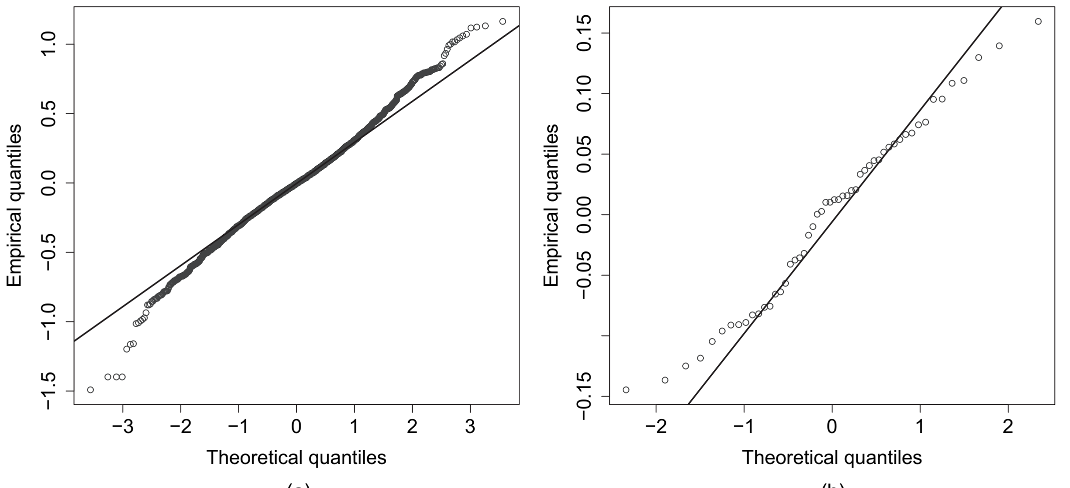


Fig. 6. Normal Q–Q-plot of (a) unit level residuals and of (b) predicted subdomain effects

Table 4. Fitting results for the model with fixed domain effects

Parameter	*Estimate*	*Standard error*	*Degrees of freedom*	*t-value*	*p-value*
β_0	9.9078	0.04165	2623	237.907	0.0000
β_1	0.1489	0.01937	2623	7.686	0.0000
β_2	−0.1291	0.02272	2623	−5.680	0.0000
β_3	0.0087	0.01936	2623	0.450	0.6526
α_1	0.0154	0.05571	46	0.277	0.7832
α_2	0.0072	0.07256	46	0.099	0.9216
α_3	−0.0416	0.07085	46	−0.587	0.5602
α_4	0.0317	0.05059	46	0.627	0.5335
α_5	0.0090	0.05043	46	0.178	0.8595

considering the domain effects as fixed, none of the domain estimated effects turn out to be significant (Table 4).

Finally, based on the twofold model, EB estimates were computed for the target domain and subdomain parameters. Table 5 reports the resulting direct and EB estimates of average income, poverty incidence and poverty gap for each province × gender. Direct estimates are actually Horvitz–Thompson estimators calculated by using the survey sampling weights. We can see that EB estimates tend to be closer to direct estimates for the domains with larger sample size, but they differ substantially for some of the domains with smaller sample sizes.

Table 6 lists the corresponding RRMSE estimates. For EB estimators, MSEs are obtained by using the parametric bootstrap procedure that is given in section 2 of the on-line supplementary material. For direct estimators, we use the usual variance estimators of the Horvitz–Thompson estimators from Särndal *et al.* (1992), pages 43, 185 and 391, considering that the survey sampling weights are the inverses of the inclusion probabilities $w_{dtj} = 1/\pi_{dtj}$, and with second-order inclusion probabilities $\pi_{dti,dtj} = \pi_{dti}\pi_{dtj}$ for all $i \neq j$ and $\pi_{dtj,dtj} = \pi_{dtj}$. According to Table 6, EB estimates have smaller RRMSE than do direct estimates for the three provinces and the two

Table 5. Direct and EB estimates of the mean income, poverty incidence and poverty gap for each province–gender

Gender	*Province*	n_d	N_d	$\hat{\bar{E}}_d^{\mathrm{dir}}$	$\hat{\bar{E}}_d^{\mathrm{EB}}$	$\hat{F}_{0,d}^{\mathrm{dir}}$	$\hat{F}_{0,d}^{\mathrm{EB}}$	$\hat{F}_{1d}^{\mathrm{dir}}$	$\hat{F}_{1d}^{\mathrm{EB}}$
Men	Castellón	122	292217	13811	14504	0.2586	0.1875	0.0836	0.0758
Men	Alicante	477	927765	13032	13711	0.2087	0.2068	0.1106	0.0835
Men	Valencia	693	1256704	13255	13758	0.2026	0.2053	0.0813	0.0829
Women	Castellón	156	293004	12735	13841	0.2884	0.2053	0.1211	0.0836
Women	Alicante	484	931713	13256	13387	0.2083	0.2156	0.1016	0.0872
Women	Valencia	746	1288874	12999	13575	0.1899	0.2087	0.0720	0.0839

Table 6. Percentage RRMSE estimates of direct and EB estimators of mean income, poverty incidence and poverty gap for each province × gender

Gender	*Province*	n_d	N_d	$\hat{\bar{E}}_d^{\mathrm{dir}}$	$\hat{\bar{E}}_d^{\mathrm{EB}}$	$\hat{F}_{0,d}^{\mathrm{dir}}$	$\hat{F}_{0,d}^{\mathrm{EB}}$	$\hat{F}_{1d}^{\mathrm{dir}}$	$\hat{F}_{1d}^{\mathrm{EB}}$
Men	Castellón	122	292217	4.21	2.48	10.54	6.22	17.75	8.55
Men	Alicante	477	927765	4.37	2.70	10.42	6.84	15.12	9.42
Men	Valencia	693	1256704	6.94	3.78	22.93	10.63	23.60	14.21
Women	Castellón	156	293004	6.99	3.27	19.59	8.73	27.06	11.78
Women	Alicante	484	931713	2.55	1.92	8.72	4.84	11.00	6.76
Women	Valencia	746	1288874	2.29	2.08	8.39	5.33	11.19	7.55

genders, and the RRMSE reductions are especially remarkable for the poverty incidence and gap in the domains with the smallest sample sizes, even if, in the case of domains, the sample sizes are not so small.

For the *comarcas*, direct and EB estimates of the poverty incidences are plotted in Fig. 7 for men (Fig. 7(a)) and women (Fig. 7(b)). In Fig. 7, the *comarcas* (on the x-axis) are sorted from smaller to larger sample size and their sample sizes are indicated in the x-axis labels. The first eight *comarcas* are those with zero sample size in the SILC and hence direct estimates are not available for them. These plots show that direct estimates are unstable across the *comarcas* compared with EB estimates, although both estimates tend to go in the same direction. Differences between the two estimates tend to be larger for the *comarcas* with smaller sample sizes (the left-hand side of the plots) but are in good agreement for the *comarcas* with sample sizes greater than 100 (right-hand side), which is true also for average income and poverty gap; see Figs 10 and 11 of the on-line supplementary material. Fig. 8 shows the estimated RRMSEs of the direct and EB estimators. It shows that EB estimators based on the twofold model have much smaller estimated RRMSEs than direct estimators for virtually all *comarcas*, and the RRMSE reductions are substantial for some of them. EB estimators also have a more stable RRMSE pattern across *comarcas*, with a small increase in the level of RRMSE for those with zero sample size as expected. Plots for average income and poverty gap present similar features (Figs 12 and 13 in the supplementary material).

Finally, Figs 9 and 10 display cartograms of EB estimates of poverty incidence and poverty gap in the *comarcas* from the autonomous community of Valencia for each gender. Alicante province is in the south of the community, the province of Valencia occupies the middle and Castellón is at the north. East of this region is the Mediterranean sea, whereas the west lies

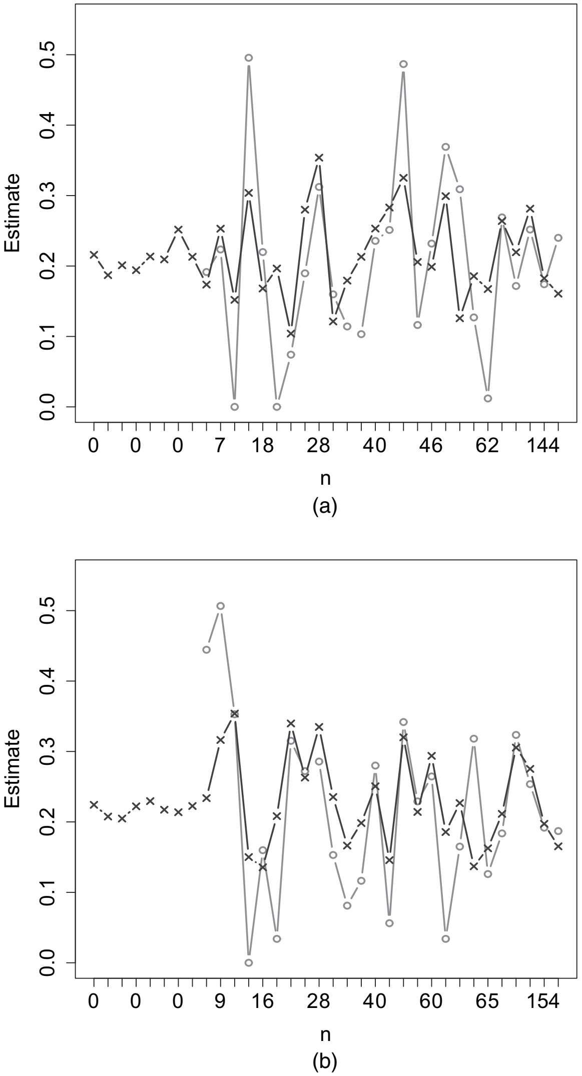


Fig. 7. Direct (○) and EB (×) estimates of poverty incidences for each *comarca* (*x*-axis) sorted from smallest to largest sample size, for (a) men and (b) women: sample sizes are shown on the *x*-axis labels

in the interior of Spain. Fig. 9 shows that, in Alicante, the *comarcas* with over 25% poverty incidence are in the interior of Alicante, together with a southern *comarca* with extreme poverty incidence (greater than 30%) for women. In Valencia province, we find one of the interior *comarcas* affected by a poverty incidence over 30%, and a small coastal *comarca* in the case of women. In Castellón, the northern *comarcas* are those suffering most from poverty, and for women there is one additional *comarca* in the interior south with extreme poverty. Poverty gaps show maps for this region very much in line with the maps for poverty incidence, indicating that the *comarcas* with greater share of people with income below the poverty line tend to be those where the depth of poverty is also greater. To summarize, our results indicate a greater variation in income across *comarcas* than across the three provinces, with a larger number of counties with extreme poverty incidences and gaps for women.

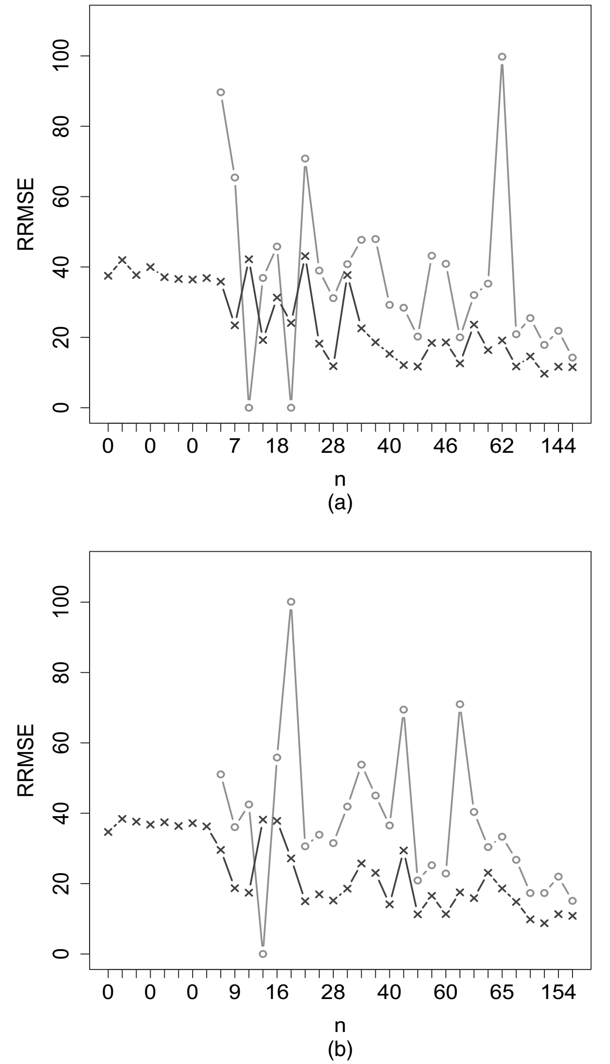


Fig. 8. Estimated RRMSEs of direct (○) and EB (×) estimators of poverty incidence for each *comarca* sorted from smallest to largest sample size, for (a) men and (b) women: sample sizes are shown on the *x*-axis labels

9. Concluding remarks

The proposed EB prediction methodology requires finding a one-to-one transformation of the target variable that follows a normal distribution. Theoretically, the transformation desired exists. If the target variable E_{dtj} is absolutely continuous with cumulative distribution function F, then we have $Y_{dtj} = T(E_{dtj}) = \Phi^{-1}\{F(E_{dtj})\} \sim N(0, 1)$, where $\Phi^{-1} : [0, 1] \mapsto R$ is the quantile function (inverse cumulative distribution function) of the standard normal distribution. In practice, finding the correct transformation might not be so easy and the transformed data might still be far from normality, leading to biased estimates. If the deviations from normality are serious, we can try to find another distribution that fits better and to derive the methodology under a twofold nested regression model with non-normal errors, for example, by considering

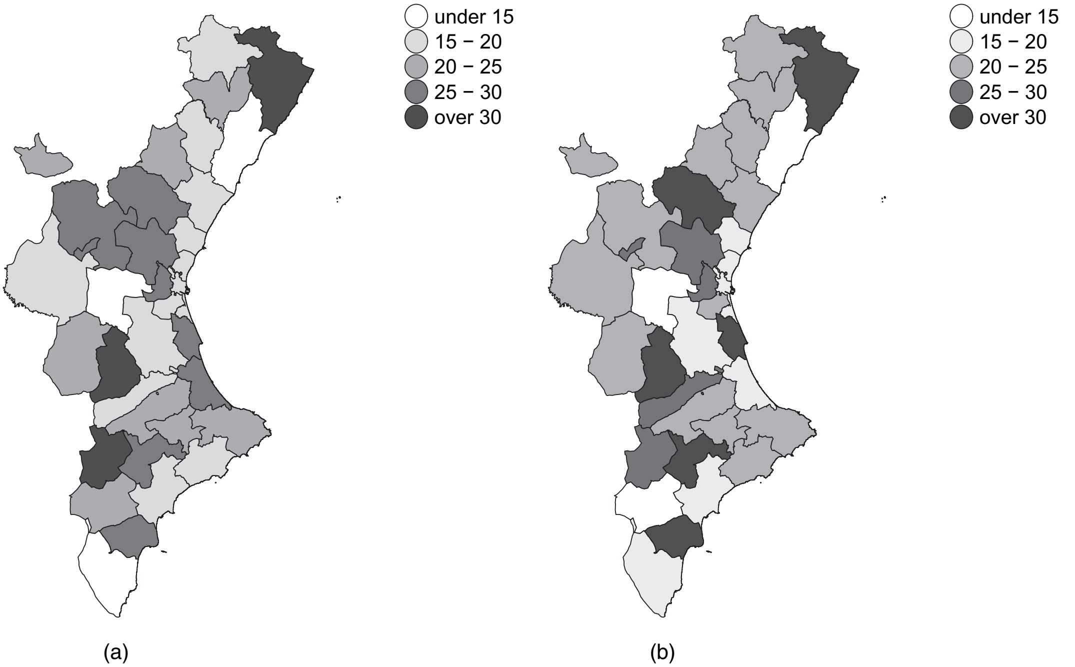


Fig. 9. Cartograms of EB estimates of poverty incidences for each *comarca* for (a) men and (b) women

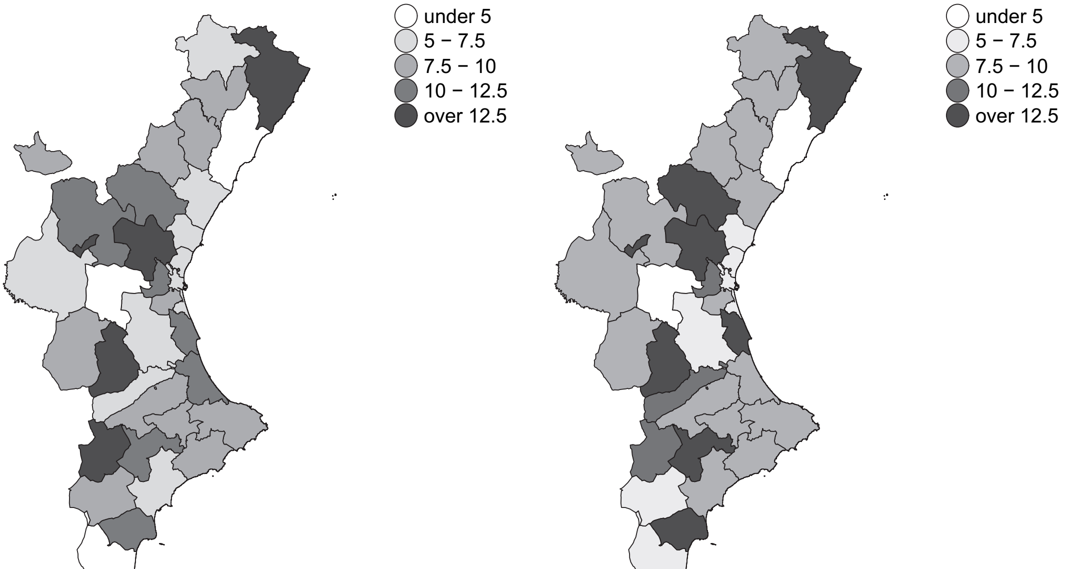


Fig. 10. Cartograms of EB estimates of poverty gaps for each *comarca* for (a) men and (b) women

the skew normal distribution or a mixture of normals. The bootstrap MSE estimators that are proposed here are also based on normality and might not be robust against departures from this distribution.

The decision to include or not random effects at any of the levels is typically based on the results of testing for the significance of the variances at each level or by comparing the models on the basis of goodness-of-fit measures. However, from our simulation study, which included the cases where only the domain or the subdomain effects are significant, which is equivalent to a case where covariates explain all the between-domain or between-subdomain variation, we have learnt that there is virtually no loss of efficiency by preserving a random factor even if it is not significant. EB estimators based on the twofold model perform very similarly to the true onefold model in those cases. However, if the user decides to apply a onefold model because of simplicity or software availability, we recommend including random effects for the domains where we want to estimate whether these domains are sampled. We recommend not to estimate in non-sampled areas because the model cannot be checked for them but, if estimates for subdomains with zero sample size are required, including the domain effects in the model leads to EB estimators for the non-sampled subdomains with greater efficiency.

For other territories that are not nested in the domains, EB estimators can also be readily obtained just by adding the corresponding observed and predicted values (through the conditional expectation) for the units in that territory.

In this application, no census or register with auxiliary microdata was available. In the case of categorical auxiliary variables, by virtue of expression (10), only counts of people in each category and subdomain are actually needed apart from survey data. This information was available from other statistical sources. However, even if models with stronger prediction power can be found in other applications, leading to highly efficient EB estimators, note that EB estimators use the regression errors (through the domain and subdomain effects) to correct the regression term. Thus, as our simulation results indicate, EB estimates perform well even in cases where the available covariates are not strong.

To calculate EB estimators in the application, we have used the counts of individuals in each labour force status category and each *comarca* by gender from other (more reliable) statistical sources. Thus, these counts have some uncertainty, which could be accounted for by considering twofold models with measurement error in the covariates. We leave this topic for further research.

Acknowledgements

We thank the Office for Social, Demographic and Economic Statistics of the Valencian Government for providing the real data that were used in the application. This work was partially supported by Spanish grant MTM2015-64842-P. J. N. K. Rao is partially supported by a grant from the Natural Sciences and Engineering Research Council of Canada.

References

Battese, G. E., Harter, R. M. and Fuller, W. A. (1988) An error-components model for prediction of county crop areas using survey and satellite data. *J. Am. Statist. Ass.*, **83**, 28–36.

Chambers, R., Salvati, N. and Tzavidis, N. (2012) M-quantile regression for binary data with application to small area estimation. *Working Paper 12-12, 2012, 24*. Centre for Statistical and Survey Methodology, University of Wollongong, Wollongong. (Available from `http://ro.uow.edu.au/cssmwp/101`.)

Chambers, R., Salvati, N. and Tzavidis, N. (2016) Semiparametric small area estimation for binary outcomes with application to unemployment estimation for local authorities in the UK. *J. R. Statist. Soc.* A, **179**, 453–479.

Citro, C. F. and Kalton, G. (eds) (2000) *Small-area Estimates of School-age Children in Poverty: Evaluation of Current Methodology*. Washington DC: National Academy Press.

Elbers, C., Lanjouw, J. O. and Lanjouw, P. (2003) Micro-level estimation of poverty and inequality. *Econometrica*, **71**, 355–364.
Esteban, M. D., Morales, D., Pérez, A. and Santamaría, L. (2012a) Two area-level time models for estimating small area poverty indicators. *J. Ind. Soc. Agric. Statist.*, **66**, 75–89.
Esteban, M. D., Morales, D., Pérez, A. and Santamaría, L. (2012b) Small area estimation of poverty proportions under area-level time models. *Computnl Statist. Data Anal.*, **56**, 2840–2855.
Fay, R. E. and Herriot, R. A. (1979) Estimation of income from small places: an application of James-Stein procedures to census data. *J. Am. Statist. Ass.*, **74**, 269–277.
Foster, J., Greer, J. and Thorbecke, E. (1984) A class of decomposable poverty measures. *Econometrica*, **52**, 761–766.
Ghosh, M. and Lahiri, P. (1988) Bayes and empirical Bayes analysis in multistage sampling. In *Statistical Decision Theory and Related Topics IV*, vol. 1 (eds S. S. Gupta and J. O. Berger), pp. 195–212. New York: Springer.
González-Manteiga, W., Lombardía, M. J., Molina, I., Morales, D. and Santamaría, L. (2007) Estimation of the mean squared error of predictors of small area linear parameters under a logistic mixed model. *Computnl Statist. Data Anal.*, **51**, 2720–2733.
González-Manteiga, W., Lombardía, M. J., Molina, I., Morales, D. and Santamaría, L. (2008) Bootstrap mean squared error of small-area EBLUP. *J. Statist. Computn Simuln*, **78**, 443–462.
Guadarrama, M., Molina, I. and Rao, J. N. K. (2014) A comparison of small area estimation methods for poverty mapping. *Statist. Transn*, **17**, 41–66.
Hobza, T. and Morales, D. (2016) Empirical best prediction under unit-level logit mixed models. *J. Off. Statist.*, **32**, 661–669.
Marhuenda, Y., Molina, I. and Morales, D. (2013) Small area estimation with spatio-temporal Fay-Herriot models. *Computnl Statist. Data Anal.*, **58**, 308–325.
Molina, I. and Marhuenda, Y. (2015) sae: an R package for small area estimation. *R J.*, **7**, 81–98.
Molina, I., Nandram, B. and Rao, J. N. K. (2015) Small area estimation of general parameters with application to poverty indicators: a hierarchical Bayes approach. *Ann. Appl. Statist.*, **8**, 852–885.
Molina, I. and Rao, J. N. K. (2010) Small area estimation of poverty indicators. *Can. J. Statist.*, **38**, 369–385.
Morales, D., Pagliarella, M. C. and Salvatore, R. (2015) Small area estimation of poverty indicators under partitioned area-level time models. *Statist. Oper. Res. Trans.*, **39**, 19–34.
Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. and R Core Team (2017) nlme: linear and nonlinear mixed effects models. *R Package Version 3.1-131*. (Available from `https://cran.r-project.org/web/packages/nlme/nlme.pdf`.)
Pratesi, M. (2016) *Analysis of Poverty Data by Small Area Estimation*. Hoboken: Wiley.
Rao, J. N. K. and Molina, I. (2015) *Small Area Estimation*, 2nd edn. Hoboken: Wiley.
Särndal, C. E., Swensson, B. and Wretman, J. (1992) *Model Assisted Survey Sampling*. Berlin: Springer.
Stukel, D. M. and Rao, J. N. K. (1997) Estimation of regression models with nested error structure and unequal error variances under two and three stage cluster sampling. *Statist. Probab. Lett.*, **35**, 401–407.
Tzavidis, N., Salvati, N., Pratesi, M. and Chambers, R. (2008) M-quantile models with application to poverty mapping. *Statist. Meth. Appl.*, **17**, 393–411.

Supporting information
Additional 'supporting information' may be found in the on-line version of this article:

'Supplementary material for manuscript "Poverty mapping in small areas under a two-fold nested error regression model"'.