Controlling the bias for M-quantile estimators for small area

Francesco Schirripa Spagnolo [a*], Gaia Bertarelli [b**], Raymond Chambers [c***], David Haziza [d****] and Nicola Salvati [e*]

[*]Dipartimento di Economia e Management, Università di Pisa, Pisa, Italy
[**]Istituto di Management, Scuola Universitaria Superiore Sant'Anna, Pisa, Italy
[***]National Institute for Applied Statistics Research Australia, University of Wollongong, Wollongong, Australia
[****]Department of mathematics and statistics, University of Ottawa, Ottawa, Canada

Abstract

In this paper we propose two bias correction approaches in order to reduce the prediction bias of the robust M-quantile predictors in small area estimation in the presence of representative outliers. A Monte-Carlo simulation study is conducted. Results confirm that our approaches improve the efficiency and reduce the prediction bias of M-quantile predictors when the population contains units that may be influential if selected in the sample.

Keywords— Robust methods, Small Area Estimation, M-quantile

1 Introduction

Outliers can arise frequently in sample surveys, for instance regarding economic variables whose distribution are highly skewed the data distribution is highly skewed. Some outliers are sample elements whose data values are recorded incorrectly or are unique, consequently they can be can be somehow corrected or removed. However, other outliers may not associated with an error: the

[a]francesco.schirripa@ec.unipi.it
[b]gaia.bertarelli@santannapisa.it
[c]ray@uow.edu.au
[d]dhaziza@uottawa.ca
[e]nicola.salvati@unipi.it

sample values associated with these outliers have been correctly recorded and they cannot be considered as unique. According to (Chambers, 1986) they are ‘representative outliers’. Such outliers values are representative of the non-sampled part of the population and they can seriously affect the survey estimates. Consequently, several methods have been developed in order to mitigate the effects of outliers on survey estimates. The representative outliers are even more concerning in the small area estimation (SAE) context, where sample sizes are very small and the estimation is often model-based Chambers et al. (2014). Robust small area estimation has received considerable attention in last years. Among other, Chambers and Tzavidis (2006) propose a robust approach based on the M-quantile regression aiming at overcoming the issue of outliers by avoiding the normal assumption. Sinha and Rao (2009) addressed the same issue from the perspective of linear mixed models. However, these approaches use plug-in robust prediction replacing parameter estimates in optimal but outlier-sensitive predictors by outlier robust versions and they may introduce a prediction biases. Dongmo-Jiongo et al. (2013) and Chambers et al. (2014) proposed a bias correction method for models with continuous response variables. The main aim of this work is to propose general bias correction methods to reduce the prediction bias of the robust M-quantile predictors in SAE in the presence of outliers. Two approaches are studied. The first estimator is a unified approach to M-quantile predictors based on a full bias correction and it could be viewed as a generalization of Chambers (1986). The second proposal is developed following the conditional bias approach by Beaumont et al. (2013) and Dongmo-Jiongo et al. (2013).

2 Bias corrected M-quantile-based estimator

Let θ_i be a finite population parameter for area i. That is, θ_i is a well-defined function of the values of a random variable Y associated with the N_i elements of such a small area finite population of interest. For ease of notation, we assume that both Y and θ_i are scalar, and we denote

$$\theta_i = f(\mathbf{y}_{U_i}),$$

where $\mathbf{y}_{U_i}$ denotes the vector of population values of Y for small area i and f is a known function. A basic sample survey inference problem is then one of predicting the value of θ_i give a sample of $n < N$ values from $\mathbf{y}_U$. Without loss of generality we put $\mathbf{y}_s$ equal to the population sub-vector defined by these values, where s denotes the set of sampled population units. We define (i) $\mathbf{y}_{U_i}$ vector of population values of Y for area i with $U = \bigcup_{i=1}^{m} U_i$ with m is the number of small areas; (ii) $\mathbf{y}_{s_i}$ vector of sampled population values in small area i with $s = \bigcup_{i=1}^{m} s_i$. Suppose that, given $\mathbf{y}_{s_i}$ we can impute the remaining values $\hat{\mathbf{y}}_{U_i}$ denote this imputed vector. A popular method of predicting the unobserved value of θ_i is via the Plug-In Predictor (PIP)

$$\hat{\theta}_i = f(\hat{\mathbf{y}}_{U_i}). \quad (1)$$

Adopting a model-based approach, the empirical PIP for θ_i based on this plug-in approximation is

$$\hat{\theta}_i = f(\mathbf{y}_{s_i}, \{\hat{y}_{ij}^{opt}; j \in U_i - s_i\}) \quad (2)$$

where the set $U_i - s_i$ contains the $N_i - n_i$ indices of the non-sampled units, $\hat{y}_{ij}^{opt} = E[y_{ij}|\mathbf{y}_s; \boldsymbol{\delta} = \hat{\boldsymbol{\delta}}]$ is the plug-in approximation of the minimum mean squared error predictor (MMSEP) of y_{ij}^{opt} for a non-sampled population unit j for area i, and $\boldsymbol{\delta}$ is a vector of unkown parameters. The above PIP (2) for small area can be also computed using the M-quantile approach. It can be obtained by using the estimated regression coefficients by M-quantile approach, $\hat{\boldsymbol{\beta}}_\tau$, leading to

$$\hat{\theta}_i^{MQ} = f\Big(\mathbf{y}_{s_i}, \{g^{-1}(\mathbf{x}_{ij}^T\hat{\boldsymbol{\beta}}_{\tau_i}); j \in U_i - s_i\}\Big), \quad (3)$$

where τ_i represents the order of M-quantile for area i. Its computation varies depending on the type of the data.

We propose two small area estimators based on Generalised version of M-quantile regression models.

The first estimator is a unified approach to M-quantile predictors based on a full bias correction. Following Chambers (1986), the first order approximation to the prediction bias of $\hat{\theta}_i^{MQ}$ is

$$E[\hat{\theta}_i^{MQ} - \theta_i] \simeq \sum_{j \notin s_i} \Big(\frac{\partial f}{\partial y_{ij}}\Big)_{\mathbf{y}_{U_i} = \mathbf{m}_{U_i}} E[\hat{y}_{ij} - y_{ij}] \simeq \sum_{i \in r_j} \Big(\frac{\partial f}{\partial y_{ij}}\Big)_{\mathbf{y}_U = \hat{\mathbf{m}}_{U\bar{q}_j}} \Big(\frac{\partial g^{-1}}{\partial \eta}\Big)_{\eta = \mathbf{x}_{ij}^T\hat{\boldsymbol{\beta}}_{\bar{q}_j}} \mathbf{x}_{ij}^T E[\hat{\beta}_q - \beta_q],$$

The bias corrected robust predictor MQC for the population average of Y in the ith area will be:

$$\theta_i^{MQC} = N_i^{-1} \left(\sum_{j \in s_i} y_{ij} + \sum_{j \in r_i} \hat{\mu}_{ij} + \sum_{j \in r_i} \Big(\frac{\partial f}{\partial y_{ij}}\Big)_{\mathbf{y}_U = \hat{\mathbf{m}}_{U\bar{q}_j}} \Big(\frac{\partial g^{-1}}{\partial \eta}\Big)_{\eta = \mathbf{x}_{ij}^T\hat{\boldsymbol{\beta}}_{\bar{q}_j}} \mathbf{x}_{ij}^T \hat{\mathbf{B}}_i \right) \quad (4)$$

where $d_{jh\bar{q}_j} = 2\{\bar{q}_j I(r_{hj} > 0) + (1 - \bar{q}_j) I(r_{hj} \leq 0)\}$ and $\hat{\mathbf{B}}_i$ has to be computed depending of the type of the response variable. If y_{ij} is continuous

$$\hat{\mathbf{B}}_i = \left(\sum_{h=1}^m \sum_{j \in s_h} \mathbf{x}_{hj} \hat{d}_{hj} \mathbf{x}_{hj}^T \right)^{-1} \sum_{h=1}^m \sum_{j \in s_h} \mathbf{x}_{hj} \hat{d}_{hj} \hat{\sigma}_{hj} \phi \left\{ \frac{y_{hj} - \mathbf{x}_{ij}^T \hat{\boldsymbol{\beta}}_{\tau_i}}{\hat{\sigma}_{hj}} \right\}. \quad (5)$$

The second proposal is developed following the conditional bias approach by Beaumont et al. (2013) and Dongmo-Jiongo et al. (2013). In a model based approach, the conditional bias attached to unit ij is

$$B_{ij} = E[\hat{\theta} - \theta | s; Y_{ij} = y_{ij}].$$

The prediction error $\hat{\theta}_i - \theta_i$ can be approximated as:

$$\hat{\theta}_i - \theta_i \simeq \sum_{j \in r_i} B_{ij}(I_{ij} = 0) + \sum_{j \in s_i} B_{ij}(I_{ij} = 1). \quad (6)$$

To determine the conditional bias, we need to distinguish two cases, whether the unit belongs to the sample or not. The main problem is that the conditional bias of a non-sampled unit can't be estimated since it depends on the Y-values on the non-sample units, which are not observed. A robust predictor of the mean in the ith area can be expressed as

$$N_i^{-1} \left(\sum_{j \in s_i} y_{ij} + \sum_{j \in r_i} g^{-1}(\mathbf{x}_{ij}^T \boldsymbol{\beta}) - \sum_{j \in s_i} B_{ij}(I_{ij} = 1) + \phi \left\{ \sum_{j \in s_i} B_{ij}(I_{ij} = 1) \right\} \right)$$

where ϕ is the Huber function. Translating the idea for MQ we have:

$$\theta_i^{MQD} = N_i^{-1} \left(\sum_{j \in s_i} y_{ij} + \sum_{j \in r_i} g^{-1}(\mathbf{x}_{ij}^T \boldsymbol{\beta}_{\bar{q}_j}) - \sum_{h=1}^{m} \sum_{j \in s_h} \hat{B}_{jh}(I_{jh} = 1) + \phi \left\{ \sum_{h=1}^{m} \sum_{j \in s_h} \hat{B}_{jh}(I_{jh} = 1) \right\} \right). \quad (7)$$

The ϕ-function in MQD depends on a tuning constant c. Using min-max method to compute the optimal tuning constant we obtain

$$\theta_i^{MQD} = N_i^{-1} \left(\sum_{j \in s_i} y_{ij} + \sum_{j \in r_i} g^{-1}(\mathbf{x}_{ij}^T \boldsymbol{\beta}_{\bar{q}_j}) - \frac{1}{2} \left(min \left\{ B_{jh}(I_{jh} = 1) \right\} + max \left\{ B_{jh}(I_{jh} = 1) \right\} \right) \right) \quad (8)$$

where the conditional bias for unit j has to be computed depending of the type of the response variable. If y_{ij} is a continuous

$$\hat{B}_{hj}(I_{hj} = 1) = \sum_{i \notin s_i} \mathbf{x}_{ij}^T \left\{ \sum_{h=1}^{m} \sum_{j \in s_h} \mathbf{x}_{hj} \hat{d}_{hj} \mathbf{x}_{hj}^T \right\}^{-1} \hat{d}_{hj} \mathbf{x}_{hj} (y_{hj} - \mathbf{x}_{hj}^T \hat{\boldsymbol{\beta}}_{\tau_i}). \quad (9)$$

3 Model-based simulations

In this section, we provide results regarding model-based simulation scenarios for continuous variables. We use a simulation setup based on Chambers et al. (2014). We consider the following outcome model for generating the finite population for $m = 40$ small areas:

$$y_{ij} = 100 + 5x_{ij} + u_i + \epsilon_{ij},$$

where i refers to the areas and j to the population units. Values for x are generated as i.i.d. from a lognormal distribution with a mean of 1 and a standard deviation of 0.5 on the log scale. The area and individual random effects are independently generated according to the following scenarios:

a) [0,0,0] - no outliers, $u \sim N(0, 3)$ and $e \sim N(0, 6)$;

b) [e,u,0] - outliers in area (fixed) and individual effects, $u \sim N(0, 3)$ for areas 1–36, $u \sim N(9, 20)$ for areas 37–40 and $e \sim \delta N(0, 6) + (1 - \delta)N(20, 150)$.

The sample data are selected by a simple random sampling without replacement within each area. The population and sample size are the same for all areas and are fixed at $N_i = 100$ and $n_i = 5$.

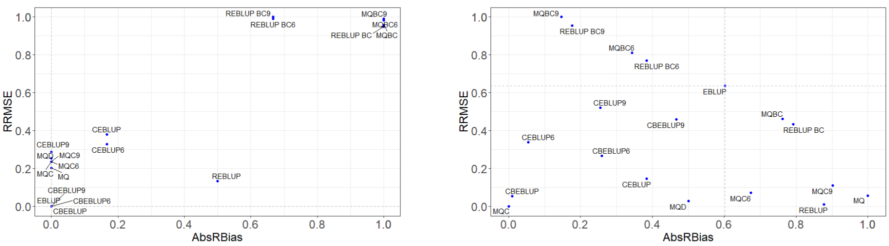
Each scenario is independently simulated 1000 times. The parameter of interest is the population mean in each small area. Nine different estimators are used for this purpose: the M-quantile estimator MQ by (Chambers and Tzavidis, 2006) which serves as a reference for the MQ regression based estimators, the bias corrected M-quantile estimator MQBC by (Chambers et al., 2014), the M-quantile estimator based on full bias correction MQC (see equation (4)), the M-quantile estimator based on conditional bias correction MQD (see equation (8)),the standard EBLUP which serves as a reference for all the considered estimators, the robust eblup REBPLUP by (Sinha and Rao, 2009) and its robust bias corrected version REBLUP–BC by (Chambers et al., 2014), the CBEBLUP and CEBLUP predictorS by (Dongmo-Jiongo et al., 2013). The influence function ϕ that is used in MQBC, MQC, REBLUP BC, CBEBLUP and CEBLUP is a Huber proposal 2 type. For each estimator, we test three different tuning constant for the bias correction part equal to 3, 6 and 9. The performance of the proposed indicators is evaluated according to min-max plots (Figure 1). The values on the x-axis and $y-$axis on plots are:

$$AbsRBias = \frac{\text{Median}[\text{AbsB}(\theta_{ki})] - \min\{\text{Median}[\text{AbsB}(\Theta_i)]\}}{\max\{\text{Median}[\text{AbsB}(\Theta_i)]\} - \min\{\text{Median}[\text{AbsB}(\Theta_i)]\}}$$

and

$$RRMSE = \frac{\text{Median}[\text{RRMSE}(\theta_{ki})] - \min\{\text{Median}[\text{RRMSE}(\Theta_i)]\}}{\max\{\text{Median}[\text{RRMSE}(\Theta_i)]\} - \min\{\text{Median}[\text{RRMSE}(\Theta_i)]\}},$$

where θ_{ki} is the kth estimator in the ith area and Θ_i is the vector all K predictors in area i.



(a) (0,0,0) (b) (e,u,0)

Figure 1: Min-Max plots for MQ, MQBC, MQC, MQD, EBLUP, REBLUP, REBLUP BC, CBEBLUP and CEBLUP under selected simulation scenarios.

Results confirm our expectations regarding the behaviour of the MQC and MQD estimators. With respect to MQ estimator, the new proposed estimators reduce the bias in the presence of outliers.

References

Beaumont, J., Haziza, D., and Ruiz-Gazen, A. (2013). A unified approach to robust estimation in finite population sampling. *Biometrika*, 100:555–569.

Chambers, R., Chandra, H., Salvati, N., and Tzavidis, N. (2014). Outlier robust small area estimation. *Journal of the Royal Statistical Society: Series B*, 76 (1):47–69.

Chambers, R. and Tzavidis, N. (2006). M-quantile models for small area estimation. *Biometrika*, 93 (2):255–268.

Chambers, R. L. (1986). Outlier robust finite population estimation. *Journal of the American Statistical Association*, 81(396):1063–1069.

Dongmo-Jiongo, V., Haziza, D., and Duchesne, P. (2013). Controlling the bias of robust small area estimators. *Biometrika*, 100:843–858.

Sinha, S. K. and Rao, J. N. K. (2009). Robust small area estimation. *The Canadian Journal of Statistics*, 37 (3):381–399.

Best Prediction of Missing Area-Level Direct Estimates via Multivariate Modelling

Anna-Lena Wölwer [a*], Jan Pablo Burgard [b*], Domingo Morales [c**] and Ralf Münnich [d*]

[*]Department of Economic and Social Statistics, Trier University, Germany
[**]Operations Research Center, University Miguel Hernández de Elche, Spain

Abstract

In the last years one could see increasing methodological research and applications of multivariate Fay-Herriot (MFH) models. The models allow for various structures of random effects and sampling variances and can further improve the quality of the model-based predictions. In applications to real data, however, MFH models can suffer from partially missing direct estimates of the variables of interest. This can frequently occur when considering direct estimates from different survey or different points in time as dependent variables. Burgard et al. (2021b, 2019) introduce a variant of the bivariate Fay-Herriot model which allows for partially missing direct estimates. They present parameter estimation (ML and REML), derive (empirical) best predictors and approximations to the corresponding MSE for the new model. We extend their work on bivariate models to arbitrary multivariate models and missing structures of the variables of interest, conduct simulation studies, and give an application to publicly available data from the *American Community Survey* (ACS).

Keywords— area-level models, multivariate models, small area estimation, missing values

1 The multivariate FH model

For fine regional and demographic domains, direct survey estimates can be associated with high variability due to small sample sizes. Model-based small

[a]woelwer@uni-trier.de
[b]burgardj@uni-trier.de
[c]d.morales@umh.es
[d]muennich@uni-trier.de

area estimation (SAE) techniques facilitate to increase the effective sample size of domain-level direct estimates by combining similar domains in a common model-based framework; a procedure which is referred to as borrowing strength. A comprehensive overview of SAE techniques is given in Rao and Molina (2015) and Morales et al. (2021). There are two main types of model-based small area estimation techniques, unit- and area-level models. In the small area context their variants are often referred to as Battese-Harter Fuller (BHF) and Fay-Herriot (FH) models respectively, following the works of Battese et al. (1988) and Fay and Herriot (1979). Even though area-level models do not directly use unit-level sampling information, but only aggregate domain statistics, there are a number of reasons for their use, several of them listed in Morales et al. (2021, Chapter 16). For non-linear statistics, the auxiliary information have to be available for the entire target population at the unit-level. This is often not given or leads to the fact that valuable but only aggregated auxiliary information cannot be used. Furthermore, researchers often do not get access to unit-level information on fine regional and demographic levels, but only aggregate statistics.

In the class of area-level small area models, multivariate Fay-Herriot (MFH) models have received more attention in recent years. With MFH models several variables of interest are modelled simultaneously, additionally profiting from the correlation structure between them. One can model one statistic over different points in time or several statistics from the same survey together. The possibility of using additional information from the same survey for SAE motivated Fay (1987) to propose a multivariate version of the FH model. He applied the model to estimate the median income of three-, four-, and five-person households in the *U.S. Current Population Survey* (CPS). The structure of the MFH model accounts for covariances of the sampling errors which is especially necessary when considering variables of interest from the same survey. Even when one is interested in one variable alone, the multivariate modeling can further increase the precision of each variable of interest when the variables are sufficiently correlated. Thus, MFH models facilitate to include further variables of interest as well as (estimated) auxiliary information for which sampling error covariances can be estimated. Further early work with MFH models is given in Datta et al. (1991) and Datta et al. (1999). To name a few applications of the MFH model, in the context of poverty estimation it is applied in Huang and Bell (2004) with further studies in Huang and Bell (2006), Morales et al. (2015), Porter et al. (2015), Benavent and Morales (2016), Arima et al. (2017), Ubaidillah et al. (2019), Benavent and Morales (2021), and Burgard et al. (2021a). We refer to Benavent and Morales (2016) for a general description of the MFH model and its parameter estimation.

2 Partially missing information

Especially when using data from different sources it can frequently occur that information which is needed for modeling is partially missing. Area-level information can be partially missing when the domains of interest are not incorporated in the sampling design (via stratification) and thus - by chance - domain-specific sample sizes can be zero such that no direct estimate can be computed. Furthermore, statistical agencies usually only publish aggregate statistics for which a statistic of the variation, e.g. the standard error or the coefficient of variation, does not exceed a certain threshold. Molina and Marhuenda (2015) recall that in official statistics the threshold for the coefficient of variation is usually set to 20%. In addition to that, for disclosure control statistical agencies set minimum cell counts for the publication of frequency tables, see Hundepool et al. (2010) for an overview. The previously mentioned reasons can lead to missing values both in the variables of interest and auxiliary data. Partially missing auxiliary data can be imputed. Then, however, the associated measurement errors should be considered in the FH model. The use of partially-imputed auxiliary information motivated Lohr and Ybarra (2002) to investigate an extension of the FH model to measurement errors, known as the *measurement error model* which is published in Ybarra and Lohr (2008). The model is extended in Burgard et al. (2021a) to bivariate FH models and a bivariate normal distribution of measurement errors.

Next to the auxiliary data, also the direct estimates of interest may be partially missing. Then, a multivariate - or a corresponding univariate - FH model can only be applied to the domains with full information. Using a model fit on domains with complete data, synthetic predictions can be calculated for domains with missing direct estimates, see e.g. Morales et al. (2021, Chapter 16).

Let there be D domains and $m > 1$ dependent variables. Then, we can partition the set of domains in subsets $\mathcal{D}_0 = \{1, \ldots, D_0\}$ and $\mathcal{D}_1 = \{D_0 + 1, \ldots, D\}$, where $D_0 < D$, such that if the vector of the m direct estimates $\boldsymbol{y}_d$ is completely observed $d \in \mathcal{D}_0$ and if at least one entry of $\boldsymbol{y}_d$ is not observed $d \in \mathcal{D}_1$. Parameters $\boldsymbol{\beta}$, i.e. the fixed effects of the model, and $\boldsymbol{\theta}$, i.e. the variance parameters of the model, can be estimated based on information from $\mathcal{D}_0$ via maximum likelihood (ML) or restricted maximum likelihood (REML). The synthetic predictor of the characteristic of interest $\boldsymbol{\mu}_d$ is given by

$$\hat{\boldsymbol{\mu}}_d^{syn} = \boldsymbol{X}_d\hat{\boldsymbol{\beta}}, \quad d = 1, \ldots, D, \tag{1}$$

where $\boldsymbol{X}_d$ is the matrix of auxiliary information for domain d. For domains with missing direct estimates the mean squared error of the synthetic predictor can be approximated by

$$MSE(\hat{\boldsymbol{\mu}}_d^{syn}) \approx \boldsymbol{X}_d(\boldsymbol{X}_0^\top \boldsymbol{V}_0^{-1}\boldsymbol{X}_0)^{-1}\boldsymbol{X}_d^\top + \boldsymbol{V}_{ud}, \quad \forall d \in \mathcal{D}_1, \tag{2}$$

where quantities $\boldsymbol{X}_0$ and $\boldsymbol{V}_0$ are defined solely based on data from $\mathcal{D}_0$, compare Morales et al. (2021, 441–442). The MSE can be estimated by plugging in $\hat{\boldsymbol{\theta}}$ for $\boldsymbol{\theta}$.

Applying the MFH model only to the domains with complete information, however, is unsatisfactory. The estimation of the parameters (apart from the correlation of the random effects) can be worse than with the corresponding univariate FH models. This occurs when there are only few domains for which no observation is missing and when the missing pattern is heterogeneous across domains. On the other hand, the univariate FH model ignores the correlation of the variables of interest, thereby only using part of the available information. Furthermore, the synthetic predictor is not considering the information of other domain-specific direct estimates in a domain which could give valuable information for the prediction of the missing values.

3 The multivariate FH model under partially missing information

Burgard et al. (2021b, 2019) introduce a bivariate Fay-Herriot model under partially missing direct estimates of the dependent variables, called missing data BFH (MBFH) model. They give ML and REML fitting algorithms to estimate model parameters. Furthermore, they introduce empirical best predictors of target values and derive approximations to the mean squared error. For the bivariate case Burgard et al. (2021b, 2019) allow some of the direct estimates y_{dk}, $k = 1, 2$, to be missing. By setting $y_{\bar{d}1} = (y_{d1}, 0)^\top$ and $y_{\bar{d}2} = (0, y_{d2})^\top$, three groups of domains can be distinguished:

$\mathbb{D}_1 = \{d \in \mathbb{N} : 1 \leq d \leq D_1\}$ containing the D_1 domains where only y_{d1} is observed.

$\mathbb{D}_2 = \{d \in \mathbb{N} : D_1 + 1 \leq d \leq D_1 + D_2\}$ containing the D_2 domains where only y_{d2} is observed.

$\mathbb{D}_3 = \{d \in \mathbb{N} : D_1 + D_2 + 1 \leq d \leq D\}$ containing the remaining domains with fully observed $y_d = (y_{d1}, y_{d2})'$.

The best predictor (BP) of $\boldsymbol{u}_d$ under the MFH model, exemplary shown for domains in $\mathbb{D}_1$, is given by

$$\hat{\boldsymbol{u}}_d^{bp} = E[\boldsymbol{u}_d|\boldsymbol{y}_d] = \boldsymbol{\Phi}_{d1} \begin{pmatrix} \sigma_{ed1}^{-2} & 0 \\ 0 & 0 \end{pmatrix} (y_{\bar{d}1} - \boldsymbol{X}_d\boldsymbol{\beta}), d \in \mathbb{D}_1 \tag{3}$$

with

$$\boldsymbol{\Phi}_{d1} = \left[\begin{pmatrix} \sigma_{ed1}^{-2} & 0 \\ 0 & 0 \end{pmatrix} + \boldsymbol{V}_{ud}^{-1} \right]^{-1}. \tag{4}$$

By considering the partially-missing direct estimates in (3), domain-specific best predictions of random effects can be given, also for the missing direct estimates. This is a significant advantage of the MBFH model compared to the synthetic predictions which could else-wise only be calculated for missing values in a FH or MFH model.

We extend the model introduced in Burgard et al. (2021b, 2019) to multivariate dependent variables, derive empirical best predictors of domain parameters and approximations to the mean squared error. The derived algorithms are