WORKING PAPERS

Time stable small area estimates of general parameters under a unit-level model

Maria **GUADARRAMA**¹
Domingo **MORALES**²
Isabel **MOLINA**³

LISER Working Papers are intended to make research findings available and stimulate comments and discussion. They have been approved for circulation but are to be considered preliminary. They have not been edited and have not been subject to any peer review.

*The views expressed in this paper are those of the author(s) and do not necessarily reflect views of LISER.
Errors and omissions are the sole responsibility of the author(s).*

Time stable small area estimates of general parameters under a unit-level model ¹

María Guadarrama

Luxembourg Institute of Socio-Economic Research (LISER).

Domingo Morales

Operations Research Center, University Miguel Hernández de Elche, Spain.

Isabel Molina

Department of Statistics, University Carlos III de Madrid, Spain.

June 2020

Abstract

Longitudinal surveys collecting information on certain phenomena at several time points are very popular because they allow to analyze the changes over time. Data coming from those surveys often present correlation over time that should be accounted for by the considered statistical procedures. In fact, methods that account for the existing time correlation are expected to yield more stable small area estimates over time. Temporal stability is a desirable property of statistics that are published regularly, specially in certain applications like in poverty mapping, where poverty estimates for the same area with big jumps from one period to the next are rarely credible. This paper considers a unit-level temporal linear mixed model for small area estimation that includes random time effects nested within the usual area effects, following an autoregressive process of order 1, AR(1). Based on the proposed model, we obtain empirical best predictors of general area parameters, giving explicit expressions for some common poverty indicators. We also propose a parametric bootstrap method for estimating their mean square errors under the model. The proposed methods are studied through simulation experiments and illustrated with an application to poverty mapping in Spanish provinces using survey data from 2004-2006.

Keywords: Small area estimation; Empirical best predictor; Linear mixed models; Time correlation; Poverty mapping.

¹ This work was supported by the Spanish grants PGC2018-096840-B-I00, MTM2015-69638-R and MTM2015-64842-P from Ministerio de Economía y Competitividad.

Time stable small area estimates of general parameters under a unit-level model

María Guadarrama^{a,*}, Domingo Morales^b and Isabel Molina^c,

^a Luxembourg Institute of Socio-Economic Research (LISER).

^bOperations Research Center, University Miguel Hernández de Elche, Spain.

^cDepartment of Statistics, University Carlos III de Madrid, Spain.

June 18, 2020

Abstract

Longitudinal surveys collecting information on certain phenomena at several time points are very popular because they allow to analyze the changes over time. Data coming from those surveys often present correlation over time that should be accounted for by the considered statistical procedures. In fact, methods that account for the existing time correlation are expected to yield more stable small area estimates over time. Temporal stability is a desirable property of statistics that are published regularly, specially in certain applications like in poverty mapping, where poverty estimates for the same area with big jumps from one period to the next are rarely credible. This paper considers a unit-level temporal linear mixed model for small area estimation that includes random time effects nested within the usual area effects, following an autoregressive process of order 1, AR(1). Based on the proposed model, we obtain empirical best predictors of general area parameters, giving explicit expressions for some common poverty indicators. We also propose a parametric bootstrap method for estimating their mean square errors under the model. The proposed methods are studied through simulation experiments and illustrated with an application to poverty mapping in Spanish provinces using survey data from 2004-2006.

Key words: Small area estimation; Empirical best predictor; Linear mixed models; Time correlation; Poverty mapping.

1 Introduction

States publish statistics on labor market, living conditions or any other topic of public interest, based on the information collected through specific periodic surveys. These statistics are required more and more often at a finer disaggregation level. However, the sample sizes at certain levels may be too small to obtain accurate direct estimators. Small area estimation (SAE) techniques address the estimation of parameters in areas with insufficient sample size. These techniques

*Corresponding author. Luxembourg Institute of Socio-Economic Research. Address: 11, Porte des Sciences, L-4366 Esch-sur-Alzette/Belval, Luxembourg, Tf: +352 585855605. E-mail: maria.guadarrama@liser.lu

complement the sample data of a given domain or area with information from the rest of the domains or areas of the population. This is done by considering models that link all the domains or areas of interest through common model parameters. These parameters represent the common factors that affect the variable/s of interest in all the areas. Models can be established either for the units (unit-level models) when unit-level auxiliary information is available, or for the areas (area-level models) when only area aggregates are available. The book by Rao & Molina (2015) gives an extended review of the main SAE procedures until the publication date.

Statistical offices wish to publish regular statistics that are stable over time, since unstable statistics have low credibility. By considering historical data in the estimation process and accounting for the potential time correlation, it is possible to obtain estimates that are more stable over time than those based solely on the cross-sectional data. Small area models accounting for time correlation have been previously considered by several authors. For example, at the area level, Choudhry & Rao (1989) considered a model with autocorrelated sampling errors. Pfeiffermann & Burck (1990) and Singh et al. (1991) studied models with time-varying random slopes, whereas Rao & Yu (1994) proposed a model with AR(1)-correlated area-time random intercepts. Ghosh et al. (1996) proposed a slightly more complicated time correlated model, Datta et al. (1999), You & Rao (2000), You et al. (2001) and Datta et al. (2002) introduced variants of the Rao-Yu model, and more recently, Esteban et al. (2012a,b), Marhuenda et al. (2013), Morales et al. (2015), Boubeta et al. (2017) and Benavent & Morales (2020) applied temporal area-level mixed models to poverty mapping.

In the case of unit-level models, the literature is more scarce, perhaps because application of these models in practice requires historical unit-level information, which may not be always available, or not in a comparable form. Note that the methodology of the surveys is updated in some periods, making the surveys performed before and after the changes not comparable. Computational issues may also limit their applicability, since the size of the unit-level historical data sets to be processed can be really large. As examples of unit-level models, we can cite the linear mixed model considered by Morales & Santamaría (2019) for estimation of linear parameters or the logit mixed model for proportions studied by Hobza et al. (2018).

The appropriate SAE technique in each particular application depends strongly on the functional form of the objective parameter/s. For a long time, the literature focused on the estimation of linear parameters or on linear combinations of the fixed and random effects of the assumed model. The theory for those types of parameters gave rise to the empirical best linear unbiased predictors (EBLUP), with a very fruitful literature on the subject. However, many parameters, such as poverty or inequality indicators, are not linear. Molina & Rao (2010) proposed empirical best predictors (EBP) of general non-linear parameters based on a unit-level model. Concretely, Molina & Rao (2010), Molina et al. (2014) and Guadarrama et al. (2016) assumed that a transformation of the study variable follows a nested error regression (NER) model. With the purpose of simultaneously estimating domain and subdomain parameters, Marhuenda et al. (2017) derived EBPs under two-fold NER models. The selection of the transformation is studied by Rojas-Perilla et al. (2020). Instead of searching for an appropriate transformation to achieve normality, which might not be easy, another approach is to consider

other distributions that fit better the data. The EBP methodology was extended by Graf et al. (2019) and Diallo & Rao (2018) to unit-level mixed models with skewed distributions or by Hobza et al. (2020) to unit-level generalized linear mixed models.

This paper fills a gap in the current literature by introducing EBPs of general domain parameters based on a unit-level temporal mixed model including area and area-time random effects with AR(1)-type time correlation. The paper is organized as follows. Section 2 introduces the considered small area model. Section 3 describes the EBPs of general domain parameters, deriving the conditional distribution required to calculate them. Section 4 deals with a particular class of domain parameters that includes poverty incidences, poverty gaps and average income, spelling out the corresponding EBPs. Section 5 proposes a parametric bootstrap approach to estimate the MSEs of the EBPs. Section 6 presents two simulation experiments designed to investigate the cross-sectional and longitudinal behavior of the new predictors. Section 7 applies the proposed methodology to data from the Spanish survey on income and living conditions (EU-SILC) from years 2004, 2005 and 2006, providing estimates of poverty incidences and gaps for women in the Spanish provinces. Section 8 draws some conclusions. Finally, Appendix A gives the details on the derivations of the expressions in Section 3 and Appendix B includes detailed results from the application to Spanish EU-SILC data.

2 Temporal unit-level model for small area estimation

This section describes a temporal unit-level model for estimation of general parameters in small domains. The model considers that the values of the target variable in the population units, for each time period, follows a unit-level linear mixed model with random domain effects following usual assumptions, and time effects nested within the domain effects, following independent AR(1) processes for each domain. Specifically, consider that we have data from $T > 0$ time instants. The total population at each time point t is supposed to be composed of m areas or domains, $U_{1t}, \dots, U_{mt}$, of sizes $N_{1t}, \dots, N_{mt}$ respectively, $t = 1, \dots, T$. Let y_{itj} be the value of the target variable for j -th unit from the population of domain i at time instant t , U_{it} . The target parameters have the form $\delta_{it} = \delta_{it}(y_{it1}, \dots, y_{itN_{it}})$ for each area $i = 1, \dots, m$ at time instant $t = 1, \dots, T$, although in some applications only those at the last time point, δ_{iT} , may be of interest.

For efficient estimation of δ_{it} for each area i at certain instant t , we need to “borrow strength”. In this case, we borrow strength from the other areas and time instants, by considering the following population model:

$$y_{itj} = \mathbf{x}_{itj}'\boldsymbol{\beta} + v_i + u_{it} + w_{itj}^{-1/2}e_{itj}, \quad i = 1, \dots, m, \quad t = 1, \dots, T, \quad j = 1, \dots, N_{it}, \quad (1)$$

where $\mathbf{x}_{itj}$ is a $p \times 1$ vector containing the values of p auxiliary variables for unit j in time t and domain i , $w_{itj} > 0$ is a known heteroscedasticity weight for the same unit and $\boldsymbol{\beta}$ is the $p \times 1$ vector of regression coefficients. The area effects v_i and the unit level errors e_{itj} are assumed to follow usual assumptions, that is, they are mutually independent, with $v_i \stackrel{iid}{\sim} N(0, \sigma_v^2)$ and $e_{itj} \stackrel{iid}{\sim} N(0, \sigma_e^2)$. The time effects within each area i are also mutually independent of the area

effects v_i and the individual errors e_{itj} , with $\mathbf{u}_i = (u_{i1}, \dots, u_{iT})' \stackrel{iid}{\sim} N(\mathbf{0}_T, \sigma_u^2 \mathbf{\Omega}_u)$, where

$$\mathbf{\Omega}_u = \frac{1}{1 - \rho^2} \begin{bmatrix} 1 & \rho & \dots & \rho^{T-2} & \rho^{T-1} \\ \rho & 1 & \ddots & & \rho^{T-2} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \rho^{T-2} & & \ddots & 1 & \rho \\ \rho^{T-1} & \rho^{T-2} & \dots & \rho & 1 \end{bmatrix},$$

for $|\rho| < 1$. Note that the area effects v_i are independent for all i , and the same occurs for the vectors $\mathbf{u}_i$ of time effects within each area i . Thus, even if the time effects u_{it} within the areas present an AR(1) correlation structure, the values y_{itj} and y_{dtj} for different areas ($d \neq i$) are independent, yielding a block-diagonal covariance structure with blocks corresponding to the areas. Model (1) is called here AR(1)-NER model.

Let us now define the following area vectors and matrices, where $\text{col}(a_i)$ means stacking the a_i 's in a column and $\text{diag}(a_i)$ denotes constructing a block-diagonal matrix with the a_i 's,

$$\begin{aligned} \mathbf{y}_i &= \text{col}_{1 \leq t \leq T} \left(\text{col}_{1 \leq j \leq N_{it}} (y_{itj}) \right), & \mathbf{e}_i &= \text{col}_{1 \leq t \leq T} \left(\text{col}_{1 \leq j \leq N_{it}} (e_{itj}) \right), \\ \mathbf{X}_i &= \text{col}_{1 \leq t \leq T} \left(\text{col}_{1 \leq j \leq N_{it}} (\mathbf{x}'_{itj}) \right), & \mathbf{W}_i &= \text{diag}_{1 \leq t \leq T} \left(\text{diag}_{1 \leq j \leq N_{it}} (w_{itj}) \right). \end{aligned}$$

Finally, define also $\mathbf{A}_i = \text{diag}_{1 \leq t \leq T}(\mathbf{1}_{N_{it}})$, where $\mathbf{1}_a$ denotes the $a \times 1$ vector of ones. The population model (1) can be expressed in terms of these area vectors and matrices, as

$$\mathbf{y}_i = \mathbf{X}_i \boldsymbol{\beta} + \mathbf{1}_{N_i} v_i + \mathbf{A}_i \mathbf{u}_i + \mathbf{W}_i^{-1/2} \mathbf{e}_i, \quad i = 1, \dots, m. \quad (2)$$

Equivalently, $\mathbf{y}_i \stackrel{ind}{\sim} N(\mathbf{X}_i \boldsymbol{\beta}, \mathbf{V}_i)$, for

$$\mathbf{V}_i = \text{Var}(\mathbf{y}_i) = \sigma_v^2 \mathbf{1}_{N_i} \mathbf{1}'_{N_i} + \sigma_u^2 \mathbf{A}_i \mathbf{\Omega}_u \mathbf{A}'_i + \sigma_e^2 \mathbf{W}_i^{-1}, \quad i = 1, \dots, m. \quad (3)$$

Defining now the population vectors and matrices

$$\begin{aligned} \mathbf{y} &= \text{col}_{1 \leq i \leq m} (\mathbf{y}_i), & \mathbf{e} &= \text{col}_{1 \leq i \leq m} (\mathbf{e}_i), & \mathbf{v} &= \text{col}_{1 \leq i \leq m} (v_i), & \mathbf{u} &= \text{col}_{1 \leq i \leq m} (\mathbf{u}_i), \\ \mathbf{X} &= \text{col}_{1 \leq i \leq m} (\mathbf{X}_i), & \mathbf{W} &= \text{diag}_{1 \leq i \leq m} (\mathbf{W}_i), & \mathbf{Z}_1 &= \text{diag}_{1 \leq i \leq m} (\mathbf{1}_{N_i}), & \mathbf{Z}_2 &= \text{diag}_{1 \leq i \leq m} (\mathbf{A}_i), \end{aligned}$$

the model can be also expressed as

$$\mathbf{y} = \mathbf{X} \boldsymbol{\beta} + \mathbf{Z}_1 \mathbf{v} + \mathbf{Z}_2 \mathbf{u} + \mathbf{W}^{-1/2} \mathbf{e}, \quad (4)$$

where $\mathbf{v} \sim N(\mathbf{0}_m, \sigma_v^2 \mathbf{I}_m)$, $\mathbf{u} \sim N(\mathbf{0}_{mT}, \mathbf{I}_{mT} \otimes \sigma_u^2 \mathbf{\Omega}_u)$ and $\mathbf{e} \sim N(\mathbf{0}_N, \sigma_e^2 \mathbf{I}_N)$ are independent. Finally, the population model (4) means that $\mathbf{y} \sim N(\mathbf{X} \boldsymbol{\beta}, \mathbf{V})$, where the covariance matrix $\mathbf{V}$ is block-

diagonal; specifically, by (4), it holds that

$$\mathbf{V} = \text{Var}(\mathbf{y}) = \mathbf{Z}_1 \text{Var}(\mathbf{v}) \mathbf{Z}_1' + \mathbf{Z}_2 \text{Var}(\mathbf{u}) \mathbf{Z}_2' + \sigma_e^2 \mathbf{W}^{-1} = \text{diag}_{1 \leq i \leq m} (\mathbf{V}_1, \dots, \mathbf{V}_m),$$

where the blocks $\mathbf{V}_i$ in the diagonal of $\mathbf{V}$ are given in (3).

In practice, inference is performed based on a sample. We assume that a sample s_{it} of size $n_{it} \leq N_{it}$ is drawn from the population in area i at time instant t , U_{it} . We denote by $s_i = s_{i1} \cup \dots \cup s_{iT}$ the sample from area i overall the study period, of size $n_i = \sum_{t=1}^T n_{it}$. We consider that an area i is either sampled for all the considered time instants $t = 1, \dots, T$, or not sampled in any of them, in which case $n_i = 0$. The overall sample is denoted as $s = s_1 \cup \dots \cup s_m$ and has size $n = \sum_{i=1}^m n_i$.

Considering that there is no sample selection bias, the sub-vector $\mathbf{y}_s$ of $\mathbf{y}$ corresponding to the sample units s follows the marginal model obtained from (1) and given by

$$y_{itj} = \mathbf{x}_{itj}' \boldsymbol{\beta} + v_i + u_{it} + w_{itj}^{-1/2} e_{itj}, \quad i = 1, \dots, m; \quad t = 1, \dots, T; \quad j \in s_{it}. \quad (5)$$

Similarly as before, we define the area vectors and matrices for the sample units

$$\begin{aligned} \mathbf{y}_{is} &= \text{col}_{1 \leq t \leq T} (\text{col}_{j \in s_{it}} (y_{itj})), & \mathbf{e}_{is} &= \text{col}_{1 \leq t \leq T} (\text{col}_{j \in s_{it}} (e_{itj})), \\ \mathbf{X}_{is} &= \text{col}_{1 \leq t \leq T} (\text{col}_{j \in s_{it}} (\mathbf{x}_{itj}')), & \mathbf{W}_{is} &= \text{diag}_{1 \leq t \leq T} (\text{diag}_{j \in s_{it}} (w_{itj}^{-1/2})) \end{aligned}$$

and the additional sample matrix $\mathbf{A}_{is} = \text{diag}_{1 \leq t \leq T} (\mathbf{1}_{n_{it}})$. Then, the sample model (5) can be written as

$$\mathbf{y}_{is} = \mathbf{X}_{is} \boldsymbol{\beta} + \mathbf{1}_{n_i} v_i + \mathbf{A}_{is} \mathbf{u}_i + \mathbf{W}_{is}^{-1/2} \mathbf{e}_{is},$$

where $\mathbf{e}_{is} \stackrel{\text{ind}}{\sim} N(\mathbf{0}_{n_i}, \sigma_e^2 \mathbf{I}_{n_i})$, or, equivalently, $\mathbf{y}_{is} \stackrel{\text{ind}}{\sim} N(\mathbf{X}_{is} \boldsymbol{\beta}, \mathbf{V}_{is})$, for

$$\mathbf{V}_{is} = \text{Var}(\mathbf{y}_{is}) = \sigma_v^2 \mathbf{1}_{n_i} \mathbf{1}_{n_i}' + \sigma_u^2 \mathbf{A}_{is} \boldsymbol{\Omega}_u \mathbf{A}_{is}' + \sigma_e^2 \mathbf{W}_{is}^{-1}, \quad i = 1, \dots, m. \quad (6)$$

Further, defining the overall sample vectors and matrices

$$\begin{aligned} \mathbf{y}_s &= \text{col}_{1 \leq i \leq m} (\mathbf{y}_{is}), & \mathbf{X}_s &= \text{col}_{1 \leq i \leq m} (\mathbf{X}_{is}), & \mathbf{e}_s &= \text{col}_{1 \leq i \leq m} (\mathbf{e}_{is}), \\ \mathbf{W}_s &= \text{diag}_{1 \leq i \leq m} (\mathbf{W}_{is}), & \mathbf{Z}_{1s} &= \text{diag}_{1 \leq i \leq m} (\mathbf{1}_{n_i}), & \mathbf{Z}_{2s} &= \text{diag}_{1 \leq i \leq m} (\mathbf{A}_i), \end{aligned}$$

the sample model (5) is now

$$\mathbf{y}_s = \mathbf{X}_s \boldsymbol{\beta} + \mathbf{Z}_{1s} \mathbf{v} + \mathbf{Z}_{2s} \mathbf{u} + \mathbf{W}_s^{-1/2} \mathbf{e}_s, \quad (7)$$

where $\mathbf{e}_s \sim N(\mathbf{0}_N, \sigma_e^2 \mathbf{I}_N)$. Equivalently, we have $\mathbf{y}_s \sim N(\mathbf{X}_s \boldsymbol{\beta}, \mathbf{V}_s)$ where, similarly as for the whole population vector $\mathbf{y}$, the covariance matrix has a block-diagonal structure, that is,

$$\mathbf{V}_s = \text{Var}(\mathbf{y}_s) = \mathbf{Z}_{1s} \text{Var}(\mathbf{v}) \mathbf{Z}_{1s}' + \mathbf{Z}_{2s} \text{Var}(\mathbf{u}) \mathbf{Z}_{2s}' + \sigma_e^2 \mathbf{W}_s^{-1} = \text{diag}_{1 \leq i \leq m} (\mathbf{V}_{1s}, \dots, \mathbf{V}_{ms}),$$

with blocks $\mathbf{V}_{is}$ given in (6).

Let us now describe the model for the units that are not in the sample. Let $r_{it} = U_{it} - s_{it}$ be the set of nonsampled units from area i at time t . Similarly, we denote by $r = U - s$ the subset of units out of s . Similarly as for the sample, the sub-vector $\mathbf{y}_r$ of $\mathbf{y}$ for the nonsampled units also follows the corresponding marginal model obtained from (1), given by

$$y_{itj} = \mathbf{x}'_{itj}\boldsymbol{\beta} + v_i + u_{it} + w_{itj}^{-1/2}e_{itj}, \quad i = 1, \dots, m, \quad t = 1, \dots, T, \quad j \in r_{it}. \quad (8)$$

Similarly as done for the population and the sample, we define now vectors and matrices for the units that are not in the sample from area i , as

$$\begin{aligned} \mathbf{y}_{ir} &= \text{col}_{1 \leq t \leq T} (\text{col}_{j \in r_{it}} (y_{itj})), & \mathbf{e}_{ir} &= \text{col}_{1 \leq t \leq T} (\text{col}_{j \in r_{it}} (e_{itj})), \\ \mathbf{X}_{ir} &= \text{col}_{1 \leq t \leq T} (\text{col}_{j \in r_{it}} (\mathbf{x}'_{itj})), & \mathbf{W}_{ir} &= \text{diag}_{1 \leq t \leq T} (\text{diag}_{j \in r_{it}} (w_{itj})), \end{aligned}$$

and $\mathbf{A}_{ir} = \text{diag}_{1 \leq t \leq T} (\mathbf{1}_{N_{it}-n_{it}})$. Then, the model for the area vector of nonsampled units $\mathbf{y}_{ir}$ is

$$\mathbf{y}_{ir} = \mathbf{X}_{ir}\boldsymbol{\beta} + \mathbf{1}_{N_i-n_i}v_i + \mathbf{A}_{ir}\mathbf{u}_i + \mathbf{W}_{ir}^{-1/2}\mathbf{e}_{ir}. \quad (9)$$

where $\mathbf{e}_{ir} \stackrel{\text{ind}}{\sim} N(\mathbf{0}_{N_i-n_i}, \sigma_e^2 \mathbf{I}_{N_i-n_i})$. This means that $\mathbf{y}_{ir} \stackrel{\text{ind}}{\sim} N(\mathbf{X}_{ir}\boldsymbol{\beta}, \mathbf{V}_{ir})$, where in this case

$$\mathbf{V}_{ir} = \text{Var}(\mathbf{y}_{ir}) = \sigma_v^2 \mathbf{1}_{N_i-n_i} \mathbf{1}'_{N_i-n_i} + \sigma_u^2 \mathbf{A}_{ir} \boldsymbol{\Omega}_u \mathbf{A}'_{ir} + \sigma_e^2 \mathbf{W}_{ir}^{-1}, \quad i = 1, \dots, m. \quad (10)$$

Defining additionally

$$\begin{aligned} \mathbf{y}_r &= \text{col}_{1 \leq i \leq m} (\mathbf{y}_{ir}), & \mathbf{X}_r &= \text{col}_{1 \leq i \leq m} (\mathbf{X}_{ir}), & \mathbf{e}_r &= \text{col}_{1 \leq i \leq m} (\mathbf{e}_{ir}), \\ \mathbf{W}_r &= \text{diag}_{1 \leq i \leq m} (\mathbf{W}_{ir}), & \mathbf{Z}_{1r} &= \text{diag}_{1 \leq i \leq m} (\mathbf{1}_{N_i-n_i}), & \mathbf{Z}_{2r} &= \text{diag}_{1 \leq i \leq m} (\mathbf{A}_i), \end{aligned}$$

the model for the nonsampled units then reads

$$\mathbf{y}_r = \mathbf{X}_r\boldsymbol{\beta} + \mathbf{Z}_{1r}\mathbf{v} + \mathbf{Z}_{2r}\mathbf{u} + \mathbf{W}_r^{-1/2}\mathbf{e}_r, \quad (11)$$

where $\mathbf{e}_r \sim N(\mathbf{0}_N, \sigma_e^2 \mathbf{I}_{N-n})$. Again, we can write $\mathbf{y}_r \sim N(\mathbf{X}_r\boldsymbol{\beta}, \mathbf{V}_r)$ where, similarly as before,

$$\mathbf{V}_r = \text{Var}(\mathbf{y}_r) = \mathbf{Z}_{1r} \text{Var}(\mathbf{v}) \mathbf{Z}'_{1r} + \mathbf{Z}_{2r} \text{Var}(\mathbf{u}) \mathbf{Z}'_{2r} + \sigma_e^2 \mathbf{W}_r^{-1} = \text{diag}_{1 \leq i \leq m} (\mathbf{V}_{1r}, \dots, \mathbf{V}_{mr}),$$

with blocks $\mathbf{V}_{ir}$ given in (10).

3 Empirical best predictors of general area parameters

For the target parameter of area i at time instant t , $\delta_{it} = \delta_{it}(y_{it1}, \dots, y_{itN_{it}})$, the best predictor is the function of the sample data $\mathbf{y}_s$ that minimizes the mean squared error (MSE) under the assumed model, $\text{MSE}(\hat{\delta}_{it} - \delta_{it})^2$. It is well known that this best predictor is given by $\hat{\delta}_{it}^{bp} =$

$E_{\mathbf{y}_r}(\delta_{it}|\mathbf{y}_s)$, where the expectation is taken with respect to the distribution of the nonsampled vector $\mathbf{y}_r$ given the sample data $\mathbf{y}_s$. Under the population model (1), the distribution of $\mathbf{y}_r|\mathbf{y}_s$ depends on the vector $\boldsymbol{\theta} = (\boldsymbol{\beta}', \sigma_v^2, \sigma_u^2, \rho, \sigma_e^2)'$ of model parameters, that is,

$$\hat{\delta}_{it}^{bp} = E_{\mathbf{y}_r}(\delta_{it}|\mathbf{y}_s; \boldsymbol{\theta}).$$

In practice, these model parameters are unknown and need to be estimated using the sample data $\mathbf{y}_s$. Let $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\beta}}', \hat{\sigma}_v^2, \hat{\sigma}_u^2, \hat{\rho}, \hat{\sigma}_e^2)'$ be a consistent estimator of $\boldsymbol{\theta}$ based on $\mathbf{y}_s$. The empirical best predictor (EBP) of δ_{it} is obtained replacing $\boldsymbol{\theta}$ by $\hat{\boldsymbol{\theta}}$ in the best predictor, that is,

$$\hat{\delta}_{it}^{ebp} = E_{\mathbf{y}_r}(\delta_{it}|\mathbf{y}_s; \hat{\boldsymbol{\theta}}). \quad (12)$$

An estimators of $\boldsymbol{\theta} = (\boldsymbol{\beta}', \sigma_v^2, \sigma_u^2, \rho, \sigma_e^2)'$ under the sample AR(1)-NER model given in (5) can be obtained e.g. by the maximum likelihood (ML) or restricted/residual ML (REML) methods, see Morales & Santamaría (2019).

The EBPs are obtained taking expectation with respect to the distribution of the nonsample vector $\mathbf{y}_r$ given the sample data $\mathbf{y}_s$ in the decomposition $\mathbf{y} = (\mathbf{y}'_s, \mathbf{y}'_r)'$. Similarly as $\mathbf{y}$, its covariance matrix $\mathbf{V} = \text{Var}(\mathbf{y})$ can be also decomposed in blocks for the covariances of pairs of sample units $\mathbf{V}_s$, covariances of sample and nonsample units $\mathbf{V}_{sr}$, and covariances of pairs of nonsample units $\mathbf{V}_r$, as

$$\mathbf{V} = \begin{pmatrix} \mathbf{V}_s & \mathbf{V}_{sr} \\ \mathbf{V}_{rs} & \mathbf{V}_r \end{pmatrix}.$$

Since $\mathbf{y} \sim N(\mathbf{X}\boldsymbol{\beta}, \mathbf{V})$, the distribution of $\mathbf{y}_r$ given $\mathbf{y}_s$ is then

$$\mathbf{y}_r|\mathbf{y}_s \sim N(\boldsymbol{\mu}_{r|s}, \mathbf{V}_{r|s}),$$

where the conditional mean vector and covariance matrix are respectively given by

$$\boldsymbol{\mu}_{r|s} = \mathbf{X}_r\boldsymbol{\beta} + \mathbf{V}_{rs}\mathbf{V}_s^{-1}(\mathbf{y}_s - \mathbf{X}_s\boldsymbol{\beta}), \quad (13)$$

$$\mathbf{V}_{r|s} = \mathbf{V}_r - \mathbf{V}_{rs}\mathbf{V}_s^{-1}\mathbf{V}_{sr}. \quad (14)$$

In the following, we spell out the expressions of the above conditional mean vector and covariance matrix. First of all, we need an expression for the inverse of the sample covariance matrix $\mathbf{V}_s$. Note that for unit-level models, and specially in this model dealing with several time periods, the sample covariance matrix may be really large. In this case, $\mathbf{V}_s$ is actually $n \times n$, where $n = \sum_{i=1}^m \sum_{t=1}^T n_{it}$. Inverting numerically this matrix in a real application may not be possible with standard software due to the large dimension, so here we give an explicit expression for the inverse, which will be also employed to spell out the above conditional mean vector and covariance matrix. Since $\mathbf{V}_s$ is block-diagonal, with blocks $\mathbf{V}_{is}$ given in (6), $\mathbf{V}_s^{-1}$ is also block-diagonal, with blocks $\mathbf{V}_{is}^{-1}$. Still, $\mathbf{V}_{is}^{-1}$ may be pretty large in some applications, so we express it here in terms of the inverse of a much smaller matrix, see Appendix A for the details. Let $w_{its} = \sum_{j \in s_{it}} w_{itj}$ be the sum of heteroscedasticity weights for the sample

units from area i at time instant t , $\mathbf{w}_{its} = \mathbf{W}_{its} \mathbf{1}_{n_{it}}$ the column vector of the same weights, $\mathbf{G}_{is} = \mathbf{A}'_{is} \mathbf{W}_{is} \mathbf{A}_{is} = \text{diag}_{1 \leq t \leq T}(w_{its})$ and

$$\Delta_i = \Omega_u^{-1} + \frac{\sigma_u^2}{\sigma_e^2} \mathbf{G}_{is}.$$

The inverse of Ω_u is given in (29). Let also define the matrices $\mathbf{B}_{is} = \mathbf{W}_{is} \mathbf{A}_{is} = \text{diag}_{1 \leq t \leq T}(\mathbf{w}_{its})$ and

$$\mathbf{W}_{is}^c = \mathbf{W}_{is} - \frac{\sigma_u^2}{\sigma_e^2} \mathbf{B}_{is} \Delta_i^{-1} \mathbf{B}_{is}',$$

the vector $\mathbf{w}_{is}^c = \mathbf{W}_{is}^c \mathbf{1}_{n_i}$ and the scalar $w_{is}^c = \mathbf{1}'_{n_i} \mathbf{W}_{is}^c \mathbf{1}_{n_i}$. Then, the inverse of $\mathbf{V}_s$ is given by

$$\mathbf{V}_{is}^{-1} = \frac{1}{\sigma_e^2} \left[\mathbf{W}_{is}^c - \frac{\gamma_{iw}^c}{w_{is}^c} \mathbf{w}_{is}^c (\mathbf{w}_{is}^c)' \right], \quad (15)$$

where $\gamma_{iw}^c = \sigma_v^2 / (\sigma_v^2 + \sigma_e^2 / w_{is}^c)$.

Now it is easy to see that $\mathbf{V}_{rs} = \text{cov}(\mathbf{y}_r, \mathbf{y}_s)$ is also block-diagonal, with blocks given by

$$\mathbf{V}_{irs} = \sigma_v^2 \mathbf{1}_{N_i - n_i} \mathbf{1}'_{n_i} + \sigma_u^2 \mathbf{A}_{ir} \Omega_u \mathbf{A}'_{is}, \quad i = 1, \dots, m, \quad (16)$$

see Appendix A.2 for details. Replacing the expressions of $\mathbf{V}_{rs}$ and $\mathbf{V}_s^{-1}$ in the conditional mean vector (13) and writing $\boldsymbol{\mu}_{r|s} = \text{col}_{1 \leq i \leq m}(\boldsymbol{\mu}_{ir|s})$, if area i is sampled (i.e., $n_i > 0$), we have

$$\boldsymbol{\mu}_{ir|s} = \mathbf{X}_{ir} \boldsymbol{\beta} + \frac{\gamma_{iw}^c}{w_{is}^c} (\mathbf{w}_{is}^c)' (\mathbf{y}_{is} - \mathbf{X}_{is} \boldsymbol{\beta}) \mathbf{1}_{N_i - n_i} + \frac{\sigma_u^2}{\sigma_e^2} \mathbf{A}_{ir} \Omega_u \left(\mathbf{B}_{is}^c - \frac{\gamma_{iw}^c}{w_{is}^c} \mathbf{w}_{is}^c (\mathbf{g}_{is}^c)' \right)' (\mathbf{y}_{is} - \mathbf{X}_{is} \boldsymbol{\beta}), \quad (17)$$

see Appendix A.2. Note also that $\boldsymbol{\mu}_{ir|s} = \text{col}_{1 \leq t \leq T}(\boldsymbol{\mu}_{itr|s})$, and we denote the elements of $\boldsymbol{\mu}_{itr|s}$ as $\mu_{itj|s}$, $j \in r_{it}$. From (17), we can see that the conditional mean $\mu_{itj|s}$ is composed of the regression term $\mathbf{x}'_{itj} \boldsymbol{\beta}$, the predicted area effect $\hat{v}_i = \gamma_{iw}^c (\mathbf{w}_{is}^c)' (\mathbf{y}_{is} - \mathbf{X}_{is} \boldsymbol{\beta}) / w_{is}^c$ and the predicted time effect $\hat{u}_{it}$ obtained as the t -th component of $\hat{\mathbf{u}}_i = \frac{\sigma_u^2}{\sigma_e^2} \Omega_u \left(\mathbf{B}_{is}^c - \frac{\gamma_{iw}^c}{w_{is}^c} \mathbf{w}_{is}^c (\mathbf{g}_{is}^c)' \right)' (\mathbf{y}_{is} - \mathbf{X}_{is} \boldsymbol{\beta})$. For a nonsampled area ($n_i = 0$), the conditional mean vector is simply $\boldsymbol{\mu}_{ir|s} = \mathbf{X}_{ir} \boldsymbol{\beta}$.

Finally, the conditional covariance matrix $\mathbf{V}_{r|s}$ is again block-diagonal, that is, $\mathbf{V}_{r|s} = \text{diag}_{1 \leq i \leq m}(\mathbf{V}_{ir|s})$, where, if area i is sampled ($n_i > 0$), then we have

$$\mathbf{V}_{ir|s} = \sigma_e^2 \mathbf{W}_{ir}^{-1} + \sigma_v^2 (1 - \gamma_{iw}^c) \mathbf{1}_{N_i - n_i} \mathbf{1}'_{N_i - n_i} + \sigma_u^2 \mathbf{A}_{ir} \mathbf{T}_{ir|s} \mathbf{A}'_{ir}, \quad (18)$$

where

$$\mathbf{T}_{ir|s} = \Omega_u - \frac{\gamma_{iw}^c}{w_{is}^c} (\mathbf{V}_{2i} \Omega_u + \Omega_u \mathbf{V}_{2i}') - \frac{\sigma_u^2}{\sigma_e^2} \Omega_u \mathbf{V}_{4i} \Omega_u,$$

for

$$\mathbf{V}_{2i} = \mathbf{1}_T (\mathbf{g}_{is}^c)', \quad \mathbf{V}_{4i} = \mathbf{G}_{is}^c - \frac{\gamma_{iw}^c}{w_{is}^c} \mathbf{g}_{is}^c (\mathbf{g}_{is}^c)'.$$

If area i is nonsampled ($n_i = 0$), we have

$$\mathbf{V}_{ir|s} = \sigma_e^2 \mathbf{W}_i^{-1} + \sigma_v^2 \mathbf{1}_{N_i} \mathbf{1}'_{N_i} + \sigma_u^2 \mathbf{A}_i \Omega_u \mathbf{A}_i'. \quad (19)$$

Let $v_{itj|s}^2$ be the element in the diagonal of $\mathbf{V}_{ir|s}$ corresponding to time period t and nonsampled unit $j \in r_{it}$. Then, we have

$$v_{itj|s}^2 = \sigma_e^2 w_{itj}^{-1} + \sigma_v^2 (1 - \gamma_{iw}^c) + \sigma_u^2 \mathbf{b}_t' \mathbf{T}_{ir|s} \mathbf{b}_t, \quad (20)$$

for $\mathbf{b}_t = (\mathbf{0}_{t-1}', 1, \mathbf{0}_{T-t}')'$, where $\mathbf{b}_t' \mathbf{T}_{ir|s} \mathbf{b}_t$ is the t -th diagonal element of $\mathbf{T}_{ir|s}$.

Even if the conditional distribution $\mathbf{y}_r | \mathbf{y}_s$ is described above, the expectation defining the EBP in (12) may not be easy to calculate in a real application, where the dimension of the multivariate vectors $\mathbf{y}_{ir}$ may be really large. In the next section, we consider a class of parameters for which the dimension can be reduced to univariate, and we give explicit formulas for the EBPs of special parameters of interest.

4 Empirical best predictors of additive parameters

Now consider that the parameters are additive in the population units for the area of interest at the given time instant, that is, they take the form

$$\delta_{it} = \frac{1}{N_{it}} \sum_{j=1}^{N_{it}} h(y_{itj}),$$

for a known measurable function h . For this type of parameters, the best predictor is obtained by preserving the values of the (transformed) welfare for the sample units and predicting only for the nonsampled ones, that is,

$$\hat{\delta}_{it}^{bp} = E_{\mathbf{y}_r} \left[\frac{1}{N_{it}} \sum_{j=1}^{N_{it}} h(y_{itj}) \middle| \mathbf{y}_s; \boldsymbol{\theta} \right] = \frac{1}{N_{it}} \left\{ \sum_{j \in s_{it}} h(y_{itj}) + \sum_{j \in r_{it}} E_{\mathbf{y}_r} [h(y_{itj}) | \mathbf{y}_s; \boldsymbol{\theta}] \right\}. \quad (21)$$

Particular additive parameters of interest are the family of FGT poverty indicators introduced by Foster et al. (1984). These indicators are defined in terms of a welfare variable such as income or expenditure. Let z_{itj} be the considered welfare variable for individual j in domain i at time instant t and let z be the poverty line. For a given power $\alpha \geq 0$, the FGT poverty indicator of order α for domain i at time instant t is defined as

$$F_{\alpha,it} = \frac{1}{N_{it}} \sum_{j=1}^{N_{it}} F_{\alpha,itj}, \quad F_{\alpha,itj} = \left(\frac{z - z_{itj}}{z} \right)^\alpha I(z_{itj} < z), \quad (22)$$

where $I(z_{itj} < z) = 1$ if $z_{itj} < z$ and $I(z_{itj} < z) = 0$ otherwise. For $\alpha = 0$, we obtain the proportion of people in domain i at time t whose welfare measure is below the poverty line z , also called poverty incidence. For $\alpha = 1$, we obtain the poverty gap, measuring the poverty intensity of the people in domain i at time t .

We assume that a one-to-one transformation of the welfare variables, $y_{itj} = T(z_{itj})$, follow the AR(1)-NER model (1). Then, using the inverse transformation $z_{itj} = T^{-1}(y_{itj})$, we can

express $F_{\alpha,itj}$ in (22) in terms of the model response variables y_{itj} , as

$$F_{\alpha,itj} = \left(\frac{z - T^{-1}(y_{itj})}{z} \right)^\alpha I(T^{-1}(y_{itj}) < z) \triangleq h_\alpha(y_{itj}), \quad (23)$$

which shows that $F_{\alpha,it}$ is indeed additive. Therefore, the best predictor of $F_{\alpha,it}$ is given by

$$\hat{F}_{\alpha,it}^{bp} = \frac{1}{N_{it}} \left\{ \sum_{j \in s_{it}} F_{\alpha,itj} + \sum_{j \in r_{it}} \hat{F}_{\alpha,itj}^{bp} \right\}, \quad (24)$$

where $\hat{F}_{\alpha,itj}^{bp} \triangleq E_{\mathbf{y}_r}[F_{\alpha,itj}|\mathbf{y}_s; \boldsymbol{\theta}]$. For $\alpha = 0, 1$ and for certain transformations T , the expectation $\hat{F}_{\alpha,itj}^{bp}$ can be calculated analytically. Concretely, for the poverty incidence ($\alpha = 0$), if T is a nondecreasing monotonous function, we obtain

$$\hat{F}_{\alpha,itj}^{bp} = E_{\mathbf{y}_r}[F_{0,itj}|\mathbf{y}_s; \boldsymbol{\theta}] = E_{\mathbf{y}_r}[I(y_{itj} < T(z))|\mathbf{y}_s; \boldsymbol{\theta}] = \Phi(\alpha_{itj}),$$

where Φ is the cumulative distribution function of a standard normal random variable and

$$\alpha_{itj} = \frac{T(z) - \mu_{itj|s}}{v_{itj|s}},$$

for $\mu_{itj|s}$ and $v_{itj|s}^2$ given in Section 3.

For the poverty gap ($\alpha = 1$), considering the log transformation after a shift, $y_{itj} = \log(z_{itj} + c)$, with inverse $z_{itj} = T^{-1}(y_{itj}) = \exp(y_{itj}) - c$, we obtain

$$\begin{aligned} \hat{F}_{1,itj}^{bp} &= E_{\mathbf{y}_r}[F_{1,itj}|\mathbf{y}_s; \boldsymbol{\theta}] = E_{\mathbf{y}_r} \left[\frac{z - T^{-1}(y_{itj})}{z} I(y_{itj} < T(z)) \middle| \mathbf{y}_s; \boldsymbol{\theta} \right] \\ &= \frac{z + c}{z} \Phi(\alpha_{itj}) - \frac{1}{z} \exp \left(\mu_{itj|s} + v_{itj|s}^2/2 \right) \Phi(\alpha_{itj} - v_{itj|s}). \end{aligned}$$

Finally, consider that we wish to estimate the average welfare of domain i at time t under the latter transformation of z_{itj} . Then, we can express the average welfare as

$$\bar{z}_{it} = \frac{1}{N_{it}} \sum_{j=1}^{N_{it}} z_{itj} = \frac{1}{N_{it}} \sum_{j=1}^{N_{it}} \exp(y_{itj}) - c.$$

which is also an additive parameter. Then, the best predictor of $\bar{z}_{it}$ becomes

$$\hat{\bar{z}}_{it}^{bp} = \frac{1}{N_{it}} \left(\sum_{j \in s_{it}} z_{itj} + \sum_{j \in r_{it}} \hat{z}_{itj}^{bp} \right),$$

where, for nonsampled unit $j \in r_{it}$, we have

$$\hat{z}_{itj}^{bp} \triangleq E_{\mathbf{y}_r}[z_{itj}|\mathbf{y}_s; \boldsymbol{\theta}] = E_{\mathbf{y}_r}[\exp(y_{itj})|\mathbf{y}_s; \boldsymbol{\theta}] - c = \exp \left(\mu_{itj|s} + v_{itj|s}^2/2 \right) - c.$$

For more complex functions $h(\cdot)$, the expected value $E_{\mathbf{y}_r}[h(y_{itj})|\mathbf{y}_s; \hat{\boldsymbol{\theta}}]$ involved in the EBP

of an additive parameter in (21) might be not tractable analytically. When this is the case, that expected value can be approximated using the following Monte Carlo simulation procedure:

- a) Estimate $\boldsymbol{\theta} = (\boldsymbol{\beta}', \sigma_v^2, \sigma_u^2, \rho, \sigma_e^2)'$ using the sample data $\mathbf{y}_s$.
- b) Draw L copies of the nonsample response variable y_{itj} , $j \in r_{it}$, from its conditional distribution given the sample data $\mathbf{y}_s$, as

$$y_{itj}^{(\ell)} \sim N(\mu_{itj|s}, v_{itj|s}^2), \quad \ell = 1, \dots, L,$$

with $\boldsymbol{\theta} = (\boldsymbol{\beta}', \sigma_v^2, \sigma_u^2, \rho, \sigma_e^2)'$ replaced by the estimate $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\beta}}', \hat{\sigma}_v^2, \hat{\sigma}_u^2, \hat{\rho}, \hat{\sigma}_e^2)'$ obtained in a).

- c) The Monte Carlo approximation of the expected value on the right-hand side of (21), with estimated parameters, is then obtained by taking average across the L replicates,

$$E_{\mathbf{y}_r}[h(y_{itj})|\mathbf{y}_s; \hat{\boldsymbol{\theta}}] \approx \frac{1}{L} \sum_{\ell=1}^L h(y_{itj}^{(\ell)}), \quad j \in r_{it}.$$

Except for the case of linear parameters δ_{it} (where $h(\cdot)$ is linear), apart from the corresponding survey sample data $(y_{itj}, \mathbf{x}_{itj})$, $j \in s_{it}$, the EBP $\hat{\delta}_{it}^{ebp}$ requires the values of the auxiliary variables $\mathbf{x}_{itj}$ for each nonsample unit $j \in r_{it}$ (microdata) for all areas i and periods t where estimation is desired. This is typically obtained from census or administrative records. In fact, the EBP requires to link the survey and census data files, in order to identify the sample and nonsample units in the census. Many times, linking these two data sets is not feasible. In fact, sometimes the survey sample is not obtained from the considered census and, in that case, the available sample data $\mathbf{y}_s$ is not actually a subset of the population vector $\mathbf{y}$. In that case, instead of the original BP of δ_{it} given in (21), we can use an approximation of it, called the Census BP (CBP). The CBP of δ_{it} is obtained by predicting all the census units based on $\mathbf{y}_s$, as

$$\hat{\delta}_{it}^{cbp} = \frac{1}{N_{it}} \sum_{j=1}^{N_{it}} E_{\mathbf{y}}[h(y_{itj})|\mathbf{y}_s; \boldsymbol{\theta}], \quad (25)$$

where the distribution of $\mathbf{y}|\mathbf{y}_s$ is the same as that of $\mathbf{y}_r|\mathbf{y}_s$ given in Section 3, with the nonsample vectors and matrices replaced by the analogous population versions. The Census EBP (CEBP) $\hat{\delta}_{it}^{cebp}$ is then obtained by replacing the unknown model parameter $\boldsymbol{\theta}$ with the estimate $\hat{\boldsymbol{\theta}}$. When the sampling fractions n_{it}/N_{it} are negligible, which is typically the case in small area applications, the CEBP is very similar to the original EBP.

5 Parametric bootstrap MSE estimator

Here we propose a bootstrap procedure similar to the one described in Molina & Rao (2010), for estimation of the MSE of the EBP $\hat{\delta}_{it}^{ebp}$ under the assumed model, $\text{MSE}(\hat{\delta}_{it}^{ebp}) = E(\hat{\delta}_{it}^{ebp} - \delta_{it})^2$, given the original sample s . This procedure is based on the parametric bootstrap method for

finite populations introduced by González-Manteiga et al. (2008). In this procedure, the MSE of $\hat{\delta}_{it}^{ebp}$ is estimated by keeping the original sample s as fixed and then proceeding as follows:

- a) Fitting the sample model (5) to the survey sample data $(\mathbf{y}_s, \mathbf{X}_s)$, obtain an estimate $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\beta}}', \hat{\sigma}_v^2, \hat{\sigma}_u^2, \hat{\rho}, \hat{\sigma}_e^2)$ of $\boldsymbol{\theta} = (\boldsymbol{\beta}', \sigma_v^2, \sigma_u^2, \rho, \sigma_e^2)'$.
- b) For $b = 1, \dots, B$, with B large, generate $v_i^{*(b)} \stackrel{iid}{\sim} N(0, \hat{\sigma}_v^2)$, $\mathbf{u}_i^{*(b)} \stackrel{iid}{\sim} N(\mathbf{0}_T, \hat{\sigma}_u^2 \boldsymbol{\Omega}_u(\hat{\rho}))$ and $e_{itj}^{*(b)} \stackrel{iid}{\sim} N(0, \hat{\sigma}_e^2)$, $j = 1, \dots, N_{it}$, $i = 1, \dots, m$, $t = 1, \dots, T$, independently.
- c) Construct B i.i.d. bootstrap population vectors $\mathbf{y}^{*(b)}$, $b = 1, \dots, B$, with elements $y_{itj}^{*(b)}$ generated as

$$y_{itj}^{*(b)} = \mathbf{x}_{itj}' \hat{\boldsymbol{\beta}} + v_i^{*(b)} + u_{it}^{*(b)} + e_{itj}^{*(b)}, \quad j = 1, \dots, N_{it}, \quad i = 1, \dots, m, \quad t = 1, \dots, T.$$

From each bootstrap population b , calculate the true value of the domain parameter at time instant t ,

$$\delta_{it}^{*(b)} = N_{it}^{-1} \sum_{j=1}^{N_{it}} h(y_{itj}^{*(b)}), \quad b = 1, \dots, B.$$

- d) From each generated bootstrap population $\mathbf{y}^{*(b)}$, take the sample part $\mathbf{y}_s^{*(b)}$, where s has the same elements as the initial sample. Using the bootstrap sample data $\mathbf{y}_s^{*(b)}$ and the known population vectors $\mathbf{x}_{itj}$, $j = 1, \dots, N_{it}$, calculate the bootstrap EBP of $\delta_{it}^{*(b)}$, denoted $\hat{\delta}_{it}^{ebp*(b)}$, $b = 1, \dots, B$.
- e) A bootstrap estimator of $\text{MSE}(\hat{\delta}_{it}^{ebp})$ is then given by

$$\text{mse}(\hat{\delta}_{it}^{ebp}) = \frac{1}{B} \sum_{b=1}^B \left(\hat{\delta}_{it}^{ebp*(b)} - \delta_{it}^{*(b)} \right)^2. \quad (26)$$

6 Simulation Experiments

6.1 General description

This section presents two simulation experiments with different purposes: In the first one, we use the considered temporal unit-level model to estimate only in the last time period, and we wish to study the cross-sectional performance of the resulting empirical best predictor (called EBT) across areas in comparison with direct estimators and EB estimators based solely on the cross-sectional data for that time point. In the second experiment, we estimate at all the considered time points and investigate the temporal stability of the estimates obtained from the considered model. Our target parameters are the FGT poverty indicators $F_{\alpha, it}$, of orders $\alpha = 0, 1$ for each area $i = 1, \dots, m$. Both simulation experiments are based on real data, the Spanish EU-SILC data for women in years 2004, 2005 and 2006 used in the application of Section 7, which yield realistic simulation scenarios.

The EB predictors require microdata from a census of the auxiliary variables, where the sample units can be identified. Since this is not always available, in our comparisons we consider

two additional predictors: the Census EB predictors $\hat{F}_{\alpha, it}^{cebt}$ based on the temporal unit-level model (called CEBT), which do not require to identify the sample units in the census file and another version, called estimated CEBT (CEBT_e) and denoted $\hat{F}_{\alpha, iT}^{cebte}$, which can be employed when no census file is available. These estimators replace the census with an expansion of the survey auxiliary data based on the survey weights.

6.2 Simulation 1: cross-sectional prediction

In this simulation experiment, we analyze the cross-sectional properties of the proposed predictors when estimating at the last time point T , using the survey data from the three time instants. The simulation experiment is purely model-based; hence, the sample s is kept fixed across simulations, and new model responses y_{itj} are generated from the assumed AR(1)-NER model (1) in each simulation replicate. As in the application of Section 7, the target parameters are the province poverty incidences and gaps for women in the last year $t = T$.

The Spanish EU-SILC data for the auxiliary variables acts as the census of these variables and the survey values are obtained drawing a sample from that “census”. Hence, the actual EU-SILC sample size for each province i at each year t becomes the population size N_{it} , and a subsample of size $n_{it} = N_{it}/4$ is drawn from each province by simple random sampling without replacement. The sample is kept fixed across the simulations. We exclude the province of Soria because of its small sample size, so we wish to estimate in the remaining $m = 51$ provinces.

Note that when predicting only for the last year $t = T$, census data are needed only for that year. The data from the previous years $t = 1, \dots, T - 1$ is used only to fit the AR(1)-NER model, so only the survey data are required for those years. The same auxiliary variables as in the application of Section 7 are considered. As true values of the model parameters $\theta = (\beta', \sigma_v^2, \sigma_u^2, \rho, \sigma_e^2)'$, we take the resulting estimates after fitting the AR(1)-NER model (5) to the log-income data from the Spanish SILC. These values are reported in Tables 2 and 3.

We run $L = 1000$ iterations, $\ell = 1, \dots, L$. In each iteration ℓ , we generate the sample values $\mathbf{y}_{its}^{(\ell)}$ for all but last year $t = 1, \dots, T - 1$ from the sample model (5), with design matrix $\mathbf{X}_{its}$ obtained from the subsample of the EU-SILC, and generating the area effects as $v_i^{(\ell)} \stackrel{iid}{\sim} N(0, \sigma_v^2)$, the time effects within each area as $\mathbf{u}_i^{(\ell)} = (u_{i1}^{(\ell)}, \dots, u_{iT}^{(\ell)})' \stackrel{iid}{\sim} N(\mathbf{0}_T, \sigma_u^2 \mathbf{\Omega}_u)$, and the model errors as $e_{itj}^{(\ell)} \stackrel{iid}{\sim} N(0, \sigma_e^2)$, $i = 1, \dots, m$, $t = 1, \dots, T - 1$. For the last year T , the population vectors for each area i , $\mathbf{y}_{iT}^{(\ell)} = (\mathbf{y}_{iT_s}^{(\ell)}, \mathbf{y}_{iT_r}^{(\ell)})'$, are generated using the census matrices $\mathbf{X}_{iT}$ obtained from the full SILC data for province i , the previously generated $v_i^{(\ell)}$ and $\mathbf{u}_{iT}^{(\ell)}$, and the individual errors $e_{iTj}^{(\ell)} \stackrel{iid}{\sim} N(0, \sigma_e^2)$, $i = 1, \dots, m$. Note that the nonsample response vector for the last year, $\mathbf{y}_{iT_r}^{(\ell)}$, uses the nonsample matrix $\mathbf{X}_{iT_r}$ obtained from the $3N_{iT}/4$ individuals that are not in the subsample of the EU-SILC for province i and the same values of the model parameters θ used to generate the sample part $\mathbf{y}_{iT_s}$.

For each Monte Carlo simulation ℓ , the poverty line is calculated as in the EU-SILC, $z^{(\ell)} = 0.6 \text{median}(z_{iTj}^{(\ell)})$, where $z_{iTj} = \exp(y_{iTj}^{(\ell)})$ represents the annual disposable income of individual j from domain i at period T . With the population data for the last year $t = T$, $\mathbf{y}_{iT}^{(\ell)}$, we compute the true values of the poverty indicators $F_{\alpha, iT}^{(\ell)}$, for $\alpha = 0, 1$ and $i = 1, \dots, m$. With the

corresponding sample values of the last year $t = T$, $\mathbf{y}_{iT}^{(\ell)}$, we compute the direct estimators $\hat{F}_{\alpha, iT}^{dir(\ell)}$ for each area $i = 1, \dots, m$. The same data is also used to fit a NER model, obtaining estimates of the NER model parameters and calculating the cross-sectional EBPs $\hat{F}_{\alpha, iT}^{eb(\ell)}$, $i = 1, \dots, m$. On the other hand, the full sample vector $\mathbf{y}_s^{(\ell)}$ is used to fit the AR(1)-NER model, yielding model parameter estimates $\hat{\boldsymbol{\theta}}^{(\ell)} = ((\hat{\boldsymbol{\beta}}^{(\ell)})', \hat{\sigma}_v^{2(\ell)}, \hat{\sigma}_u^{2(\ell)}, \hat{\rho}^{(\ell)}, \hat{\sigma}_e^{2(\ell)})'$. Then the EBT, CEBT and CEBTe estimators are calculated. We denote them respectively by $\hat{F}_{\alpha, iT}^{ebt(\ell)}$, $\hat{F}_{\alpha, iT}^{cebt(\ell)}$ and $\hat{F}_{\alpha, iT}^{cebte(\ell)}$. Finally, we compute the empirical MSE of each estimator $\hat{F}_{\alpha, iT}$ across the L Monte Carlo simulations as

$$\text{MSE}(\hat{F}_{\alpha, iT}) = \frac{1}{L} \sum_{\ell=1}^L (\hat{F}_{\alpha, iT}^{(\ell)} - F_{\alpha, iT}^{(\ell)})^2.$$

We calculate also the relative bias (RB) and relative root of mean square error (RRMSE) of each estimator $\hat{F}_{\alpha, iT}$ as

$$\text{RB}(\hat{F}_{\alpha, iT}) = \frac{\frac{1}{L} \sum_{\ell=1}^L (\hat{F}_{\alpha, iT}^{(\ell)} - F_{\alpha, iT}^{(\ell)})}{\frac{1}{L} \sum_{\ell=1}^L F_{\alpha, iT}^{(\ell)}}, \quad \text{RRMSE}(\hat{F}_{\alpha, iT}) = \frac{\sqrt{\text{MSE}[\hat{F}_{\alpha, iT}]}}{\frac{1}{L} \sum_{\ell=1}^L F_{\alpha, iT}^{(\ell)}}.$$

To summarize the results, we divided the $m = 51$ provinces in four groups according to the sample sizes n_{iT} (<35 , $[35,60)$, $[60,125)$, >125), with thirteen provinces in each group, except for the last one which contains twelve provinces. Figure 1 shows boxplots of percent RB (left) and RRMSE (right) for the direct (DIR) estimator and the EB, EBT, CEBT and CEBTe predictors of the poverty gap $F_{1, iT}$ for each group of provinces. These boxplots show a very similar performance of EB and EBT predictors in terms of RB and RRMSE. This occurs because the original SILC data has very weak time effects (see Table 2, showing a pretty small estimate of σ_u^2). This means that, when the time effects are weak, not much is lost when considering them in the model. Comparing now the other predictors, note that CEBT performs also very similarly to EBT and CEBTe. CEBTe seems to pursue the direct estimators in terms of RB, but it has clearly lower RRMSE. Results for poverty incidence are similar and then not reported here.

6.3 Simulation 2: longitudinal stability

The second simulation experiment is designed to study the temporal stability of the considered model-based predictors (EB, EBT, CEBT and CEBTe). We exclude the direct estimator from this study as we know its lower performance in terms of variability. This experiment is similar to the previous one and based on the same data set. The only difference is that we calculate the estimators for each province $i = 1, \dots, m$ and for each year $t = 1, \dots, T$ and we calculate a measure of temporal stability. We expect the methods accounting for the time effect, namely EBT, CEBT and CEBTe, provide smoother longitudinal predictions (measured by variability over time) than the EB predictors based on separate cross-sectional NER models for each time instant. As measure of temporal stability of each estimator for each area, we compute the standard deviation (sd) of the predicted poverty indicators over the T time points. To make them comparable, we divide them by the mean value of the province EBT predicted values,

Figure 1: Boxplots of percent RB (left) and RRMSE (right) of DIR, EB, EBT, CEBT and CEBTe estimators of poverty gap, F_{1iT} , for each province. Provinces are grouped in four sets of sample sizes and boxplots are shown in order from group 1 with smaller sample size to group 4 with the largest.

yielding a coefficient of variation (cv) over time for each predictor and province.

Figure 2 shows boxplots of the cv's of the model-based predictors over time, for each group of provinces as in simulation 1. Figure 2 reveals practically identical boxplots for EBT and CEBT predictors, which are both much more stable over time than the EB predictor. Except for two particular areas of group 2 of sample sizes, the CEBTe predictor is also more stable than the EB predictor based on the cross-sectional model for the last time instant.

Figure 2: Boxplots of cv over time of EB, EBT, CEBT and CEBTe predictors of poverty gap, F_{1iT} . Provinces are grouped in four sets of sample sizes and boxplots are shown from group 1 with smaller sample size to group 4 with the largest.

The main conclusions of these simulation experiments are: 1) When estimating in a given time point, there is practically no loss in terms of bias and MSE due to considering the time effects even if they are weak. 2) When estimating at several time instants, the predictors based on the temporal model (EBT, CEBT and CEBTe) are much more stable than the EB predictor

based on the cross-sectional model for each time point. 3) The CEBT performs practically the same as the original EBT when domain sampling fractions are negligible. 4) The CEBTe predictors are still better than the direct estimators when the census auxiliary data required for the EBT predictor are not available.

7 Application to poverty mapping in Spain

This section estimates the 2006 poverty incidences and gaps for the women across the $m = 52$ Spanish provinces, using the data from the Spanish EU-SILC for years 2004, 2005 and 2006. The EU-SILC collects data on income, poverty, social exclusion, labor, education and health in a timely and comparable way across the European Union. Results from this survey are used to obtain the structural index of social cohesion.

We consider the direct estimator (DIR) of Hájek (1971), the EB predictor based on the NER model for 2006, and EBT predictors based on the AR(1)-NER model for the three years. As census data are not available, we consider the largest Spanish survey, which is the Spanish Labor Force Survey (LFS) of 2006, expanded with its corresponding survey weights. Moreover, we consider the Census versions of the EB and EBT predictors. We have seen in the simulation experiments of Section 6 that the CEBT predictor performs almost the same as the original EBT predictors. Similarly, Correa et al. (2012) and Guadarrama et al. (2016) showed that the CEB is approximately equal to the EB predictors when the area sampling fractions are negligible. Note that the LFS is a much larger survey than the EU-SILC, so the estimated versions CEBTe and CEBe based on this larger survey should perform much better than those considered in the simulation experiments, which were based on the actual sample data. Note also that the actual sample size of the EU-SILC is 4 times that one considered in the simulation experiments.

The Spanish EU-SILC provides reliable direct estimates at the national level and also at the NUTS 2 level (Autonomous Communities), but for NUTS 3 (provinces), direct estimators are no longer reliable due to the small province sample sizes. Table 1 shows the quartiles and the total EU-SILC samples sizes for the Spanish provinces in 2004, 2005, 2006 and for the whole period 2004-2006, for women. This table shows that 25% of the provinces have sample size lower than 138 for the year 2006, and there is a province with only 17 women sampled.

Table 1: Summary of EU-SILC province sample sizes for women.

	2004	2005	2006	2004-2006
Min	17	13	17	47
Q1	180	161	138	489
Q2	291	270	239	800
Q3	621	527	491	1,613
Max	1,822	1,494	1,483	4,718
Total	21,597	19,119	17,739	58,455

As auxiliary variables, we consider three levels of education (primary, secondary and superior studies), nationality (being Spanish or not) and three levels of labor status (employed, unemployed and inactive); the category below 16 years old is the reference category for the level

of education and the labor status. The welfare measure z_{itj} in the EU-SILC is the equivalized disposable income. This income is the total income of a household, after tax and other deductions available for spending or saving, divided by a measure of household size calculated according to a scale defined by the OECD. The poverty line z is set to 60% of the national median of equivalized annual disposable income. We address the right skewness and overdispersion of z_{itj} by taking the transformation $y_{itj} = \log(z_{itj} + c)$, where c is selected so that the residuals obtained from the fitted model, $\hat{e}_{itj} = y_{itj} - \mathbf{x}'_{itj}\hat{\beta} - \hat{v}_i - \hat{u}_{it}$, are approximately symmetric.

Figure 3 depicts the scatter plot of the residuals $\hat{e}_{itj}$ of the fitted AR(1)-NER model against the corresponding predicted values $\hat{y}_{itj}$ (left) and the histogram of the model residuals (right). The left plot reveals no remarkable patterns for the residuals except for one outlier at the bottom. The histogram on the right shows a big concentration of residuals around 0, indicating the high predictive power of the considered model.

Figure 3: Temporal model residuals versus predicted values (left) and histogram of residuals (right)

Table 2 gives the estimates of the variance components of the fitted AR(1)-NER model (5). We include 95% confidence intervals of each parameter. Table 3 reports the regression coefficients. The estimated coefficients indicate that the Spanish individuals (natio=1) with higher education (edu3=1) and employment (situ1=1) tend to have larger incomes. On the other hand, unemployed individuals (situ2=1) with lower education level (edu1=1 or edu2=1) tend to have smaller incomes.

Table 2: Estimated variance components of the AR(1)-NER model

parameter	estimate	lower.lim	upper.lim
σ_e^2	0.1313	0.1298	0.1328
σ_v^2	0.0065	-0.0226	0.0355
σ_u^2	0.0015	-0.0031	0.0061
ρ	0.7316	0.3949	1.0683

Figure 4 plots DIR, CEBe and CEBTe estimates of women poverty incidence (left) and poverty gap (right) for in each province, where provinces are sorted by increasing sample sizes. We can see clearly that direct estimates are much more volatile than model-based ones, specially

Table 3: Estimated regression coefficients of the AR(1)-NER model

variable	estimate	std.error	z-Value	p-value
intercept	9.4002	0.0429	219.0955	0.0000
edu1	-0.1366	0.0120	-11.4006	0.0000
edu2	-0.0488	0.0128	-3.8154	0.0001
edu3	0.2073	0.0173	11.9899	0.0000
natio	0.2067	0.0217	9.5278	0.0000
situ1	0.1501	0.0110	13.5901	0.0000
situ2	-0.0913	0.0203	-4.4954	0.0000

for the provinces with smaller sample size. Although CEBTe and CEBe estimates are rather close, CEBTe based on the temporal model seem to be slightly more stable. Similar patterns are obtained for the poverty gap.

Figure 4: DIR, CEBe and CEBTe estimates of poverty incidences (left) and gaps (right) for women in each province. Provinces sorted by increasing sample sizes.

MSEs of the direct estimator and the CEBe and CEBTe predictors are estimated using the parametric bootstrap procedure of Section 5. The MSEs for the three different estimators are comparable here, because they are all calculated with respect to the distribution of the fitted AR(1)-NER model. Figure 5 displays boxplots of the estimated CVs of the three estimators for the poverty incidence (left) and poverty gap (right), where these CVs are in fact the squared root of the estimated MSEs by the estimates. These boxplots show the much larger estimated CVs for the direct estimators, which exceed 40% for the poverty incidence and 60% for the poverty gap. Both model-based estimators show a clear improvement in CV with respect to the direct ones, but even if the time effects are weak in this application, the CEBTe predictors with time effects seem to be slightly better than CEBe without time effects in terms of estimated CV. This improvement can also be seen in Figure 6, showing the estimated the MSEs of CEBTe (labeled CEBTe-MSE) against those of CEBe (labeled CEBe-MSE) for the poverty incidences (left) and gaps (right) for women in the Spanish provinces. This figure shows that the largest estimated MSEs, which are for the provinces with smaller sample sizes, are slightly smaller for CEBTe based on the model with time effects compared to those of CEBe based on the cross-sectional

NER model without time effects. For the provinces with larger sample sizes, the MSE estimates of the two model-based predictors are very similar.

Figure 5: Boxplots of estimated CVs for poverty incidences (left) and gaps (right) for women.

Figure 6: Estimated MSEs of CEBTe versus those of CEBe for poverty incidences (left) and gaps (right) for women in Spanish provinces.

In practice, National statistical offices are interested in estimators whose estimated CV do not exceed certain limits. In Table 4 we list the number of provinces in which the estimated CVs of each estimator are below 10%, in between 10% and 15%, in between 15% and 20%, and over 20%. See that both CEB and CEBT get estimated CVs larger than 20% for less provinces than the direct estimators, but CEBT has CVs in the lowest range for more provinces than CEB in the case of poverty incidence.

Finally, cartograms of CEBTe predicted poverty incidences and gaps for women in 2006 are shown in Figure 7. This figure shows that poverty is concentrated in the south-west of Spain, as well as in the provinces of Soria, Zamora and Ávila in the Autonomous Community of Castilla-León. The provinces with the largest estimated poverty incidence for women are Ávila, Albacete, Almería, Badajoz, Cádiz, Ciudad Real, Córdoba, Cuenca, Jaén and Zamora, whereas

Table 4: Number of provinces with estimated percent CVs below 10, 10-15, 15-20 and over 20, for poverty incidences and gaps for women.

estimator	poverty incidence				poverty gap			
	<10	10-15	15-20	>20	<10	10-15	15-20	>20
DIR	6	21	10	15	2	15	14	21
CEB	21	19	11	1	5	16	13	18
CEBT	27	18	6	1	5	20	9	18

the largest estimated poverty gap for women is found in Granada.

Figure 7: Cartogram of CEBTe predicted poverty incidences (left) and gaps (right) for women in 2006.

Finally, Appendix B reports the detailed results for each province. Concretely, Table 5 shows direct, CEB and CEBT estimates of poverty incidence and their estimated percent CVs for women, with provinces sorted from smaller to larger sample size. The estimated CVs of the direct estimators exceed 20% for the provinces with the smaller sample sizes such as Soria, Segovia, Ávila and Palencia. On the other hand, for CEBT, the estimated CVs are below 20% for all the provinces except for Álava in the case of poverty incidence, although for this province the estimated CV is still smaller than those of the other estimators. Table 6 includes the analogous results for the poverty gap. In this case, the estimated CVs of the direct estimator exceed 20% for all the provinces with less than 140 sampled individuals and a few more. For Soria, the CV of the direct estimator exceeds 40% for the poverty incidence and 60% for the poverty gap, which makes it completely unreliable. CEBT achieve smaller estimated CVs than CEB for the two poverty indicators in the provinces with smaller sample sizes, and take similar values for the provinces with larger sample sizes.

8 Conclusions

This paper fills a gap in the small area estimation literature by introducing EBPs of general non-linear parameters based on a unit-level temporal linear mixed model that contemplates auto-correlation structure of type AR(1). The paper gives spelled-out formulas for the EBPs of poverty incidences, poverty gaps and average incomes under the considered model. Further, it introduces a parametric bootstrap approach to estimate the MSEs of the EBPs.

In two simulation experiments based on real data from the Spanish EU-SILC, we study the cross-sectional and longitudinal properties of the EBPs based on the proposed model. The following conclusions are obtained from these simulation studies: If we are only interested in estimating for a fixed time period T , then predictors based on the AR(1)-NER model (2), which uses historical information, work similarly, in terms of bias and MSE, than those based on the NER model fitted to period T , in the case of weak time effects. On the other hand, when estimating over time, temporal predictors (EBT, CEBT and CEBTe) are much more stable than the cross-sectional EB predictor based on a separate model for each instant. These results suggest that the use of temporal model-based predictors is advisable when doing longitudinal studies or producing periodical statistics. Furthermore, our studies also support the use of CEBT when linking the census and survey data sets is not possible, or the CEBTe when census microdata on the auxiliary variables is not available.

The paper estimates poverty incidences and gaps for women in the Spanish provinces for 2006, using the data from the Spanish EU-SILC of years 2004, 2005 and 2006. This application illustrates the advantage of improving estimates for a given year using data from previous years. From a socioeconomic point of view, the obtained poverty maps show that the largest estimated rates of women suffering from poverty and also the largest values of estimated poverty intensity are for provinces in the south-west of Spain and also in some provinces of central Spain.

Acknowledgements

This work was supported by the Spanish grants PGC2018-096840-B-I00, MTM2015-69638-R and MTM2015-64842-P from Ministerio de Economía y Competitividad.

Appendix A: Calculations

A.1. Inverse of sample covariance matrix

Let us write $\mathbf{V}_{is} = \sigma_v^2 \mathbf{1}_{n_i} \mathbf{1}_{n_i}' + \mathbf{R}_{is}$, for $\mathbf{R}_{is} = \sigma_u^2 \mathbf{A}_{is} \boldsymbol{\Omega}_u \mathbf{A}_{is}' + \sigma_e^2 \mathbf{W}_{is}^{-1}$. The inverse is then $\mathbf{V}_{is}^{-1} = (\sigma_v^2 \mathbf{1}_{n_i} \mathbf{1}_{n_i}' + \mathbf{R}_{is})^{-1}$. Applying the inversion formula

$$(\mathbf{A} + \mathbf{u}\mathbf{v}')^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{u}\mathbf{v}'\mathbf{A}^{-1}}{1 + \mathbf{v}'\mathbf{A}^{-1}\mathbf{u}},$$

with $\mathbf{A} = \mathbf{R}_{is}$, $\mathbf{u} = \sigma_v^2 \mathbf{1}_{n_i}$ and $\mathbf{v} = \mathbf{1}_{n_i}$, we obtain

$$\mathbf{V}_{is}^{-1} = \mathbf{R}_{is}^{-1} - \frac{\sigma_v^2}{1 + \sigma_v^2 \mathbf{1}_{n_i}' \mathbf{R}_{is}^{-1} \mathbf{1}_{n_i}} \mathbf{R}_{is}^{-1} \mathbf{1}_{n_i} \mathbf{1}_{n_i}' \mathbf{R}_{is}^{-1}. \quad (27)$$

Now to calculate $\mathbf{R}_{is}^{-1}$, we apply the formula

$$(A + CBD)^{-1} = A^{-1} - A^{-1}C(B^{-1} + DA^{-1}C)^{-1}DA^{-1},$$

with $A = \sigma_e^2 \mathbf{W}_{is}^{-1}$, $B = \mathbf{\Omega}_u$, $C = \sigma_u^2 \mathbf{A}_{is}$ and $D = \mathbf{A}_{is}'$, obtaining

$$\begin{aligned} \mathbf{R}_{is}^{-1} &= \frac{1}{\sigma_e^2} \mathbf{W}_{is} - \frac{\sigma_u^2}{\sigma_e^4} \mathbf{W}_{is} \mathbf{A}_{is} \left(\mathbf{\Omega}_u^{-1} + \frac{\sigma_u^2}{\sigma_e^2} \mathbf{A}_{is}' \mathbf{W}_{is} \mathbf{A}_{is} \right)^{-1} \mathbf{A}_{is}' \mathbf{W}_{is} \\ &= \frac{1}{\sigma_e^2} \left[\mathbf{W}_{is} - \frac{\sigma_u^2}{\sigma_e^2} \mathbf{B}_{is} \left(\mathbf{\Omega}_u^{-1} + \frac{\sigma_u^2}{\sigma_e^2} \mathbf{G}_{is} \right)^{-1} \mathbf{B}_{is}' \right] \\ &= \frac{1}{\sigma_e^2} \mathbf{W}_{is}^c. \end{aligned} \quad (28)$$

The inverse of $\mathbf{\Omega}_u$ involved in $\mathbf{\Delta}_i$ is given by

$$\mathbf{\Omega}_u^{-1} = \begin{pmatrix} 1 & -\rho & 0 & \cdots & 0 & 0 & 0 \\ -\rho & 1 + \rho^2 & -\rho & \cdots & 0 & 0 & 0 \\ 0 & -\rho & 1 + \rho^2 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \ddots & 1 + \rho^2 & -\rho & 0 \\ 0 & 0 & \cdots & \ddots & -\rho & 1 + \rho^2 & -\rho \\ 0 & 0 & \cdots & \cdots & 0 & -\rho & 1 \end{pmatrix}. \quad (29)$$

Then, we have

$$\mathbf{\Delta}_i = \begin{pmatrix} \alpha_{i1} & -\rho & 0 & \cdots & 0 \\ -\rho & \rho^2 + \alpha_{i2} & -\rho & \cdots & 0 \\ 0 & -\rho & \rho^2 + \alpha_{i3} & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & \rho^2 + \alpha_{iT-1} & -\rho \\ 0 & \cdots & \cdots & -\rho & \alpha_{iT} \end{pmatrix},$$

where

$$\alpha_{it} = 1 + \frac{\sigma_u^2}{\sigma_e^2} w_{its}, \quad t = 1, \dots, T.$$

Since $\mathbf{\Delta}_i$ is $T \times T$ and in small area estimation T is typically small, $\mathbf{\Delta}_i^{-1}$ can be computed numerically. Anyway, its inverse is the symmetric matrix given by $\mathbf{\Delta}_i^{-1} = (\tau_{\ell r})_{\ell, r \in \{1, \dots, T\}}$, with

$$\tau_{\ell r} = \sum_{t=r}^T \frac{\rho^{t-r+t-\ell}}{\left(\prod_{j=\ell}^{r-1} \alpha_{ij} \right) \left(\prod_{j=r}^{t-1} \alpha_{ij}^2 \right) \alpha_{it}}, \quad r, \ell = 1, \dots, T.$$

Replacing (28) in (27), we obtain

$$\mathbf{V}_{is}^{-1} = \frac{1}{\sigma_e^2} \left(\mathbf{W}_{is}^c - \frac{\sigma_v^2}{\sigma_e^2 + \sigma_v^2 w_{is}^c} \mathbf{W}_{is}^c \mathbf{1}_{n_i} \mathbf{1}_{n_i}' \mathbf{W}_{is}^c \right) = \frac{1}{\sigma_e^2} \left(\mathbf{W}_{is}^c - \frac{\gamma_{iw}^c}{w_{is}^c} \mathbf{W}_{is}^c \mathbf{1}_{n_i} \mathbf{1}_{n_i}' \mathbf{W}_{is}^c \right), \quad (30)$$

where

$$\begin{aligned} w_{is}^c &= \mathbf{1}_{n_i}' \mathbf{W}_{is}^c \mathbf{1}_{n_i} = \mathbf{1}_{n_i}' \mathbf{W}_{is} \mathbf{1}_{n_i} - \frac{\sigma_u^2}{\sigma_e^2} \mathbf{1}_{n_i}' \mathbf{B}_{is} \Delta_i^{-1} \mathbf{B}_{is}' \mathbf{1}_{n_i} \\ &= w_{is} - \frac{\sigma_u^2}{\sigma_e^2} \mathbf{g}_{is}' \Delta_i^{-1} \mathbf{g}_{is}, \end{aligned} \quad (31)$$

for $\mathbf{g}_{is} = \mathbf{B}_{is}' \mathbf{1}_{n_i} = (w_{i1}, \dots, w_{iT})'$. Replacing now $\mathbf{w}_{is}^c = \mathbf{W}_{is}^c \mathbf{1}_{n_i}$ in (30), we obtain the inverse of $\mathbf{V}_{is}$ given in (15), where

$$\mathbf{w}_{is}^c = \mathbf{W}_{is}^c \mathbf{1}_{n_i} = \left(\mathbf{W}_{is} - \frac{\sigma_u^2}{\sigma_e^2} \mathbf{B}_{is} \Delta_i^{-1} \mathbf{B}_{is}' \right) \mathbf{1}_{n_i} = \mathbf{w}_{is} - \frac{\sigma_u^2}{\sigma_e^2} \mathbf{B}_{is} \Delta_i^{-1} \mathbf{g}_{is}. \quad (32)$$

A.2. Conditional mean vector

In this section we spell out the conditional mean vector $\boldsymbol{\mu}_{r|s}$ given in (13), which involves the covariance matrix between the nonsample and sample vectors $\mathbf{y}_r$ and $\mathbf{y}_s$, $\mathbf{V}_{rs} = \text{cov}(\mathbf{y}_r, \mathbf{y}_s)$. Considering the sample and nonsample models (7) and (11), the covariance matrix becomes

$$\begin{aligned} \mathbf{V}_{rs} &= E[(\mathbf{Z}_{1r} \mathbf{v} + \mathbf{Z}_{2r} \mathbf{u} + \mathbf{W}_r^{-1/2} \mathbf{e}_r)(\mathbf{Z}_{1s} \mathbf{v} + \mathbf{Z}_{2s} \mathbf{u} + \mathbf{W}_s^{-1/2} \mathbf{e}_s)'] \\ &= \sigma_v^2 \mathbf{Z}_{1r} \mathbf{Z}_{1s}' + \sigma_u^2 \mathbf{Z}_{2r} \boldsymbol{\Omega}_u \mathbf{Z}_{2s}' = \text{diag}_{1 \leq i \leq m}(\mathbf{V}_{irs}), \end{aligned}$$

that is, $\mathbf{V}_{rs}$ is block-diagonal, with blocks given by

$$\mathbf{V}_{irs} = \sigma_v^2 \mathbf{1}_{N_i - n_i} \mathbf{1}_{n_i}' + \sigma_u^2 \mathbf{A}_{ir} \boldsymbol{\Omega}_u \mathbf{A}_{is}'. \quad (33)$$

Moreover, we know that $\mathbf{V}_s^{-1} = \text{diag}_{1 \leq i \leq m}(\mathbf{V}_{is}^{-1})$, with blocks $\mathbf{V}_{is}^{-1}$ given in (15). Therefore, we have $\mathbf{V}_{rs} \mathbf{V}_s^{-1} = \text{diag}_{1 \leq i \leq m}(\mathbf{V}_{irs} \mathbf{V}_{is}^{-1})$, where the blocks $\mathbf{V}_{irs} \mathbf{V}_{is}^{-1}$ are now

$$\begin{aligned} \mathbf{V}_{irs} \mathbf{V}_{is}^{-1} &= \underbrace{(\sigma_v^2 \mathbf{1}_{N_i - n_i} \mathbf{1}_{n_i}')}_A + \underbrace{\sigma_u^2 \mathbf{A}_{ir} \boldsymbol{\Omega}_u \mathbf{A}_{is}'}_B \underbrace{\left(\frac{1}{\sigma_e^2} \mathbf{W}_{is}^c - \frac{\gamma_{iw}^c}{\sigma_e^2 w_{is}^c} \mathbf{w}_{is}^c (\mathbf{w}_{is}^c)' \right)}_{C-D} \\ &= (AC - AD) + (BC - BD). \end{aligned} \quad (34)$$

Next calculate each term in (34). Using (32), the first term reduces to

$$AC = \sigma_v^2 \mathbf{1}_{N_i - n_i} \mathbf{1}_{n_i}' \frac{1}{\sigma_e^2} \mathbf{W}_{is}^c = \frac{\sigma_v^2}{\sigma_e^2} \mathbf{1}_{N_i - n_i} (\mathbf{w}_{is}^c)'.$$

Similarly, by (31), we get

$$AD = \sigma_v^2 \mathbf{1}_{N_i-n_i} \mathbf{1}_{n_i}' \frac{1}{\sigma_e^2} \frac{\gamma_{iw}^c}{w_{is}^c} \mathbf{w}_{is}^c (\mathbf{w}_{is}^c)' = \frac{\sigma_v^2}{\sigma_e^2} \gamma_{iw}^c \mathbf{1}_{N_i-n_i} (\mathbf{w}_{is}^c)'.$$

Therefore, we have

$$AC - AD = \frac{\sigma_v^2}{\sigma_e^2} (1 - \gamma_{iw}^c) \mathbf{1}_{N_i-n_i} (\mathbf{w}_{is}^c)'.$$

But note that

$$\frac{\sigma_v^2}{\sigma_e^2} (1 - \gamma_{iw}^c) = \frac{\gamma_{iw}^c}{w_{is}^c}.$$

Then, we can write

$$AC - AD = \frac{\gamma_{iw}^c}{w_{is}^c} \mathbf{1}_{N_i-n_i} (\mathbf{w}_{is}^c)'. \quad (35)$$

Now replacing $\mathbf{W}_{is}^c = \mathbf{W}_{is} - \frac{\sigma_u^2}{\sigma_e^2} \mathbf{B}_{is} \Delta_i^{-1} \mathbf{B}_{is}'$, and using the fact that $\mathbf{A}_{is}' \mathbf{B}_{is} = \mathbf{A}_{is}' \mathbf{W}_{is} \mathbf{A}_{is} = \mathbf{G}_{is}$, we obtain

$$\begin{aligned} BC &= \frac{1}{\sigma_e^2} \sigma_u^2 \mathbf{A}_{ir} \Omega_u \mathbf{A}_{is}' \mathbf{W}_{is}^c = \frac{\sigma_u^2}{\sigma_e^2} \mathbf{A}_{ir} \Omega_u \mathbf{A}_{is}' \left(\mathbf{W}_{is} - \frac{\sigma_u^2}{\sigma_e^2} \mathbf{B}_{is} \Delta_i^{-1} \mathbf{B}_{is}' \right) \\ &= \frac{\sigma_u^2}{\sigma_e^2} \mathbf{A}_{ir} \Omega_u \left(\mathbf{B}_{is}' - \frac{\sigma_u^2}{\sigma_e^2} \mathbf{G}_{is} \Delta_i^{-1} \mathbf{B}_{is}' \right) = \frac{\sigma_u^2}{\sigma_e^2} \mathbf{A}_{ir} \Omega_u (\mathbf{B}_{is}^c)', \end{aligned}$$

for $\mathbf{B}_{is}^c = \mathbf{B}_{is} - \frac{\sigma_u^2}{\sigma_e^2} \mathbf{B}_{is} \Delta_i^{-1} \mathbf{G}_{is}$. Noting now that $\mathbf{A}_{is} \mathbf{w}_{is} = \mathbf{g}_{is}$ and using again the fact that $\mathbf{G}_{is} = \mathbf{A}_{is}' \mathbf{B}_{is}$, we obtain

$$\begin{aligned} BD &= \frac{1}{\sigma_e^2} \sigma_u^2 \mathbf{A}_{ir} \Omega_u \mathbf{A}_{is}' \frac{\gamma_{iw}^c}{w_{is}^c} \mathbf{w}_{is}^c (\mathbf{w}_{is}^c)' = \frac{\sigma_u^2}{\sigma_e^2} \frac{\gamma_{iw}^c}{w_{is}^c} \mathbf{A}_{ir} \Omega_u \mathbf{A}_{is}' \left(\mathbf{w}_{is} - \frac{\sigma_u^2}{\sigma_e^2} \mathbf{B}_{is} \Delta_i^{-1} \mathbf{g}_{is} \right) (\mathbf{w}_{is}^c)' \\ &= \frac{\sigma_u^2}{\sigma_e^2} \frac{\gamma_{iw}^c}{w_{is}^c} \mathbf{A}_{ir} \Omega_u \left(\mathbf{g}_{is} - \frac{\sigma_u^2}{\sigma_e^2} \mathbf{G}_{is} \Delta_i^{-1} \mathbf{g}_{is} \right) (\mathbf{w}_{is}^c)' = \frac{\sigma_u^2}{\sigma_e^2} \frac{\gamma_{iw}^c}{w_{is}^c} \mathbf{A}_{ir} \Omega_u \mathbf{g}_{is}^c (\mathbf{w}_{is}^c)', \end{aligned}$$

for $\mathbf{g}_{is}^c := \mathbf{g}_{is} - \frac{\sigma_u^2}{\sigma_e^2} \mathbf{G}_{is} \Delta_i^{-1} \mathbf{g}_{is}$. Finally,

$$BC - BD = \frac{\sigma_u^2}{\sigma_e^2} \mathbf{A}_{ir} \Omega_u (\mathbf{B}_{is}^c)' - \frac{\sigma_u^2}{\sigma_e^2} \frac{\gamma_{iw}^c}{w_{is}^c} \mathbf{A}_{ir} \Omega_u \mathbf{g}_{is}^c (\mathbf{w}_{is}^c)' = \frac{\sigma_u^2}{\sigma_e^2} \mathbf{A}_{ir} \Omega_u \left(\mathbf{B}_{is}^c - \frac{\gamma_{iw}^c}{w_{is}^c} \mathbf{w}_{is}^c (\mathbf{g}_{is}^c)' \right)'. \quad (36)$$

Replacing (35) and (36) in (34), we get

$$\begin{aligned} \mathbf{V}_{irs} \mathbf{V}_{is}^{-1} &= (AC - AD) + (BC - BD) \\ &= \frac{\gamma_{iw}^c}{w_{is}^c} \mathbf{1}_{N_i-n_i} (\mathbf{w}_{is}^c)' + \frac{\sigma_u^2}{\sigma_e^2} \mathbf{A}_{ir} \Omega_u \left(\mathbf{B}_{is}^c - \frac{\gamma_{iw}^c}{w_{is}^c} \mathbf{w}_{is}^c (\mathbf{g}_{is}^c)' \right)'. \end{aligned} \quad (37)$$

Replacing $\mathbf{V}_{rs} \mathbf{V}_s^{-1} = \text{diag}_{1 \leq i \leq m} (\mathbf{V}_{irs} \mathbf{V}_{is}^{-1})$ and $\boldsymbol{\mu}_{r|s} = \text{col}_{1 \leq i \leq m} (\boldsymbol{\mu}_{ir|s})$ in the conditional mean vector (13) and using (37), we obtain the expression for $\boldsymbol{\mu}_{ir|s}$ given in (17) for a sampled area ($n_i > 0$). For a nonsampled area ($n_i = 0$), we have $\boldsymbol{\mu}_{ir|s} = \mathbf{X}_{ir} \boldsymbol{\beta}$.

A.3. Conditional covariance matrix

This section spells out the expression of the conditional covariance matrix $\mathbf{V}_{r|s}$ given in (14). Since $\mathbf{V}_r$, $\mathbf{V}_{r|s}$ and $\mathbf{V}_s$ are all block-diagonal, then $\mathbf{V}_{r|s} = \text{diag}_{1 \leq i \leq m}(\mathbf{V}_{ir|s})$, where

$$\mathbf{V}_{ir|s} = \mathbf{V}_{ir} - \mathbf{V}_{irs} \mathbf{V}_{is}^{-1} \mathbf{V}_{isr}, \quad i = 1, \dots, m. \quad (38)$$

For the non-sample model (9), $\mathbf{V}_{ir}$ is given in (10). Additionally, using the expression of $\mathbf{V}_{irs}$ given in (33), we have

$$\begin{aligned} \mathbf{V}_{irs} \mathbf{V}_{is}^{-1} \mathbf{V}_{isr} &= (A + B) \mathbf{V}_{is}^{-1} (A + B)' = A \mathbf{V}_{is}^{-1} A' + A \mathbf{V}_{is}^{-1} B' + B \mathbf{V}_{is}^{-1} A' + B \mathbf{V}_{is}^{-1} B' \\ &= \mathbf{L}_{i1} + \mathbf{L}_{i2} + \mathbf{L}_{i3} + \mathbf{L}_{i4}. \end{aligned} \quad (39)$$

The first term on the right-hand side of (39) is given by

$$\mathbf{L}_{i1} = A \mathbf{V}_{is}^{-1} A' = \sigma_v^4 \mathbf{1}_{N_i - n_i} \mathbf{1}_{n_i}' \mathbf{V}_{is}^{-1} \mathbf{1}_{n_i} \mathbf{1}_{N_i - n_i}' = \sigma_v^4 (\mathbf{1}_{n_i}' \mathbf{V}_{is}^{-1} \mathbf{1}_{n_i}) \mathbf{1}_{N_i - n_i} \mathbf{1}_{N_i - n_i}' \quad (40)$$

where

$$\mathbf{1}_{n_i}' \mathbf{V}_{is}^{-1} \mathbf{1}_{n_i} = \frac{1}{\sigma_e^2} \left[\mathbf{1}_{n_i}' \mathbf{W}_{is}^c \mathbf{1}_{n_i} - \frac{\gamma_{iw}^c}{w_{is}^c} \mathbf{1}_{n_i}' \mathbf{w}_{is}^c (\mathbf{w}_{is}^c)' \mathbf{1}_{n_i} \right]. \quad (41)$$

But note that

$$\mathbf{1}_{n_i}' \mathbf{W}_{is}^c \mathbf{1}_{n_i} = \mathbf{1}_{n_i}' \mathbf{W}_{is} \mathbf{1}_{n_i} - \frac{\sigma_u^2}{\sigma_e^2} \mathbf{1}_{n_i}' \mathbf{B}_{is} \Delta_i^{-1} \mathbf{B}_{is}' \mathbf{1}_{n_i} = w_{is} - \frac{\sigma_u^2}{\sigma_e^2} \mathbf{g}_{is}' \Delta_i^{-1} \mathbf{g}_{is} = w_{is}^c. \quad (42)$$

Replacing (42) and $\mathbf{1}_{n_i}' \mathbf{w}_{is}^c = w_{is}^c$ in (41) and the result in (40), we obtain

$$\mathbf{L}_{i1} = \frac{\sigma_v^4}{\sigma_e^2} (1 - \gamma_{iw}^c) w_{is}^c \mathbf{1}_{N_i - n_i} \mathbf{1}_{N_i - n_i}' = \sigma_v^2 \gamma_{iw}^c \mathbf{1}_{N_i - n_i} \mathbf{1}_{N_i - n_i}'.$$

Concerning the second term on the right hand side of (39), noting that $\mathbf{1}_{N_i - n_i} = \mathbf{A}_{ir} \mathbf{1}_T$ and using the expression of $\mathbf{V}_{is}^{-1}$, we have

$$\begin{aligned} \mathbf{L}_{i2} &= A \mathbf{V}_{is}^{-1} B' = \sigma_v^2 \sigma_u^2 \mathbf{1}_{N_i - n_i} \mathbf{1}_{n_i}' \mathbf{V}_{is}^{-1} \mathbf{A}_{is} \Omega_u \mathbf{A}_{ir}' \\ &= \frac{\sigma_v^2 \sigma_u^2}{\sigma_e^2} \mathbf{A}_{ir} \mathbf{1}_T \mathbf{1}_{n_i}' \left[\mathbf{W}_{is}^c - \frac{\gamma_{iw}^c}{w_{is}^c} \mathbf{w}_{is}^c (\mathbf{w}_{is}^c)' \right] \mathbf{A}_{is} \Omega_u \mathbf{A}_{ir}', \end{aligned}$$

where

$$\mathbf{1}_{n_i}' \left[\mathbf{W}_{is}^c - \frac{\gamma_{iw}^c}{w_{is}^c} \mathbf{w}_{is}^c (\mathbf{w}_{is}^c)' \right] \mathbf{A}_{is} = \mathbf{1}_{n_i}' \mathbf{W}_{is}^c \mathbf{A}_{is} - \frac{\gamma_{iw}^c}{w_{is}^c} \mathbf{1}_{n_i}' \mathbf{w}_{is}^c (\mathbf{w}_{is}^c)' \mathbf{A}_{is}. \quad (43)$$

The first term in (43) is given by

$$\mathbf{1}_{n_i}' \mathbf{W}_{is}^c \mathbf{A}_{is} = \mathbf{1}_{n_i}' \mathbf{W}_{is} \mathbf{A}_{is} - \frac{\sigma_u^2}{\sigma_e^2} \mathbf{1}_{n_i}' \mathbf{B}_{is} \Delta_i^{-1} \mathbf{B}_{is}' \mathbf{A}_{is} = \mathbf{g}_{is}' - \frac{\sigma_u^2}{\sigma_e^2} \mathbf{g}_{is}' \Delta_i^{-1} \mathbf{G}_{is} = (\mathbf{g}_{is}^c)'. \quad (44)$$

The second term is

$$\begin{aligned}\mathbf{1}_{n_i}' \mathbf{w}_{is}^c (\mathbf{w}_{is}^c)' \mathbf{A}_{is} &= \left(\mathbf{1}_{n_i}' \mathbf{w}_{is} - \frac{\sigma_u^2}{\sigma_e^2} \mathbf{1}_{n_i}' \mathbf{B}_{is} \mathbf{\Delta}_i^{-1} \mathbf{g}_{is} \right) \left(\mathbf{w}_{is}' \mathbf{A}_{is} - \frac{\sigma_u^2}{\sigma_e^2} \mathbf{g}_{is}' \mathbf{\Delta}_i^{-1} \mathbf{B}_{is}' \mathbf{A}_{is} \right) \\ &= \left(w_{is.}^c - \frac{\sigma_u^2}{\sigma_e^2} \mathbf{g}_{is}' \mathbf{\Delta}_i^{-1} \mathbf{g}_{is} \right) \left(\mathbf{g}_{is}' - \frac{\sigma_u^2}{\sigma_e^2} \mathbf{g}_{is}' \mathbf{\Delta}_i^{-1} \mathbf{G}_{is} \right) = w_{is.}^c (\mathbf{g}_{is}^c)'.\end{aligned}$$

Thus, we can write $\mathbf{L}_{i2}$ as

$$\begin{aligned}\mathbf{L}_{i2} &= \frac{\sigma_v^2 \sigma_u^2}{\sigma_e^2} \mathbf{A}_{ir} \mathbf{1}_T \left[(\mathbf{g}_{is}^c)' - \frac{\gamma_{iw}^c}{w_{is.}^c} w_{is.}^c (\mathbf{g}_{is}^c)' \right] \mathbf{\Omega}_u \mathbf{A}_{ir}' \\ &= \frac{\sigma_v^2 \sigma_u^2}{\sigma_e^2} \mathbf{A}_{ir} \mathbf{1}_T (1 - \gamma_{iw}^c) (\mathbf{g}_{is}^c)' \mathbf{\Omega}_u \mathbf{A}_{ir}' \\ &= \frac{\sigma_u^2 \gamma_{iw}^c}{w_{is.}^c} \mathbf{A}_{ir} \mathbf{1}_T (\mathbf{g}_{is}^c)' \mathbf{\Omega}_u \mathbf{A}_{ir}' = \frac{\sigma_u^2 \gamma_{iw}^c}{w_{is.}^c} \mathbf{A}_{ir} \mathbf{V}_{2i} \mathbf{\Omega}_u \mathbf{A}_{ir}',\end{aligned}$$

where $\mathbf{V}_{2i} = \mathbf{1}_T (\mathbf{g}_{is}^c)'$. The third term satisfies $\mathbf{L}_{i3} = \mathbf{L}_{i2}'$.

The last term on the right hand side of (39) is given by

$$\mathbf{L}_{i4} = B \mathbf{V}_{is}^{-1} B' = \frac{\sigma_u^4}{\sigma_e^2} \mathbf{A}_{ir} \mathbf{\Omega}_u \mathbf{A}_{is}' \left[\mathbf{W}_{is}^c - \frac{\gamma_{iw}^c}{w_{is.}^c} \mathbf{w}_{is}^c (\mathbf{w}_{is}^c)' \right] \mathbf{A}_{is} \mathbf{\Omega}_u \mathbf{A}_{ir}'. \quad (44)$$

We calculate now

$$\mathbf{A}_{is}' \left[\mathbf{W}_{is}^c - \frac{\gamma_{iw}^c}{w_{is.}^c} \mathbf{w}_{is}^c (\mathbf{w}_{is}^c)' \right] \mathbf{A}_{is} = \mathbf{A}_{is}' \mathbf{W}_{is}^c \mathbf{A}_{is} - \frac{\gamma_{iw}^c}{w_{is.}^c} \mathbf{A}_{is}' \mathbf{w}_{is}^c (\mathbf{w}_{is}^c)' \mathbf{A}_{is}.$$

Replacing $\mathbf{W}_{is}^c = \mathbf{W}_{is} - \frac{\sigma_u^2}{\sigma_e^2} \mathbf{B}_{is} \mathbf{\Delta}_i^{-1} \mathbf{B}_{is}'$, the first term becomes

$$\mathbf{A}_{is}' \mathbf{W}_{is}^c \mathbf{A}_{is} = \mathbf{A}_{is}' \mathbf{W}_{is} \mathbf{A}_{is} - \frac{\sigma_u^2}{\sigma_e^2} \mathbf{A}_{is}' \mathbf{B}_{is} \mathbf{\Delta}_i^{-1} \mathbf{B}_{is}' \mathbf{A}_{is} = \mathbf{G}_{is}^c,$$

for $\mathbf{G}_{is}^c = \mathbf{G}_{is} - \frac{\sigma_u^2}{\sigma_e^2} \mathbf{G}_{is} \mathbf{\Delta}_i^{-1} \mathbf{G}_{is}$ and the second term is given by

$$\begin{aligned}\mathbf{A}_{is}' \mathbf{w}_{is}^c (\mathbf{w}_{is}^c)' \mathbf{A}_{is} &= \left(\mathbf{A}_{is}' \mathbf{w}_{is} - \frac{\sigma_u^2}{\sigma_e^2} \mathbf{A}_{is}' \mathbf{B}_{is} \mathbf{\Delta}_i^{-1} \mathbf{g}_{is} \right) \left(\mathbf{w}_{is}' \mathbf{A}_{is} - \frac{\sigma_u^2}{\sigma_e^2} \mathbf{g}_{is}' \mathbf{\Delta}_i^{-1} \mathbf{B}_{is}' \mathbf{A}_{is} \right) \\ &= \left(\mathbf{g}_{is} - \frac{\sigma_u^2}{\sigma_e^2} \mathbf{G}_{is} \mathbf{\Delta}_i^{-1} \mathbf{g}_{is} \right) \left(\mathbf{g}_{is}' - \frac{\sigma_u^2}{\sigma_e^2} \mathbf{g}_{is}' \mathbf{\Delta}_i^{-1} \mathbf{G}_{is} \right) = \mathbf{g}_{is}^c (\mathbf{g}_{is}^c)'\end{aligned}$$

Therefore,

$$\mathbf{A}_{is}' \left[\mathbf{W}_{is}^c - \frac{\gamma_{iw}^c}{w_{is.}^c} \mathbf{w}_{is}^c (\mathbf{w}_{is}^c)' \right] \mathbf{A}_{is} = \mathbf{G}_{is}^c - \frac{\gamma_{iw}^c}{w_{is.}^c} \mathbf{g}_{is}^c (\mathbf{g}_{is}^c)'.$$

Thus, replacing the above result in (44), we can write $\mathbf{L}_{i4}$ as

$$\mathbf{L}_{i4} = \frac{\sigma_u^4}{\sigma_e^2} \mathbf{A}_{ir} \mathbf{\Omega}_u \mathbf{V}_{4i} \mathbf{\Omega}_u \mathbf{A}_{ir}',$$

where $\mathbf{V}_{4i} = \mathbf{G}_{is}^c - \frac{\gamma_{iw}^c}{w_{is}^c} \mathbf{g}_{is}^c (\mathbf{g}_{is}^c)'$. Summarizing, we have obtained

$$\mathbf{L}_{i1} = \sigma_v^2 \gamma_{iw}^c \mathbf{1}_{N_i - n_i} \mathbf{1}'_{N_i - n_i}, \quad \mathbf{L}_{i2} = \mathbf{L}'_{i3} = \frac{\sigma_u^2 \gamma_{iw}^c}{w_{is}^c} \mathbf{A}_{ir} \mathbf{V}_{2i} \mathbf{\Omega}_u \mathbf{A}'_{ir}, \quad \mathbf{L}_{i4} = \frac{\sigma_u^4}{\sigma_e^2} \mathbf{A}_{ir} \mathbf{\Omega}_u \mathbf{V}_{4i} \mathbf{\Omega}_u \mathbf{A}'_{ir}.$$

Hence, for an area i that is sampled (with $n_i > 0$), replacing the expression of $\mathbf{V}_{ir}$ given in (10) and (39) in (38), we obtain

$$\begin{aligned} \mathbf{V}_{ir|s} &= \mathbf{V}_{ir} - (\mathbf{L}_{i1} + \mathbf{L}_{i2} + \mathbf{L}_{i3} + \mathbf{L}_{i4}) \\ &= \sigma_e^2 \mathbf{W}_{ir}^{-1} + \sigma_v^2 (1 - \gamma_{iw}^c) \mathbf{1}_{N_i - n_i} \mathbf{1}'_{N_i - n_i} \\ &\quad + \sigma_u^2 \mathbf{A}_{ir} \left[\mathbf{\Omega}_u - \frac{\gamma_{iw}^c}{w_{is}^c} (\mathbf{V}_{2i} \mathbf{\Omega}_u + \mathbf{\Omega}_u \mathbf{V}'_{2i}) - \frac{\sigma_u^2}{\sigma_e^2} \mathbf{\Omega}_u \mathbf{V}_{4i} \mathbf{\Omega}_u \right] \mathbf{A}'_{ir}, \end{aligned}$$

which equals (38). If area i is nonsampled ($n_i = 0$), then

$$\mathbf{V}_{ir|s} = \sigma_e^2 \mathbf{W}_i^{-1} + \sigma_v^2 \mathbf{1}_{N_i} \mathbf{1}'_{N_i} + \sigma_u^2 \mathbf{A}_i \mathbf{\Omega}_u \mathbf{A}'_i. \quad (45)$$

The diagonal elements can be written as $v_{itj|s}^2 = \mathbf{a}'_{itj} \mathbf{V}_{ir|s} \mathbf{a}_{itj}$, for the vector $\mathbf{a}_{itj} = (\mathbf{0}'_{N_{i1} - n_{i1}}, \dots, \mathbf{0}'_{N_{i,t-1} - n_{i,t-1}}, \boldsymbol{\alpha}'_{itj}, \mathbf{0}'_{N_{i,t+1} - n_{i,t+1}}, \dots, \mathbf{0}'_{N_{iT} - n_{iT}})'$, where $\boldsymbol{\alpha}_{itj}$ is a $(N_{it} - n_{it}) \times 1$ vector of zeros except for 1 in position j . Using the expression for $\mathbf{V}_{ir|s}$ given in (45), we have

$$v_{itj|s}^2 = \mathbf{a}'_{itj} [\sigma_e^2 \mathbf{W}_{ir}^{-1} + \sigma_v^2 (1 - \gamma_{iw}^c) \mathbf{1}_{N_i - n_i} \mathbf{1}'_{N_i - n_i} + \sigma_u^2 \mathbf{A}_{ir} \mathbf{T}_{ir|s} \mathbf{A}_{ir}] \mathbf{a}_{itj}.$$

But note that it holds

$$\mathbf{a}'_{itj} \mathbf{W}_{ir}^{-1} \mathbf{a}_{itj} = w_{itj}^{-1}, \quad \mathbf{a}'_{itj} \mathbf{1}_{N_i - n_i} \mathbf{1}'_{N_i - n_i} \mathbf{a}_{itj} = 1, \quad \mathbf{a}'_{itj} \mathbf{A}_{ir} \mathbf{T}_{ir|s} \mathbf{A}'_{ir} \mathbf{a}_{itj} = \mathbf{b}'_t \mathbf{T}_{ir|s} \mathbf{b}_t.$$

for $\mathbf{b}_t = (\mathbf{0}'_{t-1}, 1, \mathbf{0}'_{T-t})'$. Consequently, we get

$$v_{itj|s}^2 = \sigma_e^2 w_{itj}^{-1} + \sigma_v^2 (1 - \gamma_{iw}^c) + \sigma_u^2 \mathbf{b}'_t \mathbf{T}_{ir|s} \mathbf{b}_t, \quad (46)$$

where

$$\mathbf{b}'_t \mathbf{T}_{ir|s} \mathbf{b}_t = \mathbf{b}'_t \mathbf{\Omega}_u \mathbf{b}_t - 2 \frac{\gamma_{iw}^c}{w_{is}^c} \mathbf{b}'_t \mathbf{V}_{2i} \mathbf{\Omega}_u \mathbf{b}_t - \frac{\sigma_u^2}{\sigma_e^2} \mathbf{b}'_t \mathbf{\Omega}_u \mathbf{V}_{4i} \mathbf{\Omega}_u \mathbf{b}_t.$$

Note that if $\mathbf{A}$ is a $T \times T$ matrix, then $\mathbf{b}'_t \mathbf{A} \mathbf{b}_t$ gives the t -th diagonal element of $\mathbf{A}$. Then $\mathbf{b}'_t \mathbf{\Omega}_u \mathbf{b}_t = (1 - \rho^2)^{-1}$ for any t . Similarly, we can obtain $\mathbf{b}'_t \mathbf{V}_{2i} \mathbf{\Omega}_u \mathbf{b}_t$ and $\mathbf{b}'_t \mathbf{\Omega}_u \mathbf{V}_{4i} \mathbf{\Omega}_u \mathbf{b}_t$ as the t -th elements of $\mathbf{V}_{2i} \mathbf{\Omega}_u$ and $\mathbf{\Omega}_u \mathbf{V}_{4i} \mathbf{\Omega}_u$ respectively.

Appendix B: Province poverty incidences and gaps for women

Table 5: Direct, CEBE and CEBTe estimates of poverty incidences and estimated cv's (%) for women by province in 2006 (provinces sorted from smaller to larger sample size).

Province	n_{iT}	$\hat{F}_{0,iT}^{dir}$	$\hat{F}_{0,iT}^{cebe}$	$\hat{F}_{0,iT}^{cebt}$	$cv(\hat{F}_{0,iT}^{dir})$	$cv(\hat{F}_{0,iT}^{cebe})$	$cv(\hat{F}_{0,iT}^{cebt})$
Soria	17	55.62	33.44	34.51	42.69	19.06	14.88
Segovia	58	43.80	22.80	27.16	22.31	19.28	13.05
Ávila	60	44.03	40.30	40.23	21.73	10.70	9.05
Palencia	72	28.00	29.08	29.30	22.32	13.55	11.98
Teruel	78	15.09	22.34	23.97	31.38	17.50	13.90
Guadalajara	86	16.46	18.25	18.63	26.83	19.64	17.58
Cuenca	87	47.17	36.30	37.97	18.03	10.96	9.14
Álava	96	7.86	12.73	12.56	42.91	27.14	24.80
Zamora	100	26.82	33.44	36.49	19.36	11.61	9.33
Huesca	105	8.92	19.70	18.85	37.70	17.77	16.58
Castellón	123	12.24	17.48	20.62	31.80	16.93	13.65
Huelva	123	25.31	31.72	32.06	17.32	10.30	9.20
Lérida	133	24.10	26.99	25.44	21.00	12.51	12.31
Gerona	138	6.72	16.96	15.59	34.86	16.86	17.39
Tarragona	139	17.36	19.20	19.40	20.68	14.42	12.93
Orense	140	20.82	22.33	22.84	23.56	14.12	13.01
Salamanca	162	32.90	30.54	30.86	15.09	10.13	9.28
Burgos	167	12.43	19.71	19.30	23.47	15.23	14.16
Lugo	177	22.39	28.16	29.38	17.22	10.45	9.23
Melilla	180	25.10	26.30	25.58	14.70	13.03	12.82
Albacete	183	28.52	32.87	33.72	15.06	8.38	7.48
Almería	193	35.81	36.27	36.78	12.64	8.37	7.84
León	228	19.28	23.25	23.60	16.30	11.54	10.66
Granada	229	34.17	35.68	36.16	12.37	6.73	6.53
Córdoba	230	30.66	30.45	31.93	12.66	7.94	7.18
Jaén	230	33.86	32.66	32.68	11.78	8.11	7.92
Ciudad Real	239	32.37	30.51	32.33	12.52	8.38	7.37
Ceuta	247	38.79	33.66	34.73	11.29	10.47	9.77
Toledo	272	21.99	23.69	24.83	14.06	10.62	9.60
Guipúzcoa	291	9.95	13.85	13.46	21.63	16.55	16.88
Cáceres	302	33.06	28.32	30.05	11.47	7.89	7.33
Valladolid	306	20.64	18.18	18.35	14.32	12.76	12.49
Tenerife	392	28.53	26.04	26.45	10.44	7.67	7.96
Málaga	397	25.76	33.50	32.99	10.00	6.25	6.29
Cádiz	422	30.13	30.15	30.90	9.60	6.38	6.45
Cantabria	443	12.92	16.98	16.90	16.44	11.15	10.93
Pontevedra	462	21.31	22.08	21.95	12.02	9.31	9.52
Las Palmas	485	24.65	26.89	26.76	11.31	7.26	7.55
Sevilla	491	22.80	24.43	24.41	10.05	7.64	8.13
La Rioja	500	21.34	21.76	21.73	11.66	8.84	9.12
Badajoz	515	39.33	38.54	38.49	7.89	5.15	5.21
Vizcaya	532	9.88	11.90	11.72	15.26	15.39	15.34
La Coruña	533	23.76	23.60	23.42	9.37	8.00	8.01
Alicante	552	19.04	23.67	23.35	10.58	7.77	8.17
Zaragoza	574	13.27	15.78	15.47	13.57	12.03	12.58

Continued on next page

Table 5 – *Continued from previous page*

Dom	n_{iT}	$\hat{F}_{0,iT}^{dir}$	$\hat{F}_{0,iT}^{cebe}$	$\hat{F}_{0,iT}^{cebte}$	$cv(\hat{F}_{0,iT}^{dir})$	$cv(\hat{F}_{0,iT}^{cebe})$	$cv(\hat{F}_{0,iT}^{cebte})$
Navarra	603	9.41	11.26	10.89	15.30	16.08	17.29
Baleares	660	11.48	16.53	16.54	13.01	10.67	10.60
Valencia	742	13.88	21.34	21.27	10.87	7.79	7.68
Asturias	864	12.15	16.15	15.85	11.59	10.66	11.25
Murcia	902	25.55	29.05	28.92	7.85	5.12	5.68
Madrid	996	12.50	14.56	14.07	11.29	9.16	10.17
Barcelona	1,483	10.87	13.89	13.50	7.86	9.32	10.09

Table 6: Direct, CEBE and CEBTe estimates of poverty gaps and estimated cv's (%) for women by province in 2006 (provinces sorted from smaller to larger sample size).

Province	n_{iT}	$\hat{F}_{1,iT}^{dir}$	$\hat{F}_{1,iT}^{cebe}$	$\hat{F}_{1,iT}^{cebte}$	$cv(\hat{F}_{1,iT}^{dir})$	$cv(\hat{F}_{1,iT}^{cebe})$	$cv(\hat{F}_{1,iT}^{cebte})$
Soria	17	24.97	12.66	12.97	60.52	26.20	21.08
Segovia	58	10.91	7.54	9.29	26.24	30.96	20.04
Ávila	60	12.39	16.06	15.78	24.63	14.38	12.43
Palencia	72	6.08	10.28	10.22	30.74	20.05	17.95
Teruel	78	5.16	7.37	7.93	39.11	28.58	22.50
Guadalajara	86	2.55	5.76	5.81	27.96	32.51	30.01
Cuenca	87	10.74	13.68	14.35	20.24	15.73	12.92
Álava	96	1.53	3.73	3.61	47.22	48.12	44.04
Zamora	100	9.87	12.21	13.54	23.81	17.04	13.31
Huesca	105	3.04	6.31	5.88	44.55	30.38	29.09
Castellón	123	3.88	5.44	6.58	39.31	28.65	22.02
Huelva	123	9.13	11.46	11.46	23.65	15.01	13.31
Lérida	133	9.44	9.41	8.58	25.07	18.69	19.10
Gerona	138	2.15	5.36	4.77	40.97	29.35	32.01
Tarragona	139	2.79	6.15	6.13	24.55	24.10	21.84
Orense	140	4.79	7.34	7.44	33.29	23.02	21.64
Salamanca	162	8.85	11.09	11.08	17.48	14.62	13.58
Burgos	167	3.51	6.32	6.07	36.13	24.91	23.98
Lugo	177	4.91	9.82	10.23	23.08	15.98	14.49
Melilla	180	12.47	9.21	8.75	16.15	20.35	20.47
Albacete	183	11.72	12.06	12.31	17.64	11.56	10.40
Almería	193	11.32	13.89	13.95	16.61	11.64	10.65
León	228	7.56	7.68	7.71	21.13	18.38	17.17
Granada	229	16.81	13.52	13.58	15.15	9.85	9.53
Córdoba	230	8.01	10.84	11.39	14.99	11.36	10.46
Jaén	230	11.42	11.89	11.74	14.35	12.36	12.53
Ciudad Real	239	7.16	10.96	11.67	14.11	12.68	11.10
Ceuta	247	20.68	12.58	12.94	13.49	14.96	14.08
Toledo	272	5.85	7.87	8.25	17.19	16.97	15.35
Guipúzcoa	291	4.42	4.11	3.91	26.78	28.92	30.14
Cáceres	302	9.23	9.90	10.55	13.79	12.11	11.26
Valladolid	306	7.25	5.71	5.69	17.26	20.94	20.71
Tenerife	392	9.34	8.94	8.99	13.22	11.73	12.16
Málaga	397	10.80	12.40	11.98	13.89	9.00	9.35
Cádiz	422	10.81	10.70	10.91	12.02	9.29	9.66

Continued on next page

Table 6 – *Continued from previous page*

Dom	n_{iT}	$\hat{F}_{1,iT}^{dir}$	$\hat{F}_{1,iT}^{cebe}$	$\hat{F}_{1,iT}^{cebte}$	$cv(\hat{F}_{1,iT}^{dir})$	$cv(\hat{F}_{1,iT}^{cebe})$	$cv(\hat{F}_{1,iT}^{cebte})$
Cantabria	443	2.98	5.22	5.11	19.75	19.13	19.44
Pontevedra	462	4.38	7.23	7.08	13.95	15.08	15.83
Las Palmas	485	8.48	9.36	9.17	16.43	11.35	11.88
Sevilla	491	4.53	8.18	8.07	12.88	12.11	13.06
La Rioja	500	4.36	7.17	7.05	13.43	14.32	14.76
Badajoz	515	12.15	14.91	14.69	9.78	7.22	7.52
Vizcaya	532	2.26	3.42	3.31	18.96	27.81	28.47
La Coruña	533	8.36	7.90	7.70	11.86	12.37	12.54
Alicante	552	5.95	7.97	7.71	14.08	13.15	13.88
Zaragoza	574	4.73	4.80	4.61	20.40	20.94	21.95
Navarra	603	2.73	3.22	3.05	19.55	28.97	32.10
Baleares	660	2.88	5.10	5.03	18.44	19.16	18.55
Valencia	742	4.26	6.95	6.82	13.80	12.89	12.69
Asturias	864	3.09	4.88	4.70	16.03	18.60	20.04
Murcia	902	8.30	10.33	10.12	8.58	8.02	9.08
Madrid	996	3.54	4.39	4.15	17.18	16.32	18.11
Barcelona	1,483	3.62	4.14	3.94	10.26	17.12	18.45

References

- BENAVENT, R. & MORALES, D. (2020). Small area estimation under a temporal bivariate area-level linear mixed model with independent time effects. *STATISTICAL METHODS AND APPLICATIONS* .
- BOUBETA, M., LOMBARDÍA, M. J. & MORALES, D. (2017). Poisson mixed models for studying the poverty in small areas. *Computational Statistics & Data Analysis* **107**, 32–47.
- CHOUDHRY, G. & RAO, J. N. K. (1989). Small area estimation using models that combine time series and cross sectional data. In *Proceedings of Statistics Canada Symposium on Analysis of data in time*.
- CORREA, L., MOLINA, I. & RAO, J. N. K. (2012). Comparison of methods for estimation of poverty indicators in small areas. *Unpublished report* .
- DATTA, G., LAHIRI, P. & MAITI, T. (2002). Empirical bayes estimation of median income of four-person families by state using time series and cross-sectional data. *Journal of Statistical Planning and inference* **102**, 83–97.
- DATTA, G. S., LAHIRI, P., MAITI, T. & LU, K. L. (1999). Hierarchical bayes estimation of unemployment rates for the states of the us. *Journal of the American Statistical Association* **94**, 1074–1082.
- DIALLO, M. S. & RAO, J. N. K. (2018). Small area estimation of complex parameters under unit-level models with skew-normal errors. *Scandinavian Journal of Statistics* **45**, 1092–1116.

- ESTEBAN, M., MORALES, D., PÉREZ, A. & SANTAMARÍA, L. (2012a). Two area-level time models for estimating small area poverty indicators. *Journal of the Indian Society of Agricultural Statistics* **66**, 75–89.
- ESTEBAN, M. D., MORALES, D., PÉREZ, A. & SANTAMARÍA, L. (2012b). Small area estimation of poverty proportions under area-level time models. *Computational Statistics & Data Analysis* **56**, 2840–2855.
- GHOSH, M., NANGIA, N. & KIM, D. H. (1996). Estimation of median income of four-person families: a bayesian time series approach. *Journal of the American Statistical Association* **91**, 1423–1431.
- GONZÁLEZ-MANTEIGA, W., LOMBARDIA, M. J., MOLINA, I., MORALES, D. & SANTAMARÍA, L. (2008). Bootstrap mean squared error of a small area eblup. *Journal of Statistical Computation and Simulation* **78**, 443–462.
- GRAF, M., MARÍN, J. M. & MOLINA, I. (2019). A generalized mixed model for skewed distributions applied to small area estimation. *Test* **28**, 565–597.
- GUADARRAMA, M., MOLINA, I. & RAO, J. (2016). A comparison of small area estimation methods for poverty mapping. *Statistics in Transition new series* **1**, 41–66.
- HÁJEK, J. (1971). Comment on a paper by d. basu. *Foundations of statistical inference* **236**.
- HOBZA, T., MORALES, D. & MARHUENDA, Y. (2020). Small area estimation of additive parameters under unit-level generalized linear mixed models. *Sort* **44**.
- HOBZA, T., MORALES, D. & SANTAMARÍA, L. (2018). Small area estimation of poverty proportions under unit-level temporal binomial-logit mixed models. *Test* **27**, 270–294.
- MARHUENDA, Y., MOLINA, I. & MORALES, D. (2013). Small area estimation with spatio-temporal fay–herriot models. *Computational Statistics & Data Analysis* **58**, 308–325.
- MARHUENDA, Y., MOLINA, I., MORALES, D. & RAO, J. (2017). Poverty mapping in small areas under a twofold nested error regression model. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* **180**, 1111–1136.
- MOLINA, I., NANDRAM, B., RAO, J. et al. (2014). Small area estimation of general parameters with application to poverty indicators: A hierarchical bayes approach. *The Annals of Applied Statistics* **8**, 852–885.
- MOLINA, I. & RAO, J. N. K. (2010). Small area estimation of poverty indicators. *Canadian Journal of Statistics* **38**, 369–385.
- MORALES, D., PAGLIARELLA, M. C. & SALVATORE, R. (2015). Small area estimation of poverty indicators under partitioned area-level time models. *SORT: statistics and operations research transactions* **39**, 019–34.

- MORALES, D. & SANTAMARÍA, L. (2019). Small area estimation under unit-level temporal linear mixed models. *Journal of Statistical Computation and Simulation* **89**, 1592–1620.
- PFEFFERMANN, D. & BURCK, L. (1990). Robust small area estimation combining time series and cross- sectional data. *Survey Methodology* **16**, 217–237.
- RAO, J. N. & YU, M. (1994). Small-area estimation by combining time-series and cross-sectional data. *Canadian Journal of Statistics* **22**, 511–528.
- RAO, J. N. K. & MOLINA, I. (2015). *Small area estimation*. John Wiley & Sons.
- ROJAS-PERILLA, N., PANNIER, S., SCHMID, T. & TZAVIDIS, N. (2020). Data-driven transformations in small area estimation. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* **183**, 121–148.
- SINGH, A., MANTEL, H. & THOMAS, B. (1991). Time series generalizations of fay-herriot estimator for small areas. In *Proceedings of Survey Research Methods Section*.
- YOU, Y. & RAO, J. (2000). Hierarchical bayes estimation of small area means using multi-level models. *Survey Methodology* **26**, 173–181.
- YOU, Y., RAO, J. & GAMBINO, J. (2001). Model-based unemployment rate estimation for the canadian labour force survey: a hierarchical approach. Tech. rep., Technical Report.