Chapter 12
EBPs Under Two-Fold Nested Error Regression Models

12.1 Introduction

For the estimation of complex domain parameters such as certain poverty indicators, Molina and Rao (2010) proposed the empirical best (EB) method, based on assuming that a one-to-one transformation of the target variable follows the unit-level nested error model of Battese et al. (1988) with random effects for the domains of interest. Under that model, EB method gives approximately the "best" estimator in the sense of being unbiased with minimum variance error.

When the target population is naturally divided in subpopulations at two nested aggregation levels (e.g. in provinces and counties within provinces), or when the sampling design has two stages, as it is usual in many household surveys, it is reasonable to assume a two-fold nested error regression model including random effects at the two levels of aggregation, domains and subdomains. Marhuenda et al. (2017) developed the EB method for predicting additive parameters under the two-fold nested error regression model.

This chapter describes the EB methodology given by Marhuenda et al. (2017) for predicting additive parameters and provides analytical expressions for the EB predictors (EBP) of poverty proportions, poverty gaps, and average incomes. It gives Monte Carlo algorithms for approximating the EB predictors of more complex domain or subdomain parameters. The case of using only categorical explanatory variables is also treated, because it does not require the use of an auxiliary census data file. The obtained EB estimates of subdomain parameters have the good property of being consistent with the corresponding domain estimate.

For estimating the error variances of the EBPs, a parametric bootstrap procedure is given. The EBP methodology is illustrated with an application to the survey data file LFS20.txt. The given R codes calculate EBPs of poverty proportions, poverty gaps, and average incomes by areas and age groups.

D. Morales et al., *A Course on Small Area Estimation and Mixed Models*, Statistics for Social and Behavioral Sciences, https://doi.org/10.1007/978-3-030-63757-6_12

12.2 Two-fold Nested Error Regression Models

This section considers vectors $\boldsymbol{y} = (y_1, \dots, y_N)'$ containing the values of a target random variable associated with N units of a finite population. Let $\boldsymbol{y}_s$ be the sub-vector of $\boldsymbol{y}$ corresponding to sample elements and $\boldsymbol{y}_r$ the sub-vector of $\boldsymbol{y}$ corresponding to the out-of-sample elements; that is, $\boldsymbol{y} = (\boldsymbol{y}_s', \boldsymbol{y}_r')'$. The inference problem is to predict the value of a real-valued function $\delta = h(\boldsymbol{y})$ of the random vector $\boldsymbol{y}$ using the sample data $\boldsymbol{y}_s$. The best predictor of δ is given by $\hat{\delta}^{bp} = E_{\boldsymbol{y}_r}(\delta|\boldsymbol{y}_s)$, where the expectation is taken with respect to the conditional distribution of $\boldsymbol{y}_r$ given $\boldsymbol{y}_s$.

The population of interest is hierarchically divided in domains and subdomains. More concretely, let U be a population of size N partitioned into D domains or areas $U_1, \dots, U_D$ of sizes $N_1, \dots, N_D$, respectively. Additionally, each domain U_d is partitioned into M_d subdomains $U_{d1}, \dots, U_{dM_d}$, of sizes $N_{d1}, \dots, N_{dM_d}$, respectively, $d = 1, \dots, D$. The components of vector $\boldsymbol{y}$ are referenced with three subindexes. This is to say, y_{dtj} denotes the value that the study variable takes on the sample unit j of subdomain t and domain d.

12.2.1 The Population Model

At the population level, the two-fold nested error regression model is

$$y_{dtj} = \boldsymbol{x}_{dtj}\boldsymbol{\beta} + u_{1,d} + u_{2,dt} + w_{dtj}^{-1/2} e_{dtj},\ d = 1, \dots, D, t = 1, \dots, M_d,\ j = 1, \dots, N_{dt}, \tag{12.1}$$

where $\boldsymbol{x}_{dtj}$ is a row vector containing p auxiliary variables, $w_{dtj} > 0$ is a known heteroscedasticity weight, and the random effects and errors are all mutually independent and such that $u_{1,d} \sim N(0, \sigma_1^2)$, $u_{2,dt} \sim N(0, \sigma_2^2)$, and $e_{dtj} \sim N(0, \sigma_0^2)$.

The population model (12.1) can be written in the matrix form as (without taking into account reordering with respect to sampled and non-sampled elements)

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}_1\boldsymbol{u}_1 + \boldsymbol{Z}_2\boldsymbol{u}_2 + \boldsymbol{W}^{-1/2}\boldsymbol{e}, \tag{12.2}$$

where $\boldsymbol{u}_1 = \boldsymbol{u}_{1,D\times 1} \sim N(\boldsymbol{0}, \sigma_1^2\boldsymbol{I}_D)$, $\boldsymbol{u}_2 = \boldsymbol{u}_{2,M\times 1} \sim N(\boldsymbol{0}, \sigma_2^2\boldsymbol{I}_M)$, and $\boldsymbol{e} = \boldsymbol{e}_{N\times 1} \sim N(\boldsymbol{0}, \sigma_0^2\boldsymbol{I}_N)$ are independent, $\boldsymbol{y}{=}\boldsymbol{y}_{N\times 1} = \underset{1\le d\le D}{\text{col}}(\underset{1\le t\le M_d}{\text{col}}(\underset{1\le j\le N_{dt}}{\text{col}}(y_{dtj})))$, $\boldsymbol{\beta} = \boldsymbol{\beta}_{p\times 1}$, $\boldsymbol{X} = \boldsymbol{X}_{N\times p} = \underset{1\le d\le D}{\text{col}}(\underset{1\le t\le M_d}{\text{col}}(\underset{1\le j\le N_{dt}}{\text{col}}(\boldsymbol{x}_{dtj})))$ with $\text{rank}(\boldsymbol{X}) = p$, $M = \sum_{d=1}^D M_d$, $N = \sum_{d=1}^D N_d$, $N_d = \sum_{t=1}^{M_d} N_{dt}$, $\boldsymbol{Z}_1 = \underset{1\le d\le D}{\text{diag}}(\boldsymbol{1}_{N_d})_{N\times D}$, $\boldsymbol{Z}_2 = \underset{1\le d\le D}{\text{diag}}(\underset{1\le t\le M_d}{\text{diag}}(\boldsymbol{1}_{N_{dt}}))_{N\times M}$, $\boldsymbol{I}_a$ is the $a \times a$ identity matrix, $\boldsymbol{1}_a$ is the $a \times 1$

vector with all its elements equal to 1, $\boldsymbol{W} = \underset{1\le d\le D}{\text{diag}}\,(\boldsymbol{W}_d)$, $\boldsymbol{W}_d = \underset{1\le t\le M_d}{\text{diag}}\,(\boldsymbol{W}_{dt})$, $\boldsymbol{W}_{dt} = \underset{1\le j\le N_{dt}}{\text{diag}}\,(w_{dtj})_{N_{dt}\times N_{dt}}$ with known heteroscedasticity weights $w_{dtj} > 0$, $d = 1,\dots,D$, $t = 1,\dots,M_d$, $j = 1,\dots,N_{dt}$.

Without loss of generality we can reorder the population so that the target vector takes the form $\boldsymbol{y} = (\boldsymbol{y}_s', \boldsymbol{y}_r')'$. We describe the two corresponding sub-models more in detail.

12.2.2 The Sample Model

In practice, inference is carried out based on a sample drawn from the population. We assume that a sample s_d of size n_d is drawn from domain U_d, $d = 1,\dots,D$. Let s_{dt} be the subsample from subdomain U_{dt}, $t = 1,\dots,M_d$. We allow the existence of subdomains with no observations in the sample. Without loss of generality, we assume that these are the last $M_d - m_d$ subdomains; that is, $s_{dt} = \emptyset$, for $m_d + 1 \le t \le M_d$ whereas $s_{dt} \ne \emptyset$, for $1 \le t \le m_d$. The sample sub-vector $\boldsymbol{y}_s$ follows the marginal model derived from the population model (12.1), i.e.

$$y_{dtj} = \boldsymbol{x}_{dtj}\boldsymbol{\beta} + u_{1,d} + u_{2,dt} + w_{dtj}^{-1/2}e_{dtj},\ \ d = 1,\dots,D, t = 1,\dots,m_d, j = 1,\dots,n_{dt}, \tag{12.3}$$

where we change M, M_d and N, N_d, and N_{dt} by the sample counterparts m, m_d and n, n_d, and n_{dt}, respectively. In matrix notation, the model is

$$\boldsymbol{y}_s = \boldsymbol{X}_s\boldsymbol{\beta} + \boldsymbol{Z}_{1s}\boldsymbol{u}_1 + \boldsymbol{Z}_{2s}\boldsymbol{u}_{2s} + \boldsymbol{W}_s^{-1/2}\boldsymbol{e}_s, \tag{12.4}$$

where $\boldsymbol{u}_1 = \boldsymbol{u}_{1,D\times 1} \sim N(\boldsymbol{0}, \sigma_1^2\boldsymbol{I}_D)$, $\boldsymbol{u}_{2s} = \boldsymbol{u}_{2s,m\times 1} \sim N(\boldsymbol{0}, \sigma_2^2\boldsymbol{I}_m)$, and $\boldsymbol{e}_s = \boldsymbol{e}_{s,n\times 1} \sim N(\boldsymbol{0}, \sigma_0^2\boldsymbol{I}_n)$ are independent, $\boldsymbol{y}_s = \boldsymbol{y}_{s,n\times 1}$, $\boldsymbol{\beta} = \boldsymbol{\beta}_{p\times 1}$, $\boldsymbol{X}_s = \boldsymbol{X}_{s,n\times p} = \underset{1\le d\le D}{\text{col}}\,(\underset{1\le t\le m_d}{\text{col}}\,(\underset{1\le j\le n_{dt}}{\text{col}}\,(\boldsymbol{x}_{dtj})))$ with $\text{rank}(\boldsymbol{X}_s) = p$, $\boldsymbol{Z}_{1s} = \underset{1\le d\le D}{\text{diag}}\,(\boldsymbol{1}_{n_d})_{n\times D}$, $\boldsymbol{Z}_{2s} = \underset{1\le d\le D}{\text{diag}}\,(\underset{1\le t\le m_d}{\text{diag}}\,(\boldsymbol{1}_{n_{dt}}))_{n\times m}$, $m = \sum_{d=1}^D m_d$, $n = \sum_{d=1}^D n_d$, $n_d = \sum_{t=1}^{m_d} n_{dt}$, $\boldsymbol{W}_s = \underset{1\le d\le D}{\text{diag}}\,(\boldsymbol{W}_{ds})$, $\boldsymbol{W}_{ds} = \underset{1\le t\le m_d}{\text{diag}}\,(\boldsymbol{W}_{dts})$, $\boldsymbol{W}_{dts} = \underset{1\le j\le n_{dt}}{\text{diag}}\,(w_{dtj})_{n_{dt}\times n_{dt}}$ with known heteroscedasticity weights $w_{dtj} > 0, d = 1,\dots,D, t = 1,\dots,m_d, j = 1,\dots,n_{dt}$.

12.2.3 The Non-sample Model

Let r be the subset of units not appearing in the sample s. The corresponding sub-vector $\boldsymbol{y}_r$ follows the model (12.1), with the immediate modifications. For $d = 1,\dots,D, t = 1,\dots,M_d, j = n_{dt}+1,\dots,N_{dt}$, the non-sample model is

$$y_{dtj} = \boldsymbol{x}_{dtj}\boldsymbol{\beta} + u_{1,d} + u_{2,dt} + w_{dtj}^{-1/2}e_{dtj},$$

where we change N, N_d and N_{dt} by $N-n$, N_d-n_d, and $N_{dt}-n_{dt}$, respectively. In matrix notation, the model is

$$\boldsymbol{y}_r = \boldsymbol{X}_r\boldsymbol{\beta} + \boldsymbol{Z}_{1r}\boldsymbol{u}_1 + \boldsymbol{Z}_{2r}\boldsymbol{u}_{2r} + \boldsymbol{W}_r^{-1/2}\boldsymbol{e}_r, \tag{12.5}$$

where $\boldsymbol{u}_1 = \boldsymbol{u}_{1,D\times 1} \sim N(\boldsymbol{0}, \sigma_1^2\boldsymbol{I}_D)$, $\boldsymbol{u}_{2r} = \boldsymbol{u}_{2r,M\times 1} \sim N(\boldsymbol{0}, \sigma_2^2\boldsymbol{I}_M)$, and $\boldsymbol{e}_r = \boldsymbol{e}_{r,(N-n)\times 1} \sim N(\boldsymbol{0}, \sigma_0^2\boldsymbol{I}_{N-n})$ are independent, $\boldsymbol{y}_r = \boldsymbol{y}_{r,(N-n)\times 1}$, $\boldsymbol{\beta} = \boldsymbol{\beta}_{p\times 1}$, $\boldsymbol{X}_r = \boldsymbol{X}_{r,(N-n)\times p} = \underset{1\le d\le D}{\text{col}}\,(\underset{1\le t\le M_d}{\text{col}}\,(\underset{n_{dt}+1\le j\le N_{dt}}{\text{col}}\,(\boldsymbol{x}_{dtj})))$ with $\text{rank}(\boldsymbol{X}_r) = p$, $\boldsymbol{Z}_{1r} = \underset{1\le d\le D}{\text{diag}}\,(\boldsymbol{1}_{N_d-n_d})_{(N-n)\times D}$, $\boldsymbol{Z}_{2r} = \underset{1\le d\le D}{\text{diag}}\,(\underset{1\le t\le M_d}{\text{diag}}\,(\boldsymbol{1}_{N_{dt}-n_{dt}}))_{(N-n)\times M}$, $\boldsymbol{W}_r = \underset{1\le d\le D}{\text{diag}}\,(\boldsymbol{W}_{dr})$, $\boldsymbol{W}_{dr} = \underset{1\le t\le M_d}{\text{diag}}\,(\boldsymbol{W}_{dtr})$, $\boldsymbol{W}_{dtr} = \underset{n_{dt}+1\le j\le N_{dt}}{\text{diag}}\,(w_{dtj})_{(N_{dt}-n_{dt})\times(N_{dt}-n_{dt})}$ with known $w_{dtj} > 0$, $d = 1,\dots,D$, $t = 1,\dots,M_d$, $j = n_{dt}+1,\dots,N_{dt}$, and $n_{dt} = 0$ if $t > m_d$.

12.2.4 The Inverse of the Variance Matrix

Let $\boldsymbol{V}_s$ denote the covariance matrix of the sample vector $\boldsymbol{y}_s$. Direct calculation of $\boldsymbol{V}_s^{-1}$ is not computationally efficient because it requires the inversion of the $n\times n$ matrix $\boldsymbol{V}_s$. This is why we apply the inversion formula (cf. Appendix A)

$$(\boldsymbol{A} + \boldsymbol{u}\boldsymbol{v}')^{-1} = \boldsymbol{A}^{-1} - \frac{\boldsymbol{A}^{-1}\boldsymbol{u}\boldsymbol{v}'\boldsymbol{A}^{-1}}{1 + \boldsymbol{v}'\boldsymbol{A}^{-1}\boldsymbol{u}} \tag{12.6}$$

for deriving an expression for $\boldsymbol{V}_s^{-1}$. Note that variance of $\boldsymbol{y}_s$ is

$$\boldsymbol{V}_s = \text{var}(\boldsymbol{y}_s) = \boldsymbol{Z}_{1s}\text{var}(\boldsymbol{u}_1)\boldsymbol{Z}_{1s}' + \boldsymbol{Z}_{2s}\text{var}(\boldsymbol{u}_{2s})\boldsymbol{Z}_{2s}' + \sigma_0^2\boldsymbol{W}_s^{-1} = \text{diag}(\boldsymbol{V}_{1s},\dots,\boldsymbol{V}_{Ds}), \tag{12.7}$$

where

$$\boldsymbol{V}_{ds} = \sigma_1^2\boldsymbol{1}_{n_d}\boldsymbol{1}_{n_d}' + \sigma_2^2\underset{1\le t\le m_d}{\text{diag}}\,(\boldsymbol{1}_{n_{dt}}\boldsymbol{1}_{n_{dt}}') + \sigma_0^2\boldsymbol{W}_{ds}^{-1} = \sigma_1^2\boldsymbol{1}_{n_d}\boldsymbol{1}_{n_d}' + \boldsymbol{R}_{ds},\ d = 1,\dots,D,$$

$$\boldsymbol{R}_{ds} = \underset{1\le t\le m_d}{\text{diag}}\,(\sigma_2^2\boldsymbol{1}_{n_{dt}}\boldsymbol{1}_{n_{dt}}' + \sigma_0^2\boldsymbol{W}_{dts}^{-1}) = \underset{1\le t\le m_d}{\text{diag}}\,(\boldsymbol{R}_{dts}),\quad d = 1,\dots,D.$$

For $d = 1,\dots,D$, $t = 1,\dots,m_d$, we introduce the notation $\boldsymbol{w}_{n_{dt}} = \boldsymbol{W}_{dts}\boldsymbol{1}_{n_{dt}} = (w_{dt1},\dots,w_{dtn_{dt}})'_{n_{dt}\times 1}$, $w_{dt\cdot} = \boldsymbol{1}_{n_{dt}}'\boldsymbol{w}_{n_{dt}} = \sum_{j=1}^{n_{dt}} w_{dtj}$ and

$$\gamma_{dt} = \frac{\sigma_2^2}{\sigma_2^2 + \frac{\sigma_0^2}{w_{dt\cdot}}},\quad \varphi_d = \frac{\sigma_1^2}{\sigma_0^2 + \sigma_1^2\sum_{\ell=1}^{m_d}(1-\gamma_{d\ell})w_{d\ell\cdot}}. \tag{12.8}$$

For calculating $\boldsymbol{V}_s^{-1} = \text{diag}(\boldsymbol{V}_{1s}^{-1}, \dots, \boldsymbol{V}_{Ds}^{-1})$ it is necessary to obtain $\boldsymbol{R}_{ds}^{-1}$. Here we use twice the formula (12.6), first to calculate $\boldsymbol{R}_{dts}^{-1}$ and second to obtain $\boldsymbol{V}_{ds}^{-1}$. For calculating $\boldsymbol{R}_{dts}^{-1} = (\sigma_2^2 \mathbf{1}_{n_{dt}} \mathbf{1}'_{n_{dt}} + \sigma_0^2 \boldsymbol{W}_{dts}^{-1})^{-1}$, we put $\boldsymbol{A} = \sigma_0^2 \boldsymbol{W}_{dts}^{-1}$, $\boldsymbol{u} = \sigma_2^2 \mathbf{1}_{n_{dt}}$, $\boldsymbol{v}' = \mathbf{1}'_{n_{dt}}$ and we get

$$\boldsymbol{R}_{dts}^{-1} = \frac{1}{\sigma_0^2}\boldsymbol{W}_{dts} - \frac{\sigma_2^2}{\sigma_0^4}\frac{\boldsymbol{W}_{dts}\mathbf{1}_{n_{dt}}\mathbf{1}'_{n_{dt}}\boldsymbol{W}_{dts}}{1+\frac{\sigma_2^2}{\sigma_0^2}\mathbf{1}'_{n_{dt}}\boldsymbol{W}_{dts}\mathbf{1}_{n_{dt}}} = \frac{1}{\sigma_0^2}\left(\boldsymbol{W}_{dts} - \frac{\sigma_2^2 \boldsymbol{w}_{n_{dt}}\boldsymbol{w}'_{n_{dt}}}{\sigma_0^2\left(1+\frac{\sigma_2^2}{\sigma_0^2}w_{dt\cdot}\right)}\right)$$

$$= \frac{1}{\sigma_0^2}\left(\boldsymbol{W}_{dts} - \frac{\gamma_{dt}}{w_{dt\cdot}}\boldsymbol{w}_{n_{dt}}\boldsymbol{w}'_{n_{dt}}\right), \quad d=1,\dots,D,\ t=1,\dots,m_d.$$

If we define $\boldsymbol{B}_{dts} = \boldsymbol{W}_{dts} - \frac{\gamma_{dt}}{w_{dt\cdot}}\boldsymbol{w}_{n_{dt}}\boldsymbol{w}'_{n_{dt}}$, we have

$$\boldsymbol{R}_{ds}^{-1} = \frac{1}{\sigma_0^2}\underset{1\le t\le m_d}{\text{diag}}\,(\boldsymbol{B}_{dts}).$$

For calculating $\boldsymbol{V}_{ds}^{-1} = (\sigma_1^2 \mathbf{1}_{n_d}\mathbf{1}'_{n_d} + \boldsymbol{R}_{ds})^{-1}$, we put $\boldsymbol{A} = \boldsymbol{R}_{ds}$, $\boldsymbol{u} = \sigma_1^2\mathbf{1}_{n_d}$, $\boldsymbol{v}' = \mathbf{1}'_{n_d}$ in (12.6) and we obtain

$$\boldsymbol{V}_{ds}^{-1} = \frac{1}{\sigma_0^2}\underset{1\le t\le m_d}{\text{diag}}\,(\boldsymbol{B}_{dts}) \tag{12.9}$$
$$- \frac{\sigma_1^2}{\sigma_0^4}\frac{\underset{1\le t\le m_d}{\text{col}}\left[\boldsymbol{w}_{n_{dt}} - \frac{\gamma_{dt}}{w_{dt\cdot}}w_{dt\cdot}\boldsymbol{w}_{n_{dt}}\right]\underset{1\le t\le m_d}{\text{col}'}\left[\boldsymbol{w}'_{n_{dt}} - \frac{\gamma_{dt}}{w_{dt\cdot}}w_{dt\cdot}\boldsymbol{w}'_{n_{dt}}\right]}{1+\frac{\sigma_1^2}{\sigma_0^2}\left(\sum_{\ell=1}^{m_d} w_{d\ell\cdot} - \sum_{\ell=1}^{m_d}\gamma_{d\ell}w_{d\ell\cdot}\right)}$$
$$= \frac{1}{\sigma_0^2}\left[\underset{1\le t\le m_d}{\text{diag}}\,(\boldsymbol{B}_{dts}) - \varphi_d \underset{1\le t\le m_d}{\text{col}}\left[(1-\gamma_{dt})\boldsymbol{w}_{n_{dt}}\right]\underset{1\le t\le m_d}{\text{col}'}\left[(1-\gamma_{dt})\boldsymbol{w}'_{n_{dt}}\right]\right].$$

By applying (12.9), no matrix inversions are needed in programs. We also give an alternative formula for $\boldsymbol{V}_{ds}^{-1}$ requiring the inversion of a $m_d \times m_d$ matrix. Let us note that $\boldsymbol{V}_{ds} = \sigma_1^2 \mathbf{1}_{n_d}\mathbf{1}'_{n_d} + \boldsymbol{R}_{ds}$, where

$$\boldsymbol{R}_{ds} = \sigma_2^2 \underset{1\le t\le m_d}{\text{diag}}\,(\mathbf{1}_{n_{dt}})\boldsymbol{I}_{m_d}\underset{1\le t\le m_d}{\text{diag}}\,(\mathbf{1}'_{n_{dt}}) + \sigma_0^2 \boldsymbol{W}_{ds}^{-1}.$$

For calculating $\boldsymbol{R}_{ds}^{-1}$, we apply the formula (cf. Appendix A)

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

with $A = \sigma_0^2 \boldsymbol{W}_{ds}^{-1}$, $C = \sigma_2^2 \underset{1\le t\le m_d}{\text{diag}}\ (\mathbf{1}_{n_{dt}})$, $B = \boldsymbol{I}_{m_d}$, and $D = \underset{1\le t\le m_d}{\text{diag}}\ (\mathbf{1}'_{n_{dt}})$. We obtain

$$\boldsymbol{R}_{ds}^{-1} = \sigma_0^{-2}\boldsymbol{W}_{ds} - \sigma_0^{-2}\sigma_2^2 \boldsymbol{W}_{ds} \underset{1\le t\le m_d}{\text{diag}}\ (\mathbf{1}_{n_{dt}})$$
$$\cdot \left[\boldsymbol{I}_{m_d} + \sigma_0^{-2}\sigma_2^2 \underset{1\le t\le m_d}{\text{diag}}\ (\mathbf{1}'_{n_{dt}})\boldsymbol{W}_{ds} \underset{1\le t\le m_d}{\text{diag}}\ (\mathbf{1}_{n_{dt}})\right]^{-1} \sigma_0^{-2} \underset{1\le t\le m_d}{\text{diag}}\ (\mathbf{1}'_{n_{dt}})\boldsymbol{W}_{ds}.$$

For calculating $\boldsymbol{V}_{ds}^{-1}$, we use the formula (12.6) with $A = \boldsymbol{R}_{ds}$, $\boldsymbol{u} = \sigma_1^2 \mathbf{1}_{n_d}$, $\boldsymbol{v}' = \mathbf{1}'_{n_d}$. Finally, we obtain

$$\boldsymbol{V}_{ds}^{-1} = \boldsymbol{R}_{ds}^{-1} - \frac{\sigma_1^2}{1 + \sigma_1^2 \mathbf{1}'_{n_d} \boldsymbol{R}_{ds}^{-1} \mathbf{1}_{n_d}} \boldsymbol{R}_{ds}^{-1} \mathbf{1}_{n_d} \mathbf{1}'_{n_d} \boldsymbol{R}_{ds}^{-1}. \tag{12.10}$$

12.3 The Conditional Distribution of $\mathbf{y}_r$ given $\mathbf{y}_s$

Due to the normality assumptions of the population model (12.1), the vector $\boldsymbol{y} = (\boldsymbol{y}'_s, \boldsymbol{y}'_r)'$ is normally distributed with mean vector $\boldsymbol{\mu} = (\boldsymbol{\mu}'_s, \boldsymbol{\mu}'_r)'$ and covariance matrix

$$\boldsymbol{V} = \begin{pmatrix} \boldsymbol{V}_s & \boldsymbol{V}_{sr} \\ \boldsymbol{V}_{rs} & \boldsymbol{V}_r \end{pmatrix},$$

where $\boldsymbol{V}_s = \text{var}(\boldsymbol{y}_s)$, $\boldsymbol{V}_r = \text{var}(\boldsymbol{y}_r)$, $\boldsymbol{V}_{rs} = \text{cov}(\boldsymbol{y}_r, \boldsymbol{y}_s)$, and $\boldsymbol{V}_{sr} = \boldsymbol{V}'_{rs}$. Thus, the conditional distribution of $\boldsymbol{y}_r|\boldsymbol{y}_s$ is

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

where the conditional mean vector and covariance matrix are (see e.g. Theorem 2.2E in Rencher (1998))

$$\boldsymbol{\mu}_{r|s} = \boldsymbol{X}_r\boldsymbol{\beta} + \boldsymbol{V}_{rs}\boldsymbol{V}_s^{-1}(\boldsymbol{y}_s - \boldsymbol{X}_s\boldsymbol{\beta}), \quad \boldsymbol{V}_{r|s} = \boldsymbol{V}_r - \boldsymbol{V}_{rs}\boldsymbol{V}_s^{-1}\boldsymbol{V}_{sr}. \tag{12.11}$$

12.3.1 Conditional Mean Vector

This section derives a programmable formula for the conditional mean $\boldsymbol{\mu}_{r|s}$. We know that $\boldsymbol{V}_s^{-1} = \underset{1\le d\le D}{\text{diag}}\ (\boldsymbol{V}_{ds}^{-1})$, where $\boldsymbol{V}_{ds}^{-1}$ is given in (12.9) or alternatively in (12.10). Since the population vector $\boldsymbol{y}$ follows the model (12.1), it holds that

$$\boldsymbol{y} - E[\boldsymbol{y}] = (\boldsymbol{X\beta} + \boldsymbol{Z}_1\boldsymbol{u}_1 + \boldsymbol{Z}_2\boldsymbol{u}_2 + \boldsymbol{e}) - \boldsymbol{X\beta} = \boldsymbol{Z}_1\boldsymbol{u}_1 + \boldsymbol{Z}_2\boldsymbol{u}_2 + \boldsymbol{e}.$$

Therefore, the covariance $\boldsymbol{V}_{rs} = \operatorname{cov}(\boldsymbol{y}_r, \boldsymbol{y}_s)$ is

$$\begin{aligned}\boldsymbol{V}_{rs} &= E[(\boldsymbol{Z}_{1r}\boldsymbol{u}_1 + \boldsymbol{Z}_{2r}\boldsymbol{u}_{2r} + \boldsymbol{e}_r)(\boldsymbol{Z}_{1s}\boldsymbol{u}_1 + \boldsymbol{Z}_{2s}\boldsymbol{u}_{2s} + \boldsymbol{e}_s)'] \\ &= \boldsymbol{Z}_{1r}\sigma_1^2\boldsymbol{I}_D\boldsymbol{Z}_{1s}' + \boldsymbol{Z}_{2r}E[\boldsymbol{u}_{2r}\boldsymbol{u}_{2s}']\boldsymbol{Z}_{2s}'.\end{aligned}$$

Let us now calculate the $M \times m$ matrix $E[\boldsymbol{u}_{2r}\boldsymbol{u}_{2s}']$. As $E[u_{2,d_1t_1}u_{2,d_2t_2}] = 0$ if $d_1 \neq d_2$ or $t_1 \neq t_2$, we get

$$\begin{aligned}E[\boldsymbol{u}_{2r}\boldsymbol{u}_{2s}'] &= E\left[\underset{1\le d\le D}{\operatorname{col}}\big(\underset{1\le t\le M_d}{\operatorname{col}}(u_{2,dt})\big)\underset{1\le d\le D}{\operatorname{col}'}\big(\underset{1\le t\le m_d}{\operatorname{col}'}(u_{2,dt})\big)\right] \\ &= \underset{1\le d\le D}{\operatorname{diag}}\left(E\left[\operatorname{col}\Big\{\underset{1\le t\le m_d}{\operatorname{col}}(u_{2,dt}),\underset{m_d+1\le t\le M_d}{\operatorname{col}}(u_{2,dt})\Big\}\underset{1\le t\le m_d}{\operatorname{col}'}(u_{2,dt})\right]\right) \\ &= \sigma_2^2\underset{1\le d\le D}{\operatorname{diag}}\left(\operatorname{col}\big\{\boldsymbol{I}_{m_d}, \boldsymbol{0}_{M_d-m_d\times m_d}\big\}\right).\end{aligned}$$

As $\boldsymbol{Z}_{2r} = \underset{1\le d\le D}{\operatorname{diag}}\big(\operatorname{diag}\{\underset{1\le t\le m_d}{\operatorname{diag}}(\mathbf{1}_{N_{dt}-n_{dt}}), \underset{m_d+1\le t\le M_d}{\operatorname{diag}}(\mathbf{1}_{N_{dt}})\}\big)$, we get

$$\begin{aligned}&\boldsymbol{Z}_{2r}E[\boldsymbol{u}_{2r}\boldsymbol{u}_{2s}']\boldsymbol{Z}_{2s}' \\ &= \sigma_2^2\underset{1\le d\le D}{\operatorname{diag}}\left(\underset{1\le t\le M_d}{\operatorname{diag}}(\mathbf{1}_{N_{dt}-n_{dt}})\operatorname{col}\big\{\boldsymbol{I}_{m_d}, \boldsymbol{0}_{M_d-m_d\times m_d}\big\}\underset{1\le t\le m_d}{\operatorname{diag}}(\mathbf{1}_{n_{dt}}')\right) \\ &= \sigma_2^2\underset{1\le d\le D}{\operatorname{diag}}\left(\operatorname{col}\big\{\underset{1\le t\le m_d}{\operatorname{diag}}(\mathbf{1}_{N_{dt}-n_{dt}}), \boldsymbol{0}_{N_d-N_{ds}\times m_d}\big\}\underset{1\le t\le m_d}{\operatorname{diag}}(\mathbf{1}_{n_{dt}}')\right) \\ &= \sigma_2^2\underset{1\le d\le D}{\operatorname{diag}}\left(\operatorname{col}\big\{\underset{1\le t\le m_d}{\operatorname{diag}}(\mathbf{1}_{N_{dt}-n_{dt}}\mathbf{1}_{n_{dt}}'), \boldsymbol{0}_{N_d-N_{ds}\times n_d}\big\}\right),\end{aligned}$$

where $N_{ds} = \sum_{t=1}^{m_d} N_{dt}$. We have obtained that $\boldsymbol{V}_{rs} = \underset{1\le d\le D}{\operatorname{diag}}(\boldsymbol{V}_{drs})$, with

$$\boldsymbol{V}_{drs} = \sigma_1^2\mathbf{1}_{N_d-n_d}\mathbf{1}_{n_d}' + \operatorname{col}\left\{\sigma_2^2\underset{1\le t\le m_d}{\operatorname{diag}}(\mathbf{1}_{N_{dt}-n_{dt}}\mathbf{1}_{n_{dt}}'), \boldsymbol{0}_{(N_d-N_{ds})\times n_d}\right\} = A + B. \tag{12.12}$$

Moreover, formula (12.9) can be written in the form

$$\begin{aligned}\boldsymbol{V}_{ds}^{-1} &= \frac{1}{\sigma_0^2}\underset{1\le t\le m_d}{\operatorname{diag}}(\boldsymbol{B}_{dts}) \\ &\quad - \frac{1}{\sigma_0^2}\varphi_d\underset{1\le t\le m_d}{\operatorname{col}}\big[(1-\gamma_{dt})\boldsymbol{w}_{n_{dt}}\big]\underset{1\le t\le m_d}{\operatorname{col}'}\big[(1-\gamma_{dt})\boldsymbol{w}_{n_{dt}}'\big] = C - D,\end{aligned}$$

where $\boldsymbol{B}_{dts} = \boldsymbol{W}_{dts} - \frac{\gamma_{dt}}{w_{dt.}}\boldsymbol{w}_{n_{dt}}\boldsymbol{w}'_{n_{dt}}$. Then $\boldsymbol{V}_{rs}\boldsymbol{V}_s^{-1} = \underset{1\leq d\leq D}{\text{diag}}\,(\boldsymbol{V}_{drs}\boldsymbol{V}_{ds}^{-1})$, where

$$\boldsymbol{V}_{drs}\boldsymbol{V}_{ds}^{-1} = (A+B)(C-D) = (AC-AD)+(BC-BD).$$

The intermediate calculations are

$$\begin{aligned}
AC &= \sigma_1^2 \mathbf{1}_{N_d-n_d} \underset{1\leq t\leq m_d}{\text{col}'}(\mathbf{1}'_{n_{dt}})\frac{1}{\sigma_0^2}\underset{1\leq t\leq m_d}{\text{diag}}\Big(\boldsymbol{W}_{dts} - \frac{\gamma_{dt}}{w_{dt.}}\boldsymbol{w}_{n_{dt}}\boldsymbol{w}'_{n_{dt}}\Big)\\
&= \frac{\sigma_1^2}{\sigma_0^2}\mathbf{1}_{N_d-n_d}\underset{1\leq t\leq m_d}{\text{col}'}\Big[\mathbf{1}'_{n_{dt}}\Big(\boldsymbol{W}_{dts} - \frac{\gamma_{dt}}{w_{dt.}}\boldsymbol{w}_{n_{dt}}\boldsymbol{w}'_{n_{dt}}\Big)\Big]\\
&= \frac{\sigma_1^2}{\sigma_0^2}\mathbf{1}_{N_d-n_d}\underset{1\leq t\leq m_d}{\text{col}'}\Big[\boldsymbol{w}'_{n_{dt}} - \gamma_{dt}\boldsymbol{w}'_{n_{dt}}\Big] = \frac{\sigma_1^2}{\sigma_0^2}\mathbf{1}_{N_d-n_d}\underset{1\leq t\leq m_d}{\text{col}'}\Big[(1-\gamma_{dt})\boldsymbol{w}'_{n_{dt}}\Big],\\
AD &= \sigma_1^2\mathbf{1}_{N_d-n_d}\underset{1\leq t\leq m_d}{\text{col}'}(\mathbf{1}'_{n_{dt}})\frac{\varphi_d}{\sigma_0^2}\underset{1\leq t\leq m_d}{\text{col}}\Big[(1-\gamma_{dt})\boldsymbol{w}_{n_{dt}}\Big]\underset{1\leq t\leq m_d}{\text{col}'}\Big[(1-\gamma_{dt})\boldsymbol{w}'_{n_{dt}}\Big]\\
&= \frac{\sigma_1^2}{\sigma_0^2}\varphi_d\mathbf{1}_{N_d-n_d}\Big(\sum_{\ell=1}^{m_d}(1-\gamma_{d\ell})\mathbf{1}'_{n_{d\ell}}\boldsymbol{w}_{n_{d\ell}}\Big)\underset{1\leq t\leq m_d}{\text{col}'}\Big[(1-\gamma_{dt})\boldsymbol{w}'_{n_{dt}}\Big]\\
&= \frac{\sigma_1^2}{\sigma_0^2}\varphi_d\Big(\sum_{\ell=1}^{m_d} w_{d\ell.}(1-\gamma_{d\ell})\Big)\mathbf{1}_{N_d-n_d}\underset{1\leq t\leq m_d}{\text{col}'}\Big[(1-\gamma_{dt})\boldsymbol{w}'_{n_{dt}}\Big],
\end{aligned}$$

$$\begin{aligned}
BC &= \text{col}\Big\{\sigma_2^2\underset{1\leq t\leq m_d}{\text{diag}}(\mathbf{1}_{N_{dt}-n_{dt}}\mathbf{1}'_{n_{dt}}),\ \mathbf{0}_{(N_d-N_{ds})\times n_d}\Big\}\\
&\quad\cdot\frac{1}{\sigma_0^2}\underset{1\leq t\leq m_d}{\text{diag}}\Big(\boldsymbol{W}_{dts} - \frac{\gamma_{dt}}{w_{dt.}}\boldsymbol{w}_{n_{dt}}\boldsymbol{w}'_{n_{dt}}\Big)\\
&= \frac{\sigma_2^2}{\sigma_0^2}\text{col}\Big\{\underset{1\leq t\leq m_d}{\text{diag}}\Big(\mathbf{1}_{N_{dt}-n_{dt}}\boldsymbol{w}'_{n_{dt}} - \frac{\gamma_{dt}}{w_{dt.}}\mathbf{1}_{N_{dt}-n_{dt}}w_{dt.}\boldsymbol{w}'_{n_{dt}}\Big),\ \mathbf{0}_{(N_d-N_{ds})\times n_d}\Big\}\\
&= \frac{\sigma_2^2}{\sigma_0^2}\text{col}\Big\{\underset{1\leq t\leq m_d}{\text{diag}}\Big((1-\gamma_{dt})\mathbf{1}_{N_{dt}-n_{dt}}\boldsymbol{w}'_{n_{dt}}\Big),\ \mathbf{0}_{(N_d-N_{ds})\times n_d}\Big\},\\
BD &= \text{col}\Big\{\sigma_2^2\underset{1\leq t\leq m_d}{\text{diag}}(\mathbf{1}_{N_{dt}-n_{dt}}\mathbf{1}'_{n_{dt}}),\ \mathbf{0}_{(N_d-N_{ds})\times n_d}\Big\}\\
&\quad\cdot\frac{\varphi_d}{\sigma_0^2}\underset{1\leq t\leq m_d}{\text{col}}\Big[(1-\gamma_{dt})\boldsymbol{w}_{n_{dt}}\Big]\underset{1\leq t\leq m_d}{\text{col}'}\Big[(1-\gamma_{dt})\boldsymbol{w}'_{n_{dt}}\Big]\\
&= \frac{\varphi_d}{\sigma_0^2}\text{col}\Big\{\sigma_2^2\underset{1\leq t\leq m_d}{\text{diag}}(\mathbf{1}_{N_{dt}-n_{dt}}\mathbf{1}'_{n_{dt}})\underset{1\leq t\leq m_d}{\text{col}}[(1-\gamma_{dt})\boldsymbol{w}_{n_{dt}}],\ \mathbf{0}_{(N_d-N_{ds})\times 1}\Big\}
\end{aligned}$$

$$\cdot \operatorname*{col'}_{1\le t\le m_d}[(1-\gamma_{dt})\boldsymbol{w}'_{n_{dt}}] = \frac{\sigma_2^2\varphi_d}{\sigma_0^2}$$

$$\cdot \operatorname{col}\Big\{\operatorname*{col}_{1\le t\le m_d}[\mathbf{1}_{N_{dt}-n_{dt}}(1-\gamma_{dt})w_{dt.}],\ \mathbf{0}_{(N_d-N_{ds})\times 1}\Big\}\operatorname*{col'}_{1\le t\le m_d}[(1-\gamma_{dt})\boldsymbol{w}'_{n_{dt}}].$$

Note that

$$\frac{\sigma_1^2}{\sigma_0^2}\Big[1-\varphi_d\sum_{\ell=1}^{m_d} w_{d\ell.}(1-\gamma_{d\ell})\Big] = \frac{\sigma_1^2}{\sigma_0^2}\Big[1-\frac{\sigma_1^2\sum_{\ell=1}^{m_d} w_{d\ell.}(1-\gamma_{d\ell})}{\sigma_0^2+\sigma_1^2\sum_{\ell=1}^{m_d}(1-\gamma_{d\ell})w_{d\ell.}}\Big]$$
$$= \frac{\sigma_1^2}{\sigma_0^2+\sigma_1^2\sum_{\ell=1}^{m_d}(1-\gamma_{d\ell})w_{d\ell.}} = \varphi_d.$$

Therefore,

$$\begin{aligned}
AC-AD &= \frac{\sigma_1^2}{\sigma_0^2}\mathbf{1}_{N_d-n_d}\operatorname*{col'}_{1\le t\le m_d}\Big[(1-\gamma_{dt})\boldsymbol{w}'_{n_{dt}}\Big] \\
&\quad - \frac{\sigma_1^2}{\sigma_0^2}\varphi_d\Big(\sum_{\ell=1}^{m_d} w_{d\ell.}(1-\gamma_{d\ell})\Big)\mathbf{1}_{N_d-n_d}\operatorname*{col'}_{1\le t\le m_d}[(1-\gamma_{dt})\boldsymbol{w}'_{n_{dt}}] \\
&= \frac{\sigma_1^2}{\sigma_0^2}\Big[1-\varphi_d\sum_{\ell=1}^{m_d} w_{d\ell.}(1-\gamma_{d\ell})\Big]\mathbf{1}_{N_d-n_d}\operatorname*{col'}_{1\le t\le m_d}[(1-\gamma_{dt})\boldsymbol{w}'_{n_{dt}}] \\
&= \varphi_d\mathbf{1}_{N_d-n_d}\operatorname*{col'}_{1\le t\le m_d}[(1-\gamma_{dt})\boldsymbol{w}'_{n_{dt}}]
\end{aligned}$$

and

$$\begin{aligned}
\boldsymbol{V}_{drs}\boldsymbol{V}_{ds}^{-1} &= (AC-AD)+BC-BD = \varphi_d\,\mathbf{1}_{N_d-n_d}\operatorname*{col'}_{1\le t\le m_d}[(1-\gamma_{dt})\boldsymbol{w}'_{n_{dt}}] \\
&\quad + \frac{\sigma_2^2}{\sigma_0^2}\operatorname{col}\left\{\operatorname*{diag}_{1\le t\le m_d}((1-\gamma_{dt})\mathbf{1}_{N_{dt}-n_{dt}}\boldsymbol{w}'_{n_{dt}}),\ \mathbf{0}_{(N_d-N_{ds})\times n_d}\right\} \\
&\quad - \frac{\sigma_2^2\varphi_d}{\sigma_0^2}\operatorname{col}\left\{\operatorname*{col}_{1\le t\le m_d}[w_{dt.}(1-\gamma_{dt})\mathbf{1}_{N_{dt}-n_{dt}}],\ \mathbf{0}_{(N_d-N_{ds})\times 1}\right\} \\
&\quad \cdot \operatorname*{col'}_{1\le t\le m_d}[(1-\gamma_{dt})\boldsymbol{w}'_{n_{dt}}].
\end{aligned}$$

The conditional mean vector is $\boldsymbol{\mu}_{r|s} = \operatorname*{col}_{1\le d\le D}(\boldsymbol{\mu}_{dr|s})$, where

$$\begin{aligned}
\boldsymbol{\mu}_{dr|s} &= \boldsymbol{X}_{dr}\boldsymbol{\beta}+\boldsymbol{V}_{drs}\boldsymbol{V}_{ds}^{-1}(\boldsymbol{y}_{ds}-\boldsymbol{X}_{ds}\boldsymbol{\beta}) = \operatorname*{col}_{1\le t\le M_d}(\boldsymbol{\mu}_{dtr|s}) \\
&= \operatorname*{col}_{1\le t\le M_d}[\boldsymbol{X}_{dtr}\boldsymbol{\beta}]+\boldsymbol{V}_{drs}\boldsymbol{V}_{ds}^{-1}\operatorname*{col}_{1\le t\le m_d}[\boldsymbol{y}_{dts}-\boldsymbol{X}_{dts}\boldsymbol{\beta}]
\end{aligned}$$

and, correspondingly to the notation of this chapter, $\boldsymbol{y}_{dts} = \underset{1\le j\le n_{dt}}{\text{col}}(y_{dtj})$, $\boldsymbol{y}_{ds} = \underset{1\le t\le m_d}{\text{col}}(\boldsymbol{y}_{dts})$, $\boldsymbol{X}_{dts} = \underset{1\le j\le n_{dt}}{\text{col}}(\boldsymbol{x}_{dtj})$, $\boldsymbol{X}_{ds} = \underset{1\le t\le m_d}{\text{col}}(\boldsymbol{X}_{dts})$, $\boldsymbol{X}_{dtr} = \underset{n_{dt}+1\le j\le N_{dt}}{\text{col}}(\boldsymbol{x}_{dtj})$, and $\boldsymbol{X}_{dr} = \underset{1\le t\le M_d}{\text{col}}(\boldsymbol{X}_{dtr})$. By doing some algebra, we get

$$\begin{aligned}
\boldsymbol{\mu}_{dr|s} &= \underset{1\le t\le M_d}{\text{col}}[\boldsymbol{X}_{dtr}\boldsymbol{\beta}] + \varphi_d\,\boldsymbol{1}_{N_d-n_d}\sum_{\ell=1}^{m_d}(1-\gamma_{d\ell})\boldsymbol{w}'_{n_{d\ell}}(\boldsymbol{y}_{d\ell s}-\boldsymbol{X}_{d\ell s}\boldsymbol{\beta}) \\
&+ \frac{\sigma_2^2}{\sigma_0^2}\,\text{col}\left\{\underset{1\le t\le m_d}{\text{col}}\Big[(1-\gamma_{dt})\boldsymbol{1}_{N_{dt}-n_{dt}}\boldsymbol{w}'_{n_{dt}}(\boldsymbol{y}_{dts}-\boldsymbol{X}_{dts}\boldsymbol{\beta})\Big],\ \boldsymbol{0}_{(N_d-N_{ds})\times 1}\right\} \\
&- \frac{\sigma_2^2}{\sigma_0^2}\,\varphi_d\,\text{col}\left\{\underset{1\le t\le m_d}{\text{col}}[w_{dt.}(1-\gamma_{dt})\boldsymbol{1}_{N_{dt}-n_{dt}}],\ \boldsymbol{0}_{(N_d-N_{ds})\times 1}\right\} \\
&\cdot \sum_{\ell=1}^{m_d}(1-\gamma_{d\ell})\boldsymbol{w}'_{n_{d\ell}}(\boldsymbol{y}_{d\ell s}-\boldsymbol{X}_{d\ell s}\boldsymbol{\beta}).
\end{aligned}$$

Note that

$$1-\gamma_{dt} = 1-\frac{\sigma_2^2}{\sigma_0^2}\,w_{dt.}(1-\gamma_{dt}), \quad \gamma_{dt} = w_{dt.}(1-\gamma_{dt})\frac{\sigma_2^2}{\sigma_0^2}$$

and let us denote by $\bar{y}_{d\ell s} = w_{d\ell.}^{-1}\sum_{j\in s_{d\ell}} w_{d\ell j}y_{d\ell j}$ and $\bar{\boldsymbol{x}}_{d\ell s} = w_{d\ell.}^{-1}\sum_{j\in s_{d\ell}} w_{d\ell j}\boldsymbol{x}_{d\ell j}$ the weighted sample means of the response and auxiliary variables in subdomain ℓ from domain d. For $1\le t\le m_d$, the conditional mean vector is

$$\begin{aligned}
\boldsymbol{\mu}_{dtr|s} &= \boldsymbol{X}_{dtr}\boldsymbol{\beta} + \boldsymbol{1}_{N_{dt}-n_{dt}}\,\varphi_d\sum_{\ell=1}^{m_d}(1-\gamma_{d\ell})w_{d\ell.}(\bar{y}_{d\ell s}-\bar{\boldsymbol{x}}_{d\ell s}\boldsymbol{\beta}) \\
&+ \boldsymbol{1}_{N_{dt}-n_{dt}}\frac{\sigma_2^2}{\sigma_0^2}(1-\gamma_{dt})w_{dt.}(\bar{y}_{dts}-\bar{\boldsymbol{x}}_{dts}\boldsymbol{\beta}) \\
&- \boldsymbol{1}_{N_{dt}-n_{dt}}\frac{\sigma_2^2}{\sigma_0^2}\,\varphi_d w_{dt.}(1-\gamma_{dt})\sum_{\ell=1}^{m_d}(1-\gamma_{d\ell})w_{d\ell.}(\bar{y}_{d\ell s}-\bar{\boldsymbol{x}}_{d\ell s}\boldsymbol{\beta}) \\
&= \boldsymbol{X}_{dtr}\boldsymbol{\beta}+\boldsymbol{1}_{N_{dt}-n_{dt}}\varphi_d\sum_{\ell=1}^{m_d}(1-\gamma_{d\ell})w_{d\ell.}(\bar{y}_{d\ell s}-\bar{\boldsymbol{x}}_{d\ell s}\boldsymbol{\beta})\Big[1-\frac{\sigma_2^2}{\sigma_0^2}w_{dt.}(1-\gamma_{dt})\Big] \\
&+ \boldsymbol{1}_{N_{dt}-n_{dt}}\,w_{dt.}(1-\gamma_{dt})\frac{\sigma_2^2}{\sigma_0^2}(\bar{y}_{dts}-\bar{\boldsymbol{x}}_{dts}\boldsymbol{\beta}) \\
&= \boldsymbol{X}_{dtr}\boldsymbol{\beta} + \boldsymbol{1}_{N_{dt}-n_{dt}}\,w_{dt.}(1-\gamma_{dt})\frac{\sigma_2^2}{\sigma_0^2}
\end{aligned}$$

$$\cdot \left\{ \bar{y}_{dts} - \bar{\boldsymbol{x}}_{dts}\boldsymbol{\beta} + \frac{\varphi_d \sigma_0^2}{w_{dt.}\sigma_2^2} \sum_{\ell=1}^{m_d} (1-\gamma_{d\ell}) w_{d\ell.} (\bar{y}_{d\ell s} - \bar{\boldsymbol{x}}_{d\ell s}\boldsymbol{\beta}) \right\}$$

$$= \boldsymbol{X}_{dtr}\boldsymbol{\beta} + \mathbf{1}_{N_{dt}-n_{dt}} \gamma_{dt} \left\{ \bar{y}_{dts} - \bar{\boldsymbol{x}}_{dts}\boldsymbol{\beta} + \frac{\sigma_0^4}{\sigma_2^4} \frac{\varphi_d}{w_{dt.}} \sum_{\ell=1}^{m_d} \gamma_{d\ell} \left(\bar{y}_{d\ell s} - \bar{\boldsymbol{x}}_{d\ell s}\boldsymbol{\beta} \right) \right\}.$$

For $m_d + 1 \le t \le M_d$, the conditional mean vector is

$$\boldsymbol{\mu}_{dtr|s} = \boldsymbol{X}_{dtr}\boldsymbol{\beta} + \mathbf{1}_{N_{dt}} \varphi_d \sum_{\ell=1}^{m_d} (1-\gamma_{d\ell}) w_{d\ell.} (\bar{y}_{d\ell s} - \bar{\boldsymbol{x}}_{d\ell s}\boldsymbol{\beta})$$

$$= \boldsymbol{X}_{dtr}\boldsymbol{\beta} + \mathbf{1}_{N_{dt}} \frac{\sigma_0^2}{\sigma_2^2} \varphi_d \sum_{\ell=1}^{m_d} \gamma_{d\ell} \left(\bar{y}_{d\ell s} - \bar{\boldsymbol{x}}_{d\ell s}\boldsymbol{\beta} \right).$$

Finally, if $m_d = 0$ or equivalently if $n_d = 0$, the conditional mean vector is the marginal mean vector, i.e.

$$\boldsymbol{\mu}_{dtr|s} = \boldsymbol{\mu}_{dt} = \boldsymbol{X}_{dt}\boldsymbol{\beta}, \quad t = 1, \dots, M_d.$$

In summary, for $m_d > 0$ and $j \in r_{dt} = \{n_{dt}+1, \dots, N_{dt}\}$, we have

$$\mu_{dtj|s} = \boldsymbol{x}_{dtj}\boldsymbol{\beta} + \gamma_{dt} \left\{ \bar{y}_{dts} - \bar{\boldsymbol{x}}_{dts}\boldsymbol{\beta} + \frac{\sigma_0^4}{\sigma_2^4} \frac{\varphi_d}{w_{dt.}} \sum_{\ell=1}^{m_d} \gamma_{d\ell} (\bar{y}_{d\ell s} - \bar{\boldsymbol{x}}_{d\ell s}\boldsymbol{\beta}) \right\}, \; 1 \le t \le m_d,$$

$$\mu_{dtj|s} = \boldsymbol{x}_{dtj}\boldsymbol{\beta} + \frac{\sigma_0^2}{\sigma_2^2} \varphi_d \sum_{\ell=1}^{m_d} \gamma_{d\ell} \left(\bar{y}_{d\ell s} - \bar{\boldsymbol{x}}_{d\ell s}\boldsymbol{\beta} \right), \quad m_d + 1 \le t \le M_d.$$

For $m_d = 0$ and $j \in U_{dt}$, we have $\mu_{dtj|s} = \boldsymbol{x}_{dtj}\boldsymbol{\beta}$.

12.3.2 Conditional Covariance Matrix

This section derives a programmable formula for the conditional covariance matrix $\boldsymbol{V}_{r|s}$. By (12.11), since $\boldsymbol{V}_r$, $\boldsymbol{V}_{rs}$, and $\boldsymbol{V}_s^{-1}$ are all block-diagonal, it holds that $\boldsymbol{V}_{r|s} = \underset{1 \le d \le D}{\text{diag}} (\boldsymbol{V}_{dr|s})$, where

$$\boldsymbol{V}_{dr|s} = \boldsymbol{V}_{dr} - \boldsymbol{V}_{drs}\boldsymbol{V}_{ds}^{-1}\boldsymbol{V}_{dsr}.$$

Under the non-sample model (12.5), we have

$$\boldsymbol{V}_{dr} = \sigma_0^2 \boldsymbol{W}_{dr}^{-1} + \sigma_1^2 \mathbf{1}_{N_d-n_d}\mathbf{1}'_{N_d-n_d} + \sigma_2^2 \underset{1 \le t \le M_d}{\text{diag}} (\mathbf{1}_{N_{dt}-n_{dt}}\mathbf{1}'_{N_{dt}-n_{dt}}). \tag{12.13}$$

Moreover, using the expression of $\boldsymbol{V}_{drs}$ given in (12.12), we have

$$\begin{aligned}\boldsymbol{V}_{drs}\boldsymbol{V}_{ds}^{-1}\boldsymbol{V}_{dsr} &= (A+B)\boldsymbol{V}_{ds}^{-1}(A+B)' \\ &= A\boldsymbol{V}_{ds}^{-1}A' + A\boldsymbol{V}_{ds}^{-1}B' + B\boldsymbol{V}_{ds}^{-1}A' + B\boldsymbol{V}_{ds}^{-1}B' \\ &= \boldsymbol{L}_{d1} + \boldsymbol{L}_{d2} + \boldsymbol{L}_{d3} + \boldsymbol{L}_{d4}. \end{aligned} \tag{12.14}$$

The first term on the right hand side of (12.14) is

$$\begin{aligned}\boldsymbol{L}_{d1} &= A\boldsymbol{V}_{ds}^{-1}A = \sigma_1^4\mathbf{1}_{N_d-n_d}\mathbf{1}'_{n_d}\boldsymbol{V}_{ds}^{-1}\mathbf{1}_{n_d}\mathbf{1}'_{N_d-n_d} \\ &= \sigma_1^4(\mathbf{1}'_{n_d}\boldsymbol{V}_{ds}^{-1}\mathbf{1}_{n_d})\mathbf{1}_{N_d-n_d}\mathbf{1}'_{N_d-n_d},\end{aligned} \tag{12.15}$$

where

$$\begin{aligned}\mathbf{1}'_{n_d}\boldsymbol{V}_{ds}^{-1}\mathbf{1}_{n_d} &= \frac{1}{\sigma_0^2}\sum_{t=1}^{m_d}\mathbf{1}'_{n_{dt}}\boldsymbol{B}_{dts}\mathbf{1}_{n_{dt}} \\ &\quad - \frac{\varphi_d}{\sigma_0^2}\mathbf{1}'_{n_d}\underset{1\le t\le m_d}{\text{col}}((1-\gamma_{dt})\boldsymbol{w}_{n_{dt}})\underset{1\le t\le m_d}{\text{col}'}((1-\gamma_{dt})\boldsymbol{w}'_{n_{dt}})\mathbf{1}_{n_d}.\end{aligned} \tag{12.16}$$

Note that

$$\boldsymbol{B}_{dts} = \boldsymbol{W}_{dts} - \frac{\gamma_{dt}}{w_{dt.}}\boldsymbol{w}_{n_{dt}}\boldsymbol{w}'_{n_{dt}},$$

$$\underset{1\le t\le m_d}{\text{col}'}((1-\gamma_{dt})\boldsymbol{w}'_{n_{dt}})\mathbf{1}_{n_d} = \sum_{t=1}^{m_d} w_{dt.}(1-\gamma_{dt}), \tag{12.17}$$

$$\begin{aligned}\mathbf{1}'_{n_{dt}}\boldsymbol{B}_{dts}\mathbf{1}_{n_{dt}} &= \mathbf{1}'_{n_{dt}}\left(\boldsymbol{W}_{dts} - \frac{\gamma_{dt}}{w_{dt.}}\boldsymbol{w}_{n_{dt}}\boldsymbol{w}'_{n_{dt}}\right)\mathbf{1}_{n_{dt}} \\ &= w_{dt.} - \frac{\gamma_{dt}}{w_{dt.}}w_{dt.}^2 = w_{dt.}(1-\gamma_{dt}).\end{aligned} \tag{12.18}$$

Replacing (12.17) and (12.18) in (12.16), we get

$$\begin{aligned}\mathbf{1}'_{n_d}\boldsymbol{V}_{ds}^{-1}\mathbf{1}_{n_d} &= \frac{1}{\sigma_0^2}\left[\sum_{t=1}^{m_d} w_{dt.}(1-\gamma_{dt}) - \varphi_d\left\{\sum_{t=1}^{m_d} w_{dt.}(1-\gamma_{dt})\right\}^2\right] \\ &= \frac{1}{\sigma_0^2}\sum_{t=1}^{m_d} w_{dt.}(1-\gamma_{dt})\left\{1-\varphi_d\sum_{t=1}^{m_d} w_{dt.}(1-\gamma_{dt})\right\} \\ &= \frac{\varphi_d}{\sigma_1^2}\sum_{t=1}^{m_d} w_{dt.}(1-\gamma_{dt}).\end{aligned} \tag{12.19}$$

Replacing (12.19) in (12.15), we finally obtain

$$\boldsymbol{L}_{d1} = \sigma_1^2\,\varphi_d \sum_{t=1}^{m_d} w_{dt.}(1-\gamma_{dt})\mathbf{1}_{N_d-n_d}\mathbf{1}'_{N_d-n_d}.$$

The second term on the right hand side of (12.14) is

$$\begin{aligned}
\boldsymbol{L}_{d2} &= \sigma_1^2\mathbf{1}_{N_d-n_d}\mathbf{1}'_{n_d}\boldsymbol{V}_{ds}^{-1}\mathrm{col}'\Big\{\sigma_2^2 \underset{1\le t\le m_d}{\mathrm{diag}}\,(\mathbf{1}_{n_{dt}}\mathbf{1}'_{N_{dt}-n_{dt}}),\,\mathbf{0}_{n_d\times(N_d-N_{ds})}\Big\}\\
&= \sigma_1^2\sigma_2^2 \underset{1\le t\le M_d}{\mathrm{diag}}\,(\mathbf{1}_{N_{dt}-n_{dt}})\mathbf{1}_{M_d}\mathbf{1}'_{n_d}\boldsymbol{V}_{ds}^{-1}\mathrm{col}'\Big\{\underset{1\le t\le m_d}{\mathrm{diag}}\,(\mathbf{1}_{n_{dt}}),\,\mathbf{0}_{n_d\times(M_d-m_d)}\Big\}\\
&\cdot \underset{1\le t\le M_d}{\mathrm{diag}}\,(\mathbf{1}'_{N_{dt}-n_{dt}}) = \sigma_1^2\sigma_2^2 \underset{1\le t\le M_d}{\mathrm{diag}}\,(\mathbf{1}_{N_{dt}-n_{dt}})\mathbf{1}_{M_d}\underset{1\le t\le m_d}{\mathrm{col}'}\,(\mathbf{1}'_{n_{dt}})\\
&\cdot \frac{1}{\sigma_0^2}\Big[\underset{1\le t\le m_d}{\mathrm{diag}}\,(\boldsymbol{B}_{dts}) - \varphi_d \underset{1\le t\le m_d}{\mathrm{col}}\big[(1-\gamma_{dt})\boldsymbol{w}_{n_{dt}}\big]\underset{1\le t\le m_d}{\mathrm{col}'}\big[(1-\gamma_{dt})\boldsymbol{w}'_{n_{dt}}\big]\Big]\\
&\cdot \mathrm{col}'\Big\{\underset{1\le t\le m_d}{\mathrm{diag}}\,(\mathbf{1}_{n_{dt}}),\,\mathbf{0}_{n_d\times(M_d-m_d)}\Big\}\underset{1\le t\le M_d}{\mathrm{diag}}\,(\mathbf{1}'_{N_{dt}-n_{dt}})\\
&= \frac{\sigma_1^2\sigma_2^2}{\sigma_0^2}\underset{1\le t\le M_d}{\mathrm{diag}}\,(\mathbf{1}_{N_{dt}-n_{dt}})\mathbf{1}_{M_d}\Big[\mathrm{col}'\Big\{\underset{1\le t\le m_d}{\mathrm{col}'}\,(\mathbf{1}'_{n_{dt}}\boldsymbol{B}_{dts}\mathbf{1}_{n_{dt}}),\ \mathbf{0}_{1\times(M_d-m_d)}\Big\}\\
&- \mathrm{col}'\Big\{\varphi_d\Big(\sum_{\ell=1}^{m_d} w_{d\ell.}(1-\gamma_{d\ell})\Big)\underset{1\le t\le m_d}{\mathrm{col}'}\,(w_{dt.}(1-\gamma_{dt})),\ \mathbf{0}_{1\times(M_d-m_d)}\Big\}\Big]\\
&\cdot \underset{1\le t\le M_d}{\mathrm{diag}}\,(\mathbf{1}'_{N_{dt}-n_{dt}}).
\end{aligned}$$

Using (12.18), we obtain

$$\begin{aligned}
\boldsymbol{L}_{d2} = &\underset{1\le t\le M_d}{\mathrm{diag}}\,(\mathbf{1}_{N_{dt}-n_{dt}})\mathbf{1}_{M_d}\,\sigma_2^2\,\varphi_d\,\mathrm{col}'\Big\{\underset{1\le t\le m_d}{\mathrm{col}'}\,(w_{dt.}(1-\gamma_{dt})),\ \mathbf{0}_{1\times(M_d-m_d)}\Big\}\\
&\cdot \underset{1\le t\le M_d}{\mathrm{diag}}\,(\mathbf{1}'_{N_{dt}-n_{dt}}).
\end{aligned}$$

The third term satisfies $\boldsymbol{L}_{d3} = \boldsymbol{L}'_{d2}$. The last term on the right hand side of (12.14) is

$$\begin{aligned}
\boldsymbol{L}_{d4} = &\sigma_2^4\mathrm{col}\Big\{\underset{1\le t\le m_d}{\mathrm{diag}}\,(\mathbf{1}_{N_{dt}-n_{dt}}\mathbf{1}'_{n_{dt}}),\,\mathbf{0}_{(N_d-N_{ds})\times n_d}\Big\}\boldsymbol{V}_{ds}^{-1}\\
&\cdot \mathrm{col}'\Big\{\underset{1\le t\le m_d}{\mathrm{diag}}\,(\mathbf{1}_{n_{dt}}\mathbf{1}'_{N_{dt}-n_{dt}}),\,\mathbf{0}_{n_d\times(N_d-N_{ds})}\Big\}
\end{aligned}$$

$$= \frac{\sigma_2^4}{\sigma_0^2} \operatorname*{diag}_{1\le t\le M_d}(\mathbf{1}_{N_{dt}-n_{dt}}) \operatorname{diag}\Big\{ \operatorname*{diag}_{1\le t\le m_d}(\mathbf{1}'_{n_{dt}} \boldsymbol{B}_{dts} \mathbf{1}_{n_{dt}})$$
$$- \varphi_d \operatorname*{col}_{1\le t\le m_d}(w_{dt.}(1-\gamma_{dt})) \operatorname*{col'}_{1\le t\le m_d}(w_{dt.}(1-\gamma_{dt})),\ \mathbf{0}_{(M_d-m_d)\times(M_d-m_d)} \Big\}$$
$$\cdot \operatorname*{diag}_{1\le t\le M_d}(\mathbf{1}'_{N_{dt}-n_{dt}}).$$

Using (12.18) again, we obtain

$$\boldsymbol{L}_{d4} = \frac{\sigma_2^4}{\sigma_0^2} \operatorname*{diag}_{1\le t\le M_d}(\mathbf{1}_{N_{dt}-n_{dt}}) \operatorname{diag}\Big\{ \operatorname*{diag}_{1\le t\le m_d}(w_{dt.}(1-\gamma_{dt}))$$
$$- \varphi_d \operatorname*{col}_{1\le t\le m_d}(w_{dt.}(1-\gamma_{dt})) \operatorname*{col'}_{1\le t\le m_d}(w_{dt.}(1-\gamma_{dt})),\ \mathbf{0}_{(M_d-m_d)\times(M_d-m_d)} \Big\}$$
$$\cdot \operatorname*{diag}_{1\le t\le M_d}(\mathbf{1}'_{N_{dt}-n_{dt}}).$$

Summarizing, we have obtained

$$\boldsymbol{L}_{d1} = \sigma_1^2 \varphi_d \sum_{t=1}^{m_d} w_{dt.}(1-\gamma_{dt}) \mathbf{1}_{N_d-n_d} \mathbf{1}'_{N_d-n_d},$$
$$\boldsymbol{L}_{d2} = \operatorname*{diag}_{1\le t\le M_d}(\mathbf{1}_{N_{dt}-n_{dt}}) \boldsymbol{V}_{2d} \operatorname*{diag}_{1\le t\le M_d}(\mathbf{1}'_{N_{dt}-n_{dt}}), \quad \boldsymbol{L}_{d3} = \boldsymbol{L}'_{d2},$$
$$\boldsymbol{L}_{d4} = \operatorname*{diag}_{1\le t\le M_d}(\mathbf{1}_{N_{dt}-n_{dt}}) \boldsymbol{V}_{4d} \operatorname*{diag}_{1\le t\le M_d}(\mathbf{1}'_{N_{dt}-n_{dt}}),$$

where

$$\boldsymbol{V}_{2d} = \mathbf{1}_{M_d} \sigma_2^2 \varphi_d \operatorname{col}'\Big\{ \operatorname*{col'}_{1\le t\le m_d}(w_{dt.}(1-\gamma_{dt})),\ \mathbf{0}_{1\times(M_d-m_d)} \Big\},$$
$$\boldsymbol{V}_{4d} = \frac{\sigma_2^4}{\sigma_0^2} \operatorname{diag}\Big\{ \operatorname*{diag}_{1\le t\le m_d}(w_{dt.}(1-\gamma_{dt}))$$
$$- \varphi_d \operatorname*{col}_{1\le t\le m_d}(w_{dt.}(1-\gamma_{dt})) \operatorname*{col'}_{1\le t\le m_d}(w_{dt.}(1-\gamma_{dt})),\ \mathbf{0}_{(M_d-m_d)\times(M_d-m_d)} \Big\}.$$

Recalling (12.13), we obtain

$$\boldsymbol{V}_{dr|s} = \boldsymbol{V}_{dr} - \boldsymbol{V}_{drs} \boldsymbol{V}_{ds}^{-1} \boldsymbol{V}_{dsr} \tag{12.20}$$
$$= \sigma_0^2 \boldsymbol{W}_{dr}^{-1} + \sigma_1^2 \Big\{ 1 - \varphi_d \sum_{t=1}^{m_d} w_{dt.}(1-\gamma_{dt}) \Big\} \mathbf{1}_{N_d-n_d} \mathbf{1}'_{N_d-n_d}$$

$$+ \underset{1\leq t\leq M_d}{\text{diag}}(\mathbf{1}_{N_{dt}-n_{dt}})\left(\sigma_2^2 \boldsymbol{I}_{M_d} - \boldsymbol{V}_{2d} - \boldsymbol{V}'_{2d} - \boldsymbol{V}_{4d}\right)\underset{1\leq t\leq M_d}{\text{diag}}(\mathbf{1}'_{N_{dt}-n_{dt}})$$
$$= \sigma_0^2 \boldsymbol{W}_{dr}^{-1} + \sigma_0^2\,\varphi_d\,\mathbf{1}_{N_d-n_d}\mathbf{1}'_{N_d-n_d} + \underset{1\leq t\leq M_d}{\text{diag}}(\mathbf{1}_{N_{dt}-n_{dt}})\boldsymbol{T}_{dr|s}\underset{1\leq t\leq M_d}{\text{diag}}(\mathbf{1}'_{N_{dt}-n_{dt}}),$$

where

$$\boldsymbol{T}_{dr|s} = \sigma_2^2 \boldsymbol{I}_{M_d} - (\boldsymbol{V}_{2d} + \boldsymbol{V}'_{2d}) - \boldsymbol{V}_{4d}.$$

12.3.3 Conditional Variances

For any subdomain $t = 1, \ldots, M_d$ and unit $j = n_{dt}+1, \ldots, N_{dt}$, let $v_{dtj|s}$ be the corresponding diagonal element in the matrix $\boldsymbol{V}_{dr|s}$. The diagonal elements can be written in the form

$$v_{dtj|s} = \boldsymbol{a}'_{dtj}\boldsymbol{V}_{dr|s}\boldsymbol{a}_{dtj}, \quad \text{where } \boldsymbol{a}'_{dtj} = \underset{1\leq \ell\leq M_d}{\text{col}'}(\delta_{t\ell}\underset{n_{d\ell}+1\leq i\leq N_{d\ell}}{\text{col}'}(\delta_{ij}))$$

and δ_{ij} is the delta of Kronecker, i.e. $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$. Thus, $\boldsymbol{a}'_{dtj}$ is a $1 \times (N_d - n_d)$ vector with a 1 in the position (t, j) and with 0's in the remaining positions. Replacing the expression of $\boldsymbol{V}_{dr|s}$ given in (12.20), we have

$$v_{dtj|s} = \boldsymbol{a}'_{dtj}\Big[\sigma_0^2 \boldsymbol{W}_{dr}^{-1} + \sigma_0^2\,\varphi_d\,\mathbf{1}_{N_d-n_d}\mathbf{1}'_{N_d-n_d}$$
$$+ \underset{1\leq \ell\leq M_d}{\text{diag}}(\mathbf{1}_{N_{d\ell}-n_{d\ell}})\boldsymbol{T}_{dr|s}\underset{1\leq \ell\leq M_d}{\text{diag}}(\mathbf{1}'_{N_{d\ell}-n_{d\ell}})\Big]\boldsymbol{a}_{dtj}.$$

By defining $\boldsymbol{a}'_{dt} = \underset{1\leq \ell\leq M_d}{\text{col}'}(\delta_{t\ell})$, it holds

$$\boldsymbol{a}'_{dtj}\boldsymbol{W}_{dr}^{-1}\boldsymbol{a}_{dtj} = w_{dtj}^{-1}, \quad \boldsymbol{a}'_{dtj}\mathbf{1}_{N_d-n_d}\mathbf{1}'_{N_d-n_d}\boldsymbol{a}_{dtj} = 1,$$
$$\boldsymbol{a}'_{dtj}\underset{1\leq \ell\leq M_d}{\text{diag}}(\mathbf{1}_{N_{d\ell}-n_{d\ell}})\boldsymbol{T}_{dr|s}\underset{1\leq \ell\leq M_d}{\text{diag}}(\mathbf{1}'_{N_{d\ell}-n_{d\ell}})\boldsymbol{a}_{dtj} = \boldsymbol{a}'_{dt}\boldsymbol{T}_{dr|s}\boldsymbol{a}_{dt}.$$

Consequently, we get

$$v_{dtj|s} = \sigma_0^2(w_{dtj}^{-1} + \varphi_d) + \boldsymbol{a}'_{dt}\boldsymbol{T}_{dr|s}\boldsymbol{a}_{dt}. \qquad (12.21)$$

Further, we have

$$\boldsymbol{a}'_{dt}\boldsymbol{T}_{dr|s}\boldsymbol{a}_{dt} = \boldsymbol{a}'_{dt}\left(\sigma_2^2 \boldsymbol{I}_{M_d} - \boldsymbol{V}_{2d} - \boldsymbol{V}'_{2d} - \boldsymbol{V}_{4d}\right)\boldsymbol{a}_{dt}$$
$$= \sigma_2^2\,\boldsymbol{a}'_{dt}\boldsymbol{I}_{M_d}\boldsymbol{a}_{dt} - 2\boldsymbol{a}'_{dt}\boldsymbol{V}_{2d}\boldsymbol{a}_{dt} - \boldsymbol{a}'_{dt}\boldsymbol{V}_{4d}\boldsymbol{a}_{dt}. \qquad (12.22)$$

Moreover, for any $M_d \times M_d$ matrix $\boldsymbol{A}$, the product $\boldsymbol{a}'_{dt}\boldsymbol{A}\boldsymbol{a}_{dt}$ gives the t-th diagonal element of $\boldsymbol{A}$. Then $\boldsymbol{a}'_{dt}\boldsymbol{I}_{M_d}\boldsymbol{a}_{dt} = 1$ for any t and

$$\boldsymbol{a}'_{dt}\boldsymbol{V}_{2d}\boldsymbol{a}_{dt} = \boldsymbol{a}'_{dt}\boldsymbol{V}'_{2d}\boldsymbol{a}_{dt} = \boldsymbol{a}'_{dt}\boldsymbol{V}'_{4d}\boldsymbol{a}_{dt} = 0 \quad \text{for} \quad m_d + 1 \le t \le M_d.$$

On the other hand, by defining $b_{dt} = w_{dt.}(1 - \gamma_{dt})$, we get for $1 \le t \le m_d$

$$\begin{aligned}
\boldsymbol{a}'_{dt}\boldsymbol{V}_{2d}\boldsymbol{a}_{dt} &= \sigma_2^2\, \varphi_d\, \boldsymbol{a}'_{dt}\, \boldsymbol{1}_{M_d}\, \text{col}'\left\{\underset{1\le t\le m_d}{\text{col}'}(w_{dt.}(1-\gamma_{dt})),\ \boldsymbol{0}_{1\times(M_d-m_d)}\right\}\boldsymbol{a}_{dt} \\
&= \sigma_2^2\, \varphi_d\, b_{dt}, \qquad (12.23)
\end{aligned}$$

$$\begin{aligned}
\boldsymbol{a}'_{dt}\boldsymbol{V}_{4d}\boldsymbol{a}_{dt} &= \frac{\sigma_2^4}{\sigma_0^2}\, \boldsymbol{a}'_{dt}\text{diag}\Big\{\underset{1\le t\le m_d}{\text{diag}}(w_{dt.}(1-\gamma_{dt})) - \varphi_d \underset{1\le t\le m_d}{\text{col}}(w_{dt.}(1-\gamma_{dt})) \\
&\quad \cdot \underset{1\le t\le m_d}{\text{col}'}(w_{dt.}(1-\gamma_{dt})), \boldsymbol{0}_{(M_d-m_d)\times(M_d-m_d)}\Big\}\boldsymbol{a}_{dt} \\
&= \frac{\sigma_2^4}{\sigma_0^2}\left[w_{dt.}(1-\gamma_{dt}) - \varphi_d\, w_{dt.}^2(1-\gamma_{dt})^2\right] \\
&= \frac{\sigma_2^4}{\sigma_0^2}\, b_{dt}(1 - \varphi_d\, b_{dt}). \qquad (12.24)
\end{aligned}$$

Substituting (12.23) and (12.24) in (12.22), and recalling that $b_{dt} = w_{dt.}(1-\gamma_{dt}) = \gamma_{dt}\frac{\sigma_0^2}{\sigma_2^2}$, we obtain

$$\begin{aligned}
\boldsymbol{a}'_{dt}\boldsymbol{T}_{dr|s}\boldsymbol{a}_{dt} &= \sigma_2^2\left\{1 - 2\varphi_d\, b_{dt} - \frac{\sigma_2^2}{\sigma_0^2} b_{dt}(1 - \varphi_d\, b_{dt})\right\} \\
&= \sigma_2^2\left\{1 - 2\varphi_d\gamma_{dt}\frac{\sigma_0^2}{\sigma_2^2} - \gamma_{dt}\left(1 - \varphi_d\,\gamma_{dt}\frac{\sigma_0^2}{\sigma_2^2}\right)\right\} \\
&= \sigma_2^2 - 2\varphi_d\gamma_{dt}\sigma_0^2 - \gamma_{dt}\sigma_2^2 + \gamma_{dt}^2\varphi_d\sigma_0^2 \\
&= \sigma_0^2\varphi_d\gamma_{dt}(\gamma_{dt} - 2) + \sigma_2^2(1 - \gamma_{dt}).
\end{aligned}$$

Replacing this expression in (12.21), we finally obtain for $1 \le t \le m_d$ and $n_{dt}+1 \le j \le N_{dt}$ that

$$\begin{aligned}
v_{dtj|s} &= \sigma_0^2(w_{dtj}^{-1} + \varphi_d) + \sigma_0^2\varphi_d\gamma_{dt}(\gamma_{dt} - 2) + \sigma_2^2(1 - \gamma_{dt}) \\
&= \sigma_0^2\left[w_{dtj}^{-1} + \varphi_d\left\{1 + \gamma_{dt}(\gamma_{dt} - 2)\right\}\right] + \sigma_2^2(1 - \gamma_{dt}).
\end{aligned}$$

For $m_d + 1 \le t \le M_d$, we have

$$v_{dtj|s} = \sigma_0^2(w_{dtj}^{-1} + \varphi_d) + \sigma_2^2, \quad j = 1, \ldots, N_{dt}.$$

If $m_d = 0$ or equivalently if $n_d = 0$, the conditional variance is the marginal variance, i.e.

$$v_{dtj|s} = v_{dtj} = w_{dtj}^{-1}\sigma_0^2 + \sigma_1^2 + \sigma_2^2, \quad t = 1, \ldots, M_d, \; j = 1, \ldots, N_{dt}.$$

12.4 Monte Carlo EBP of an Additive Parameter

12.4.1 Introduction

The target of this section is to estimate a small area additive parameter of the form

$$\delta_d = \frac{1}{N_d}\sum_{t=1}^{M_d}\sum_{j=1}^{N_{dt}} h(y_{dtj}),$$

where h is a known measurable function. The best predictor of δ_d is

$$\begin{aligned}\hat{\delta}_d^{bp} &= E_{\boldsymbol{y}_r}\Big[\frac{1}{N_d}\sum_{t=1}^{M_d}\sum_{j=1}^{N_{dt}} h(y_{dtj})|\boldsymbol{y}_s\Big] \\ &= \frac{1}{N_d}\sum_{t=1}^{M_d}\Big\{\sum_{j\in s_{dt}} h(y_{dtj}) + \sum_{j\in U_{dt}-s_{dt}} E_{\boldsymbol{y}_r}\big[h(y_{dtj})|\boldsymbol{y}_s\big]\Big\}.\end{aligned}$$

The conditional distribution of $\boldsymbol{y}_r|\boldsymbol{y}_s$ depends on the vector $\boldsymbol{\theta} = (\boldsymbol{\beta}', \sigma_0^2, \sigma_1^2, \sigma_2^2)'$ of unknown model parameters, which must be estimated, that is,

$$E_{\boldsymbol{y}_r}\big[h(y_{dtj})|\boldsymbol{y}_s\big] = E_{\boldsymbol{y}_r}\big[h(y_{dtj})|\boldsymbol{y}_s; \boldsymbol{\theta}\big].$$

Let $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\beta}}', \hat{\sigma}_0^2, \hat{\sigma}_1^2, \hat{\sigma}_2^2)'$ be an estimator based on sample data $\boldsymbol{y}_s$. The result of replacing $\boldsymbol{\theta}$ by $\hat{\boldsymbol{\theta}}$ in the formula of the best predictor is called empirical best predictor, that is,

$$\hat{\delta}_d^{ebp} = \frac{1}{N_d}\sum_{t=1}^{M_d}\Big\{\sum_{j\in s_{dt}} h(y_{dtj}) + \sum_{j\in U_{dt}-s_{dt}} E_{\boldsymbol{y}_r}\big[h(y_{dtj})|\boldsymbol{y}_s; \hat{\boldsymbol{\theta}}\big]\Big\}.$$

For a general function $h(\cdot)$, the expected value above might be not tractable analytically. When this occurs, the following Monte Carlo procedure can be applied.

(a) Estimate the unknown parameter $\boldsymbol{\theta} = (\boldsymbol{\beta}', \sigma_0^2, \sigma_1^2, \sigma_2^2)'$ using sample data $\boldsymbol{y}_s$.

(b) Replacing $\boldsymbol{\theta} = (\boldsymbol{\beta}', \sigma_0^2, \sigma_1^2, \sigma_2^2)'$ by the estimate $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\beta}}', \hat{\sigma}_0^2, \hat{\sigma}_1^2, \hat{\sigma}_2^2)'$, obtained in (a), draw L copies of each non-sample variable y_{dtj}. That means, for $d = 1, \ldots, D$, $t = 1, \ldots, M_d$, $j \in U_{dt} - s_{dt}$, $\ell = 1, \ldots, L$, generate $y_{dtj}^{(\ell)} \sim N(\hat{\mu}_{dtj|s}, \hat{v}_{dtj|s})$ with $\hat{\mu}_{dtj|s}$ and $\hat{v}_{dtj|s}$ obtained by replacing $\boldsymbol{\theta}$ by $\hat{\boldsymbol{\theta}}$ in the formulas for $\mu_{dtj|s}$ and $v_{dtj|s}$, where

$$\mu_{dtj|s}=\boldsymbol{x}_{dtj}\boldsymbol{\beta}+\gamma_{dt}\left\{\bar{y}_{dts}-\bar{\boldsymbol{x}}_{dts}\boldsymbol{\beta}+\frac{\sigma_0^4}{\sigma_2^4}\frac{\varphi_d}{w_{dt.}}\sum_{\ell=1}^{m_d}\gamma_{d\ell}(\bar{y}_{d\ell s}-\bar{\boldsymbol{x}}_{d\ell s}\boldsymbol{\beta})\right\},\ 1\le t\le m_d,$$

$$\mu_{dtj|s} = \boldsymbol{x}_{dtj}\boldsymbol{\beta} + \frac{\sigma_0^2}{\sigma_2^2}\varphi_d\sum_{\ell=1}^{m_d}\gamma_{d\ell}\left(\bar{y}_{d\ell s} - \bar{\boldsymbol{x}}_{d\ell s}\boldsymbol{\beta}\right), \quad \text{if } m_d + 1 \le t \le M_d,$$

$$\mu_{dtj|s} = \boldsymbol{x}_{dtj}\boldsymbol{\beta}, \quad \text{if } m_d = 0,$$

$$v_{dtj|s} = \begin{cases} \sigma_0^2\left[w_{dtj}^{-1} + \varphi_d\{1+\gamma_{dt}(\gamma_{dt}-2)\}\right] + \sigma_2^2(1-\gamma_{dt}), & 1 \le t \le m_d, \\ \sigma_0^2(w_{dtj}^{-1} + \varphi_d) + \sigma_2^2, & m_d + 1 \le t \le M_d, \\ w_{dtj}^{-1}\sigma_0^2 + \sigma_1^2 + \sigma_2^2, & m_d = 0, \end{cases}$$

$\bar{y}_{d\ell s} = w_{d\ell.}^{-1}\sum_{j\in s_{d\ell}} w_{d\ell j}y_{d\ell j}$, $\bar{\boldsymbol{x}}_{d\ell s} = w_{d\ell.}^{-1}\sum_{j\in s_{d\ell}} w_{d\ell j}\boldsymbol{x}_{d\ell j}$ and

$$\varphi_d = \frac{\sigma_1^2}{\sigma_0^2 + \sigma_1^2\sum_{\ell=1}^{m_d}(1-\gamma_{d\ell})w_{d\ell.}}, \quad \gamma_{dt} = \frac{\sigma_2^2}{\sigma_2^2 + \sigma_0^2/w_{dt.}}.$$

(c) For $d = 1, \ldots, D$, $t = 1, \ldots, M_d$, $j \in U_{dt} - s_{dt}$, the Monte Carlo approximation of the expected value is

$$E_{\boldsymbol{y}_r}\left[h(y_{dtj})|\boldsymbol{y}_s; \hat{\boldsymbol{\theta}}\right] \approx \frac{1}{L}\sum_{\ell=1}^{L} h(y_{dtj}^{(\ell)})$$

and the Monte Carlo approximation of the EBP of the additive domain parameter δ_d is

$$\hat{\delta}_d^{ebp} \approx \frac{1}{N_d}\sum_{t=1}^{M_d}\left\{\sum_{j\in s_{dt}} h(y_{dtj}) + \sum_{j\in U_{dt}-s_{dt}}\left(\frac{1}{L}\sum_{\ell=1}^{L} h(y_{dtj}^{(\ell)})\right)\right\} = \frac{1}{L}\sum_{\ell=1}^{L}\delta_d^{(\ell)},$$

where

$$\delta_d^{(\ell)} = \frac{1}{N_d}\sum_{t=1}^{M_d}\Big\{\sum_{j\in s_{dt}} h(y_{dtj}) + \sum_{j\in U_{dt}-s_{dt}} h(y_{dtj}^{(\ell)})\Big\}.$$

If we were interested in estimating the subdomain additive parameter

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

the corresponding Monte Carlo approximation of the EBP is

$$\hat{\delta}_{dt}^{ebp} \approx \frac{1}{N_{dt}}\Big\{\sum_{j\in s_{dt}} h(y_{dtj}) + \sum_{j\in U_{dt}-s_{dt}}\Big(\frac{1}{L}\sum_{\ell=1}^{L} h(y_{dtj}^{(\ell)})\Big)\Big\} = \frac{1}{L}\sum_{\ell=1}^{L}\delta_{dt}^{(\ell)},$$

where

$$\delta_{dt}^{(\ell)} = \frac{1}{N_{dt}}\Big\{\sum_{j\in s_{dt}} h(y_{dtj}) + \sum_{j\in U_{dt}-s_{dt}} h(y_{dtj}^{(\ell)})\Big\}.$$

Remark 12.1 If the selected two-fold NER model contains continuous auxiliary variables and there is no available census file, it is possible to use a design-based approximation to $\delta_{dt}^{(\ell)}$ when $n_{dt} > 0$. This approximation is

$$\delta_{dt}^{(\ell)} = \frac{1}{N_{dt}}\Big\{\sum_{j\in s_{dt}} \big(h(y_{dtj}) - h(y_{dtj}^{(\ell)})\big) + \sum_{j\in s_{dt}} \omega_{dtj} h(y_{dtj}^{(\ell)})\Big\}, \qquad (12.25)$$

where ω_{dtj}'s are the calibrated sample weights and

$$y_{dtj}^{(\ell)} \sim N(\hat{\mu}_{dtj|s}, \hat{v}_{dtj|s}), \quad j \in s_{dt},\ t = 1,\dots,m_d,\ d = 1,\dots,D,\ \ell = 1,\dots,L,$$

with $\hat{\mu}_{dtj|s}$ and $\hat{v}_{dtj|s}$ obtained by replacing $\boldsymbol{\theta}$ by $\hat{\boldsymbol{\theta}}$ in the formulas for $\mu_{dtj|s}$ and $v_{dtj|s}$ given above. Let us note that the second sum in the formula (12.25) stands in fact for summation over the whole population, this is why the sum of $h(y_{dj}^{(\ell)})$ over the sample is subtracted. For more details concerning use of the described approximation we refer to Remark 10.2.

12.4.2 Auxiliary Variables with Finite Number of Values

In many practical cases the values of the auxiliary variables are not available for all the population units. If in addition some of the variables are continuous, the EBP method is not applicable. An important particular case, where this method is applicable, is when the model is homoscedastic and the number of values of the vector of auxiliary variables is finite. More concretely, suppose that the covariates are categorical such that $\boldsymbol{x}_{dtj} \in \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_K\}$ and that $w_{dtj} = 1$ for all d, t, j. Then, we can calculate $\delta_d^{(\ell)}$ as

$$\delta_d^{(\ell)} = \frac{1}{N_d}\Big\{\sum_{t=1}^{m_d}\sum_{k=1}^{K}\sum_{j=1}^{n_{dtk}} h(y_{dtkj}) + \sum_{t=1}^{M_d}\sum_{k=1}^{K}\sum_{j=n_{dtk}+1}^{N_{dtk}} h(y_{dtkj}^{(\ell)})\Big\}, \tag{12.26}$$

where $N_{dtk} = \#\{j \in U_{dt} : \boldsymbol{x}_{dtj} = \boldsymbol{x}_k\}$ is supposed to be available from external data sources (aggregated auxiliary information), $n_{dtk} = \#\{j \in s_{dt} : \boldsymbol{x}_{dtj} = \boldsymbol{x}_k\}$, $y_{dtkj}^{(\ell)} \sim N(\hat{\mu}_{dtk|s}, \hat{v}_{dt|s})$, $k = 1, \ldots, K$, $t = 1, \ldots, M_d$, $d = 1, \ldots, D$, $\ell = 1, \ldots, L$. Further,

$$\hat{\mu}_{dtk|s} = \begin{cases} \boldsymbol{x}_k\hat{\boldsymbol{\beta}} + \hat{\gamma}_{dt}\big\{\bar{y}_{dts} - \bar{\boldsymbol{x}}_{dts}\hat{\boldsymbol{\beta}} + \frac{\hat{\sigma}_0^4}{\hat{\sigma}_2^4}\frac{\hat{\varphi}_d}{n_{dt}}\sum_{\ell=1}^{m_d}\hat{\gamma}_{d\ell}(\bar{y}_{d\ell s} - \bar{\boldsymbol{x}}_{d\ell s}\hat{\boldsymbol{\beta}})\big\}, & 1 \le t \le m_d, \\ \boldsymbol{x}_k\hat{\boldsymbol{\beta}} + \frac{\hat{\sigma}_0^2}{\hat{\sigma}_2^2}\hat{\varphi}_d\sum_{\ell=1}^{m_d}\hat{\gamma}_{d\ell}(\bar{y}_{d\ell s} - \bar{\boldsymbol{x}}_{d\ell s}\hat{\boldsymbol{\beta}}), & m_d + 1 \le t \le M_d, \\ \boldsymbol{x}_k\hat{\boldsymbol{\beta}}, & m_d = 0, \end{cases}$$

$$\hat{v}_{dt|s} = \begin{cases} \hat{\sigma}_0^2\big[1 + \hat{\varphi}_d\{1 + \hat{\gamma}_{dt}(\hat{\gamma}_{dt} - 2)\}\big] + \hat{\sigma}_2^2(1 - \hat{\gamma}_{dt}), & 1 \le t \le m_d, \\ \hat{\sigma}_0^2(1 + \hat{\varphi}_d) + \hat{\sigma}_2^2, & m_d + 1 \le t \le M_d, \\ \hat{\sigma}_0^2 + \hat{\sigma}_1^2 + \hat{\sigma}_2^2, & m_d = 0, \end{cases}$$

where

$$\bar{y}_{dts} = \frac{1}{n_{dt}}\sum_{k=1}^{K}\sum_{j=1}^{n_{dtk}} y_{dtkj}, \qquad \bar{\boldsymbol{x}}_{dts} = \frac{1}{n_{dt}}\sum_{k=1}^{K} n_{dtk}\boldsymbol{x}_k$$

and

$$\hat{\varphi}_d = \frac{\hat{\sigma}_1^2}{\hat{\sigma}_0^2 + \hat{\sigma}_1^2\sum_{\ell=1}^{m_d}(1 - \hat{\gamma}_{d\ell})n_{d\ell}}, \quad \hat{\gamma}_{dt} = \frac{n_{dt}\hat{\sigma}_2^2}{n_{dt}\hat{\sigma}_2^2 + \hat{\sigma}_0^2}.$$

Similarly, we can calculate $\delta_{dt}^{(\ell)}$ as

$$\delta_{dt}^{(\ell)} = \frac{1}{N_{dt}}\Big\{\sum_{k=1}^{K}\sum_{j=1}^{n_{dtk}} h(y_{dtkj}) + \sum_{k=1}^{K}\sum_{j=n_{dtk}+1}^{N_{dtk}} h(y_{dtkj}^{(\ell)})\Big\}. \tag{12.27}$$

12.5 EBPs of Poverty Indicators

Let z_{dtj} be a welfare variable (e.g. income or expenditure) for individual j from subdomain t within domain d and let z be the poverty line. For a given power $\alpha \geq 0$, the FGT poverty indicator of order α (Foster et al. 1984) for subdomain t within domain d is defined as

$$F_{\alpha,dt} = \frac{1}{N_{dt}} \sum_{j=1}^{N_{dt}} F_{\alpha,dtj}, \quad F_{\alpha,dtj} = \left(\frac{z - z_{dtj}}{z}\right)^{\alpha} I(z_{dtj} < z), \tag{12.28}$$

where $I(z_{dtj} < z) = 1$ if $z_{dtj} < z$ and $I(z_{dtj} < z) = 0$ otherwise. For $\alpha = 0$, we obtain the subdomain poverty proportion, which measures the proportion of people in the subdomain whose welfare is below the poverty line z. For $\alpha = 1$, we obtain the subdomain poverty gap, measuring the degree of poverty of the people in that subdomain. These indicators are defined analogously for domains.

For estimating these indicators in domains or subdomains, we assume that a one-to-one transformation of the welfare variable for each unit, $y_{dtj} = T(z_{dtj})$, follows the two-fold nested error model (12.1). Then, using the inverse transformation $z_{dtj} = T^{-1}(y_{dtj})$, we can express $F_{\alpha,dtj}$ in terms of the model response variables y_{dtj} as

$$F_{\alpha,dtj} = \left(\frac{z - T^{-1}(y_{dtj})}{z}\right)^{\alpha} I(T^{-1}(y_{dtj}) < z) = h_{\alpha}(y_{dtj}), \tag{12.29}$$

which means that $F_{\alpha,dt}$ is an additive parameter. Therefore, the best predictor of $F_{\alpha,dt}$ is

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

where $\hat{F}_{\alpha,dtj}^{bp} = E_{\boldsymbol{y}_r}\left[F_{\alpha,dtj} | \boldsymbol{y}_s, \boldsymbol{\theta}\right]$. For $\alpha = 0, 1$ and for certain transformations T, the expectation $\hat{F}_{\alpha,dtj}^{bp}$ can be calculated analytically, avoiding Monte Carlo simulation.

12.5.1 Poverty Proportion

For the poverty proportion, $\alpha = 0$, we have $h_0(y_{dtj}) = I(T^{-1}(y_{dtj}) < z) = I\left(y_{dtj} < T(z)\right)$. If T is a nondecreasing monotonous function, we obtain

$$\begin{aligned}\hat{F}_{0,dtj}^{bp} &= E_{\boldsymbol{y}_r}\left[F_{0,dtj} | \boldsymbol{y}_s; \boldsymbol{\theta}\right] = E_{\boldsymbol{y}_r}\left[h_0(y_{dtj}) | \boldsymbol{y}_s; \boldsymbol{\theta}\right] = E_{\boldsymbol{y}_r}\left[I\left(y_{dtj} < T(z)\right) | \boldsymbol{y}_s; \boldsymbol{\theta}\right] \\ &= P_{\boldsymbol{y}_r}\left(y_{dtj} < T(z) | \boldsymbol{y}_s; \boldsymbol{\theta}\right) = P\left(N(0,1) < \alpha_{dtj}\right) = \Phi(\alpha_{dtj}),\end{aligned}$$

where

$$\alpha_{dtj} = \frac{T(z) - \mu_{dtj|s}}{v_{dtj|s}^{1/2}}$$

and Φ denotes the cumulative distribution function of a standard normal random variable.

12.5.2 Poverty Gap

For the poverty gap, $\alpha = 1$, we have $h_1(y_{dtj}) = \frac{z-T^{-1}(y_{dtj})}{z} I(T^{-1}(y_{dtj}) < z)$. In this section we assume that $y = T(z) = \log(z + c)$ or $z = T^{-1}(y) = e^y - c$. First, we obtain

$$\begin{aligned}
\hat{F}_{1,dtj}^{bp} &= E_{\boldsymbol{y}_r}\left[F_{1,dtj}|\boldsymbol{y}_s;\boldsymbol{\theta}\right] = E_{\boldsymbol{y}_r}\left[h_1(y_{dtj})|\boldsymbol{y}_s;\boldsymbol{\theta}\right] \\
&= E_{\boldsymbol{y}_r}\left[\frac{z - T^{-1}(y_{dtj})}{z} I\big(y_{dtj} < T(z)\big)\Big|\boldsymbol{y}_s;\boldsymbol{\theta}\right] \\
&= E_{\boldsymbol{y}_r}\left[I\big(y_{dtj} < T(z)\big)\Big|\boldsymbol{y}_s;\boldsymbol{\theta}\right] - \frac{1}{z}E_{\boldsymbol{y}_r}\left[T^{-1}(y_{dtj})I\big(y_{dtj} < T(z)\big)\Big|\boldsymbol{y}_s;\boldsymbol{\theta}\right] \\
&= S_1 - \frac{1}{z}S_2.
\end{aligned}$$

We have proved that the first summand is

$$S_1 = E_{\boldsymbol{y}_r}\left[I\big(y_{dtj} < T(z)\big)\Big|\boldsymbol{y}_s;\boldsymbol{\theta}\right] = \Phi(\alpha_{dtj}).$$

For calculating S_2, we simplify the notation, i.e. $y_{dtj} = y$, $\mu_{dtj|s} = \mu$, $v_{dtj|s}^{1/2} = \sigma$, and $\alpha_{dtj} = \alpha$ and in the integrals below we apply the following changes of variables:

$$x = \frac{y-\mu}{\sigma}, \quad y = \sigma x + \mu, \quad dy = \sigma dx, \quad y = T(z) \Leftrightarrow x = \frac{T(z)-\mu}{\sigma} = \alpha,$$

$$u = x - \sigma, \quad x = u + \sigma, \quad dx = du, \quad x = \alpha \Leftrightarrow u = \alpha - \sigma.$$

It holds that

$$\begin{aligned}
S_2 &= E_{\boldsymbol{y}_r}\left[T^{-1}(y_{dtj})I\big(y_{dtj} < T(z)\big)\Big|\boldsymbol{y}_s;\boldsymbol{\theta}\right] = \int_{-\infty}^{T(z)} T^{-1}(y) f_{N(\mu,\sigma^2)}(y)\, dy \\
&= \int_{-\infty}^{\alpha} T^{-1}(\sigma x + \mu) f_{N(\mu,\sigma^2)}(\sigma x + \mu)\, \sigma dx
\end{aligned}$$

$$
\begin{aligned}
&= \int_{-\infty}^{\alpha} \left(e^{\mu} \exp\{\sigma x\} - c\right) \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left\{-\frac{1}{2}x^2\right\} \sigma\, dx \\
&= -c\Phi(\alpha) + e^{\mu} \int_{-\infty}^{\alpha} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}(x^2 - 2\sigma x + \sigma^2)\right\} \exp\left\{\frac{1}{2}\sigma^2\right\} dx \\
&= -c\Phi(\alpha) + \exp\left\{\frac{1}{2}\sigma^2 + \mu\right\} \int_{-\infty}^{\alpha} f_{N(\sigma,1)}(x)\, dx \\
&= -c\Phi(\alpha) + \exp\left\{\frac{1}{2}\sigma^2 + \mu\right\} \int_{-\infty}^{\alpha-\sigma} f_{N(0,1)}(u)\, du.
\end{aligned}
$$

Therefore

$$
S_2 = -c\Phi(\alpha) + \exp\left\{\frac{1}{2}\sigma^2 + \mu\right\} \Phi(\alpha - \sigma)
$$

and

$$
\begin{aligned}
\hat{F}^{bp}_{1,dtj} &= \Phi(\alpha_{dtj}) - \frac{1}{z}\left(\exp\left\{\frac{1}{2} v_{dtj|s} + \mu_{dtj|s}\right\} \Phi(\alpha_{dtj} - v^{1/2}_{dtj|s}) - c\Phi(\alpha_{dtj})\right) \\
&= \frac{z+c}{z} \Phi(\alpha_{dtj}) - \frac{1}{z} \exp\left\{\frac{1}{2} v_{dtj|s} + \mu_{dtj|s}\right\} \Phi\left(\alpha_{dtj} - v^{1/2}_{dtj|s}\right).
\end{aligned}
$$

12.6 EBPs of Average Income Indicators

Let z_{dtj} be a welfare variable (e.g. income or expenditure) for individual j from subdomain t within domain d. The average income of subdomain t within domain d is

$$
\bar{z}_{dt} = \frac{1}{N_{dt}} \sum_{j=1}^{N_{dt}} z_{dtj}.
$$

This indicator is defined analogously for domains. For estimating the average incomes in domains or subdomains, we assume that a one-to-one transformation of the welfare variable for each unit, $y_{dtj} = T(z_{dtj})$, follows the two-fold nested error model (12.1). Then, using the inverse transformation we can express z_{dtj} in terms of the model response variables y_{dtj} as

$$
z_{dtj} = T^{-1}(y_{dtj}) = h(y_{dtj}),
$$

which means that $\overline{z}_{dt}$ is an additive parameter. Therefore, the best predictor of $\overline{z}_{dt}$ is

$$\hat{\overline{z}}_{dt}^{bp} = \frac{1}{N_{dt}} \left(\sum_{j \in s_{dt}} z_{dtj} + \sum_{j \in U_{dt}-s_{dt}} \hat{z}_{dtj}^{bp} \right),$$

where

$$\hat{z}_{dtj}^{bp} = E_{\boldsymbol{y}_r}\left[z_{dtj}|\boldsymbol{y}_s, \boldsymbol{\theta}\right] = E_{\boldsymbol{y}_r}\left[h(y_{dtj})|\boldsymbol{y}_s; \boldsymbol{\theta}\right] = E_{\boldsymbol{y}_r}[T^{-1}(y_{dtj})|\boldsymbol{y}_s; \boldsymbol{\theta}].$$

Let us now assume that $y = T(z) = \log(z + c)$ or $z = T^{-1}(y) = e^y - c$. For calculating $\hat{z}_{dtj}^{bp}$, we simplify the notation, i.e. $y_{dtj} = y$, $\mu_{dtj|s} = \mu$, $v_{dtj|s}^{1/2} = \sigma$ and in the integral below we do the following change of variables:

$$x = \frac{y-\mu}{\sigma}, \quad y = \sigma x + \mu, \quad dy = \sigma dx, \quad y = T(z) \Leftrightarrow x = \frac{T(z)-\mu}{\sigma} = \alpha.$$

It holds that

$$\begin{aligned}
\hat{z}_{dtj}^{bp} &= \int_{-\infty}^{\infty} T^{-1}(y) f_{N(\mu,\sigma^2)}(y)\, dy = \int_{-\infty}^{\infty} T^{-1}(\sigma x + \mu) f_{N(\mu,\sigma^2)}(\sigma x + \mu)\, \sigma dx \\
&= \int_{-\infty}^{\infty} \left(e^{\mu} \exp\{\sigma x\} - c\right) \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left\{-\frac{1}{2} x^2\right\} \sigma dx \\
&= -c + e^{\mu} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}(x^2 - 2\sigma x + \sigma^2)\right\} \exp\left\{\frac{1}{2}\sigma^2\right\} dx \\
&= -c + \exp\left\{\frac{1}{2}\sigma^2 + \mu\right\} \int_{-\infty}^{\infty} f_{N(\sigma,1)}(x)\, dx = -c + \exp\left\{\frac{1}{2}\sigma^2 + \mu\right\}.
\end{aligned}$$

Therefore

$$\hat{z}_{dtj}^{bp} = -c + \exp\left\{\frac{1}{2} v_{dtj|s} + \mu_{dtj|s}\right\}.$$

12.7 Parametric Bootstrap MSE Estimator

Analytical approximations to the MSE of empirical best predictors are difficult to derive in the case of complex parameters such as the FGT poverty measures. We therefore present a parametric bootstrap MSE estimator by following the bootstrap method for finite populations of González-Manteiga et al. (2008a). This bootstrap method can be readily applied to other complex parameters not necessarily of the additive form as the FGT measures. Steps for implementing this method are given now.

1. Fit the model (12.1) to sample data $(\boldsymbol{y}_s, \boldsymbol{X}_s)$ and obtain an estimator $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\beta}}', \hat{\sigma}_0^2, \hat{\sigma}_1^2, \hat{\sigma}_2^2)'$ of $\boldsymbol{\theta} = (\boldsymbol{\beta}', \sigma_0^2, \sigma_1^2, \sigma_2^2)'$.
2. Repeat B times $(b = 1, \ldots, B)$:
   (a) For $d = 1, \ldots, D$, $t = 1, \ldots, M_d$, $j = 1, \ldots, N_{dt}$, generate independently $u_{1,d}^* \sim N(0, \hat{\sigma}_1^2)$, $u_{2,dt}^* \sim N(0, \hat{\sigma}_2^2)$ and $e_{dtj}^* \sim N(0, \hat{\sigma}_0^2)$.
   (b) For $d = 1, \ldots, D$, $t = 1, \ldots, M_d$, $j = 1, \ldots, N_{dt}$, generate independently the bootstrap population

$$y_{dtj}^{*(b)} = \boldsymbol{x}_{dtj}\hat{\boldsymbol{\beta}} + u_{1,d}^* + u_{2,dt}^* + w_{dtj}^{-1/2} e_{dtj}^*$$

   and calculate the bootstrap population parameters

$$\delta_d^{*(b)} = \frac{1}{N_d}\sum_{t=1}^{M_d}\sum_{j=1}^{N_{dt}} h(y_{dtj}^{*(b)}), \quad \delta_{dt}^{*(b)} = \frac{1}{N_{dt}}\sum_{j=1}^{N_{dt}} h(y_{dtj}^{*(b)}).$$

   (c) From the bootstrap population generated in Step (b), take the sample with the same indices $s \subset U$ as the initial sample, and calculate the bootstrap EBPs, $\hat{\delta}_d^{ebp*(b)}$ and $\hat{\delta}_{dt}^{ebp*(b)}$, as described in Sect. 12.4 using the bootstrap sample data $\boldsymbol{y}_s^{*(b)}$ and the known values $\boldsymbol{x}_{dtj}$.

3. Output: the bootstrap estimators of $MSE(\hat{\delta}_d^{ebp})$ and $MSE(\hat{\delta}_{dt}^{ebp})$

$$mse^*(\hat{\delta}_d^{ebp}) = \frac{1}{B}\sum_{b=1}^{B}\left(\hat{\delta}_d^{ebp*(b)} - \delta_d^{*(b)}\right)^2,$$

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

Remark 12.2 The described bootstrap estimator is applicable if a census file fulfilling properties (A), (B), and (C), given in Remark 10.1, is available. Nevertheless, it can be easily modified to the cases that only a census file fulfilling property (A) is available or no census file is available but the auxiliary variables are categorical and the sizes of the population classes are known. This modification can be done in the same manner as described in Sect. 10.8 for the NER model.

12.8 R Codes for EBPs

This section gives R codes for fitting the NER2 model to the survey data file `LFS20.txt`. We employ the R package `lme4`. The domains are defined by the variable AREA and the subdomains are obtained by crossing this variable with the age groups. The age groups (ageG) are defined by ageG $= 1$ if AGE < 25, ageG $= 2$ if $25 \leq$ AGE < 54, and ageG $= 3$ if AGE ≥ 54. The first target parameters are the proportions of poor people by domains or subdomains. The second and the third target parameters are the domain and subdomain poverty gaps and average incomes, respectively. As auxiliary variables, we take the dichotomic variables defining the three categories of the variable EDUCATION (primary or less, secondary and superior). The categories are named edu1, edu2, and edu3, respectively. The following R code reads the unit-level data files, load some R packages and defines some variables.

```
if(!require(Matrix)){
  install.packages("Matrix")
  library(Matrix)
}
if(!require(lme4)){
  install.packages("lme4")
  library(lme4)
}
dat <- read.table("LFS20.txt", header=TRUE, sep = "\t", dec = ".")
z0 <- 36500                                  # poverty threshold
ns <- nrow(dat)                              # global sample size
poor <- as.numeric(dat$INCOME<z0)            # variable poor
gap <- (z0-dat$INCOME)*poor/z0               # variable gap
one <- rep(1,nrow(dat))                      # variable one
Ga <- cut(dat$AGE, breaks=c(0,25,54,max(dat$AGE)), labels=c(1,2,3),
          right=TRUE)
ageG <- as.numeric(Ga)                       # age group
edu2 <- as.numeric(dat$EDUCATION==2)         # secondary education
edu3 <- as.numeric(dat$EDUCATION==3)         # superior education
y <- log(dat$INCOME)                         # variable y=log(income)
```

The following code reads the auxiliary data file and renames some variables.

```
aux <- read.table("Ndsa20.txt", header=TRUE, sep = "\t", dec = ".")
# Sort aux by sex, age and area
aux <- aux[order(aux$sex, aux$age, aux$area), ]
```

We calculate sample sizes, counts, and means by subdomains.

```
# Sizes
ndt <- tapply(X=one, INDEX=list(dat$AREA,ageG), FUN=sum)
# Sample counts of edu3
ndtedu3 <- tapply(edu3, list(dat$AREA,ageG), sum)
# Sample counts of edu2
ndtedu2 <- tapply(edu2, list(dat$AREA,ageG), sum)
# Sample counts of edu1
ndtedu1 <- ndt - ndtedu3 - ndtedu2
# Sample counts of poor people
ndtpoor <- tapply(poor, list(dat$AREA,ageG), sum)
# Sample poverty proportion
mdtpoor <- ndtpoor/ndt
# Sample sum of gap variable
ndtgap <- tapply(gap, list(dat$AREA,ageG), sum)
# Sample poverty gap
mdtgap <- ndtgap/ndt
# Sample sum of log-income
ndty <- tapply(y, list(dat$AREA,ageG), sum)
# Sample log-income mean
mdty <- ndty/ndt
```

We calculate direct estimators of sizes, poverty proportions and gaps and income means by subdomains, by using `dir2` function described in Sect. 2.8.4.

```
dir.poor <- dir2(data=poor, w=dat$WEIGHT, domain=list(area=dat$AREA,
                 ageG=ageG))
dir.gap <- dir2(data=gap, w=dat$WEIGHT, domain=list(area=dat$AREA,
                ageG=ageG))
dir.income <- dir2(data=dat$INCOME, w=dat$WEIGHT, domain=list(area=dat$AREA,
                   ageG=ageG))
hatNdt <- dir.poor$Nd.hat    # sizes
dirp <- dir.poor$mean        # poverty proportions
dirg <- dir.gap$mean         # poverty gaps
diri <- dir.income$mean      # income means
```

We fit a two-fold nested error regression model (NER2) to the variable $y = \log(\text{INCOME})$ under the assumption $w_{dtj} = 1$ for all d, t, j. We apply R function `lmer` with the REML fitting method.

```
lmm <- lmer(formula=y ~ edu3 + edu2 + (1|AREA/ageG), data=dat, REML=TRUE)
summary(lmm)                             # summary of the fitting procedure
anova(lmm)                               # analysis of variance table
beta <- fixef(lmm)                       # regression parameters
bedu3 <- beta[1] + beta[2]               # beta for x1=1, x2=1, x3=0 (edu3)
bedu2 <- beta[1] + beta[3]               # beta for x1=1, x2=0, x3=1 (edu2)
bedu1 <- beta[1]                         # beta for x1=1, x2=0, x3=0 (edu0)
var <- as.data.frame(VarCorr(lmm))       # variance parameters
sigmau2 <- var$sdcor[1]                  # standard deviation of u_{2,dt}
sigmau1 <- var$sdcor[2]                  # standard deviation of u_{1,d}
sigmae <- var$sdcor[3]                   # residual standard deviation
ranef(lmm)                               # modes of the random effects
ypred <- fitted(lmm)                     # predictions
residuals <- resid(lmm)                  # residuals
p.values <- 2*pnorm(abs(coef(summary(lmm))[,3]), low=F)      # p values
```

Table 12.1 gives the estimates of the regression parameters. The standard deviations of $u_{1,d}$, $u_{2,dt}$, and e_{dtj} are $\sigma_1 = 0.02114$, $\sigma_2 = 0.02321$, and $\sigma_0 = 0.27282$.

Figure 12.1 (left) plots a dispersion graph of model residuals. The residuals are situated symmetrically around zero. Figure 12.1 (right) plots an histogram of the model residuals that shows some lack of normality.

We first calculate the means and the variances of unobserved values of the income variable conditioned to observed ones. The following R code calculates γ_{dt}, ϕ_d, and $v_{dtj|s}$.

```
# Calculation of gammadt, gammadt, by subdomains
gammadt <- sigmau2^2*ndt/(sigmau2^2*ndt+sigmae^2)
# Calculation of deltad
gammad <- apply((1-gammadt)*ndt, 1, sum)
# phid by domains
phid <- sigmau1^2/(sigmae^2+sigmau1^2*gammad)
# Calculation of the conditioned variances, vdt, by subdomains
vdt <- sigmae^2*(1+phid*(1+gammadt*(gammadt-2))) + sigmau2^2*(1-gammadt)
```

Table 12.1 Estimated parameters of NER2 model

Parameter	Estimate	Std. error	z-value	p-value
Intercept	10.5406	0.01475	714.8	0.00
edu3	0.4282	0.02454	17.4	0.00
edu2	0.2275	0.01888	12.0	0.00

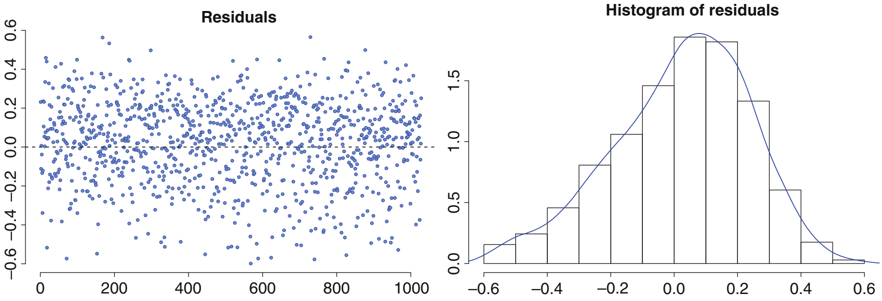

Fig. 12.1 Dispersion graph and histogram of residuals

The following R code calculates μ_{dtj} and α_{dtj}. We recall that for each area d and age group t, there are only three different values of μ_{dtj} corresponding to edu3, edu2, and edu1, respectively. The same happens for α_{dtj}.

```
# Preliminary calculations
mdtxbeta <- (bedu3*ndtedu3 + bedu2*ndtedu2 + bedu1*ndtedu1)/ndt
gammayxd <- apply(gammadt*(mdty-mdtxbeta), 1, sum)
uudt <- gammadt*(mdty-mdtxbeta+(sigmae^2/sigmau2^2)^2*(phid/ndt)*gammayxd)
# Calculation of the conditioned means
muedu3 <- bedu3 + uudt; muedu2 <- bedu2 + uudt; muedu1 <- bedu1 + uudt
# alphadt
y0 <- log(z0)
alphadedu3 <- vdt^(-1/2)*(y0-muedu3)
alphadedu2 <- vdt^(-1/2)*(y0-muedu2)
alphadedu1 <- vdt^(-1/2)*(y0-muedu1)
```

The following R code calculates $\Phi(\alpha_{dtj})$ and the EBPs of the poverty proportions by subdomains.

```
# Normal CDF
noredu3 <- pnorm(alphadedu3, mean=0, sd=1)
noredu2 <- pnorm(alphadedu2, mean=0, sd=1)
noredu1 <- pnorm(alphadedu1, mean=0, sd=1)
# Populations sizes by subdomains
Ndt <- tapply(X=aux$N, INDEX=list(aux$area,aux$age), FUN=sum)    # global
Nedu3 <- tapply(aux$edu3, list(aux$area,aux$age), sum)           # edu3
Nedu2 <- tapply(aux$edu2, list(aux$area,aux$age), sum)           # edu2
Nedu1 <- Ndt - Nedu3 - Nedu2                                     # edu1
# Poverty proportion EBPs by subdomains
ebpptot <- ndtpoor + (Nedu3-ndtedu3)*noredu3 + (Nedu2-ndtedu2)*noredu2 +
           (Nedu1-ndtedu1)*noredu1
ebp.poor <- ebpptot/Ndt; ebp.poor
```

The following R code calculates the poverty gaps by subdomains.

```
gap3 <- noredu3 - exp(vdt/2+muedu3)*pnorm(alphadedu3-vdt^{1/2})/z0
gap2 <- noredu2 - exp(vdt/2+muedu2)*pnorm(alphadedu2-vdt^{1/2})/z0
gap1 <- noredu1 - exp(vdt/2+muedu1)*pnorm(alphadedu1-vdt^{1/2})/z0
# Poverty gap EBPs by subdomains
ebpgtot <- ndtgap + (Nedu3-ndtedu3)*gap3 + (Nedu2-ndtedu2)*gap2 +
                    (Nedu1-ndtedu1)*gap1
ebp.gap <- ebpgtot/Ndt; ebp.gap
```

The following R code calculates the income means by subdomains.

```
inc3 <- exp(vdt/2+muedu3)
inc2 <- exp(vdt/2+muedu2)
```

```
inc1 <- exp(vdt/2+muedu1)
# EBPs of income means by subdomains
ebpitot <- ndty + (Nedu3-ndtedu3)*inc3 + (Nedu2-ndtedu2)*inc2 +
                  (Nedu1-ndtedu1)*inc1
ebpi <- ebpitot/Ndt; ebpi
```

Summary of results for poverty proportions

```
output1 <- data.frame(n1=ndt[,1], n2=ndt[,2], n3=ndt[,3],
            ebpp1=round(ebp.poor[,1],4), ebpp2=round(ebp.poor[,2],4),
            ebpp3=round(ebp.poor[,3],4), dirp1=round(subset(dir.poor,
            dom.ageG=="1")$mean,4), dirp2=round(subset(dir.poor,
            dom.ageG=="2")$mean,4), dirp3=round(subset(dir.poor,
            dom.ageG=="3")$mean,4))
head(output1, 10)
```

Summary of results for poverty gaps

```
output2 <- data.frame(n1=ndt[,1], n2=ndt[,2], n3=ndt[,3],
            ebpg1=round(ebp.gap[,1],4), ebpg2=round(ebp.gap[,2],4),
            ebpg3=round(ebp.gap[,3],4), dirg1=round(subset(dir.gap,
            dom.ageG=="1")$mean,4), dirg2=round(subset(dir.gap,
            dom.ageG=="2")$mean,4), dirg3=round(subset(dir.gap,
            dom.ageG=="3")$mean,4))
head(output2, 10)
```

Summary of results for income means

```
output3 <- data.frame(n1=ndt[,1], n2=ndt[,2], n3=ndt[,3],
            ebpi1=round(ebpi[,1],0), ebpi2=round(ebpi[,2],0),
            ebpi3=round(ebpi[,3],0), diri1=round(subset(dir.income,
            dom.ageG=="1")$mean,0), diri2=round(subset(dir.income,
            dom.ageG=="2")$mean,0), diri3=round(subset(dir.income,
            dom.ageG=="3")$mean,0))
head(output3, 10)
```

Table 12.2 (left) gives the sample sizes by subdomains (AREA crossed by ageG). We note that sample sizes are very small. The columns labeled by ebpp1, ebpp2, and ebpp3 contain the EBPs of poverty proportions by areas and age groups 1, 2, and 3, respectively. The columns labeled by dirp1, dirp2, and dirp3 contain the direct estimates of poverty proportions by areas and age groups 1, 2, and 3, respectively.

Table 12.3 gives the EBPs of poverty gap by subdomains. The columns labeled by ebpg1, ebpg2, and ebpg3 contain the EBPs of poverty gaps by areas and age groups 1, 2, and 3, respectively. The columns labeled by dirg1, dirg2, and dirg3

Table 12.2 EBPs and direct estimates of subdomain poverty proportions

d	n_1	n_2	n_3	ebpp1	ebpp2	ebpp3	dirp1	dirp2	dirp3
1	9	36	15	0.2813	0.2372	0.4019	0.4229	0.1098	0.2662
2	9	13	15	0.2737	0.2278	0.4192	0.2049	0.2576	0.3303
3	5	23	19	0.2341	0.1922	0.2846	0.0000	0.1330	0.1614
4	14	31	10	0.2627	0.2385	0.3659	0.3113	0.1481	0.2961
5	10	31	9	0.1628	0.2233	0.4273	0.1868	0.0900	0.3060
6	9	24	10	0.1369	0.2519	0.3670	0.0000	0.1699	0.1719
7	10	23	15	0.1834	0.2281	0.3848	0.0816	0.2272	0.2660
8	10	27	11	0.3040	0.3358	0.3513	0.1972	0.2774	0.2622
9	20	68	37	0.1688	0.2021	0.3246	0.1295	0.1169	0.1421
10	10	20	11	0.2163	0.2768	0.4002	0.0797	0.1889	0.6420

Table 12.3 EBPs and direct estimates of subdomain poverty gaps

d	n_1	n_2	n_3	ebpg1	ebpg2	ebpg3	dirg1	dirg2	dirg3
1	9	36	15	0.0459	0.0371	0.0704	0.1017	0.0120	0.0710
2	9	13	15	0.0434	0.0343	0.0732	0.0070	0.0205	0.0849
3	5	23	19	0.0362	0.0297	0.0462	0.0000	0.0238	0.0252
4	14	31	10	0.0426	0.0382	0.0623	0.0196	0.0301	0.0279
5	10	31	9	0.0227	0.0339	0.0745	0.0105	0.0109	0.0779
6	9	24	10	0.0175	0.0392	0.0620	0.0000	0.0273	0.0414
7	10	23	15	0.0256	0.0361	0.0658	0.0160	0.0421	0.0296
8	10	27	11	0.0507	0.0569	0.0603	0.0322	0.0633	0.0446
9	20	68	37	0.0234	0.0311	0.0541	0.0243	0.0141	0.0155
10	10	20	11	0.0322	0.0454	0.0701	0.0182	0.0440	0.1315

Table 12.4 EBPs and direct estimates of subdomain income means

d	n_1	n_2	n_3	ebpi1	ebpi2	ebpi3	diri1	diri2	diri3
1	9	36	15	45,690	47,671	40,832	42,479	49,380	40,651
2	9	13	15	45,450	47,216	40,073	45,717	48,783	40,990
3	5	23	19	47,249	50,873	45,895	50,648	51,913	50,338
4	14	31	10	46,994	48,262	42,302	43,178	49,605	42,803
5	10	31	9	50,928	47,672	39,560	53,155	47,561	42,339
6	9	24	10	51,667	46,360	41,912	53,973	46,726	44,807
7	10	23	15	48,929	48,586	41,576	49,804	48,159	44,327
8	10	27	11	44,948	43,477	43,236	44,286	42,933	43,187
9	20	68	37	49,995	50,100	44,219	49,650	50,432	44,069
10	10	20	11	47,694	46,111	40,910	46,483	47,648	37,116

contain the direct estimates of poverty gaps by areas and age groups 1, 2, and 3, respectively.

Table 12.4 gives the EBPs of income means by subdomains. The columns labeled by ebpi1, ebpi2, and ebpi3 contain the EBPs of income means by areas and age groups 1, 2, and 3, respectively. The columns labeled by diri1, diri2, and diri3 contain the direct estimates of income means by areas and age groups 1, 2, and 3, respectively.

In general, the EBP method produces estimates that are, across domains, smoother than direct estimates. This in an interesting property when dealing with real data applications for doing poverty mapping.

References

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

Foster, J., Greer, J., Thorbecke, E.: A class of decomposable poverty measures. Econometrica **52**, 761–766 (1984)

González-Manteiga, W., Lombardía, M.J., Molina, I., Morales, D., Santamaría, L.: Bootstrap mean squared error of small-area EBLUP. J. Stat. Comput. Simul. **78**, 443–462 (2008a)

Marhuenda, Y., Molina, I., Morales, D., Rao, J.N.K.: Poverty mapping in small areas under a two-fold nested error regression model. J. R. Stat. Soc. A **180**(4), 1111–1136 (2017)

Molina, I., Rao, J.N.K.: Small area estimation of poverty indicators. Can. J. Stat. **38**, 369–385 (2010)

Rencher, A.C.: Multivariate statistical inference and applications. Wiley, New York (1998)