UNIVERSITY OF Southampton

Southampton Statistical Sciences Research Institute

Working Paper M11/07

Methodology

Outlier Robust Small Area Estimation

Iray Chambers,Hukum Chandra,Nicola Salvati,Nikos Tzavidis

Abstract

[Recently proposed outlier robust small area estimators can be substantially biased when outliers are drawn from a distribution that has a different mean from that of the rest of the survey data. This naturally leads to the idea of an outlier robust bias correction for these estimators. In this paper we develop this idea and also propose two different analytical mean squared error estimators for the ensuing bias corrected outlier robust estimators. Simulations based on realistic outlier contaminated data show that the proposed bias correction often leads to more efficient estimators. Furthermore, the proposed mean squared error estimators appear to perform well with a variety of outlier robust small area estimators.]

Outlier Robust Small Area Estimation

Ray Chambers
University of Wollongong, Australia
Hukum Chandra
University of Wollongong, Australia
Nicola Salvati
University of Pisa, Italy
and Nikos Tzavidis
University of Southampton, United Kingdom.

Summary. Recently proposed outlier robust small area estimators can be substantially biased when outliers are drawn from a distribution that has a different mean from that of the rest of the survey data. This naturally leads to the idea of an outlier robust bias correction for these estimators. In this paper we develop this idea and also propose two different analytical mean squared error estimators for the ensuing bias corrected outlier robust estimators. Simulations based on realistic outlier contaminated data show that the proposed bias correction often leads to more efficient estimators. Furthermore, the proposed mean squared error estimators appear to perform well with a variety of outlier robust small area estimators.

Keywords: Bias-variance trade-off; Linear mixed model; M-estimation; M-quantile model; Robust prediction; Robust bias correction.

1. Introduction

Outliers are a fact of life for any survey and as a result, a variety of methods have been devised to mitigate the effects of outlier values on survey estimates. Some of these methods, like identification and removal of these data values by 'experienced' data experts during survey processing, can be effective in ensuring that the resulting survey estimates are unaffected by outliers. However, being somewhat subjective, such methods are not amenable to scientific evaluation. As a consequence there are a number of 'objective' methods for survey estimation that use statistical rules to decide whether an observation is a potential outlier and to down-weight its contribution to the survey estimates if this is the case. Generally, an outlier robust estimator of this type is based on the assumption that the non-outlier sample values all follow a well-behaved working model and so it generally involves prediction of the sum (or mean) of these values under this working model. In practice,

---

Address for correspondence: Ray Chambers, Centre for Statistical and Survey Methodology, School of Mathematics and Applied Statistics, University of Wollongong, Wollongong, Australia. Email: ray@uow.edu.au

this often involves replacement of an outlying sample value by an estimate of what it should have been if in fact it had been generated under the working model. We refer to such methods as *Robust Projective* in what follows since they project sample non-outlier behaviour on to the non-sampled part of the survey population.

Robust Projective methods essentially emulate the subjective approach described earlier, and typically lead to biased estimators with lower variances than would otherwise be the case. The reason for the bias is not difficult to find – it is extremely unlikely that the non-sampled values in the target population are drawn from a distribution with the same mean as the sample non-outliers, and yet these methods are built on precisely this assumption. Chambers (1986) recognised this dilemma and coined the concept of a 'representative outlier', i.e. a sample outlier that is potentially drawn from a group of population outliers and hence cannot be unit-weighted in estimation. He noted that representative outliers cannot be treated on the same basis in estimation as other sample data more consistent with the working model for the population, since such values can seriously destabilise the survey estimates, and suggested addition of an outlier robust bias correction term to a Robust Projective survey estimator, e.g. one based on outlier-robust estimates of the model parameters. Welsh and Ronchetti (1998) expand on this idea, applying it more generally to estimation of the finite population distribution of a survey variable in the presence of representative outliers. A similar idea is implicit in the approach described in Chambers *et al*. (1993), where a nonparametric bias correction is suggested. In what follows, we refer to methods that allow for contributions from representative sample outliers as *Robust Predictive* since they attempt to predict the contribution of the population outliers to the population quantity of interest.

If outliers are a concern for estimation of population quantities, it is safe to say that they are even more of a concern in small area estimation (SAE), where sample sizes are considerably smaller and model-dependent estimation is the norm. It is easy to see that an outlier that destabilises a population estimate based on a large survey sample will almost certainly destroy the validity of the corresponding direct estimate for the small area from which the outlier is sourced since this estimate will be based on a much smaller sample. This problem does not disappear when the small area estimator is an indirect one, e.g. an Empirical Best Linear Unbiased Predictor (EBLUP), since the weights underpinning this estimator will still put most emphasis on data from the small area of interest, and the estimates of the model parameters underpinning the estimator will themselves be destabilised by the sample outliers. Consequently, it is of some interest to see how outlier robust survey estimation can be adapted to this situation.

Chambers and Tzavidis (2006) explicitly address this issue of outlier robustness, using an approach based on fitting outlier robust M-quantile models to the survey data. Recently, Sinha and Rao (2009) have also addressed this issue from the perspective of linear mixed models. Both these approaches, however, use plug-in robust prediction. That is, they replace parameter estimates in optimal, but outlier sensitive, predictors by outlier robust versions (a Robust Projective approach). Unfortunately, this approach may involve an unacceptable prediction bias (but a low prediction variance) in situations where the outliers are drawn from a distribution that has a different mean to the rest of the survey data.

After discussing Robust Projective estimators for small areas in Section 2, we explore the extension of Chambers (1986) Robust Predictive approach to the SAE situation in Section 3. In Section 4 we propose two different analytical mean squared error (MSE) estimators for outlier robust predictors of small area means. In particular, the first proposal is based on the bias-robust mean squared error estimation approach described in Chambers *et al*. (2009) and represents an extension of the ideas in Royall and Cumberland (1978). The second MSE estimator is based on first order approximations to the variances of solutions of estimating equations. We show how these two approaches can be useful for estimating the MSE of various small area predictors considered in this paper. In Sections 5 and 6 we use model-based simulations based on realistic outlier contaminated data scenarios as well as design-based simulations to evaluate how these two different approaches compare, both in terms of point estimation performance as well as in terms of MSE estimation performance. Section 7 concludes the paper with some final remarks, and a discussion of future research aimed at outlier robust small area inference.

2. Robust Projective Estimation for Small Areas

In what follows we assume that unit record data are available at small area level. For the sampled units in the population this consists of indicators of small area affiliation, values y_j of the variable of interest, values $\mathbf{x}_j$ of a $p\times1$ vector of individual level covariates, and values $\mathbf{z}_j$ of a vector of area level covariates. For the non-sampled population units we do not know the values of y_j. However, it is assumed that all areas are sampled and that we know the numbers of such units in each small area and the respective small area averages of $\mathbf{x}_j$ and $\mathbf{z}_j$. We also assume that there is a linear relationship between y_j and $\mathbf{x}_j$ and that sampling is non-informative for the small area distribution of y_j given $\mathbf{x}_j$, allowing us to use population level models with the sample data.

A popular way of using the above data in SAE is to assume a linear mixed model, with random effects for the small areas of interest (see Rao, 2003). Let $\mathbf{y}$, $\mathbf{X}$ and $\mathbf{Z}$ denote the population level vector and matrices defined by y_j, $\mathbf{x}_j$ and $\mathbf{z}_j$ respectively. Then

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}, \tag{1}$$

where $\mathbf{u} = \left(\mathbf{u}_1^T, \ldots, \mathbf{u}_m^T\right)^T$ is a vector of dimension mq made up of m independent realisations $\{\mathbf{u}_i; i = 1, \ldots, m\}$ of a q-dimensional random area effect with $\mathbf{u} \sim N(\mathbf{0}, \boldsymbol{\Sigma}_u)$ and $\mathbf{e} \sim N(\mathbf{0}, \boldsymbol{\Sigma}_e)$ is a vector of N individual specific random effects. It is also assumed that $\mathbf{u}$ is distributed independently of $\mathbf{e}$. Here m is the total number of small areas that make up the population and q is the dimension of $\mathbf{z}_j$ so that $\mathbf{Z}$ is a $N \times mq$ matrix of fixed known constants. We assume that the covariance matrices $\boldsymbol{\Sigma}_u$ and $\boldsymbol{\Sigma}_e$ are defined in terms of a lower dimensional set of parameters $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_K)$, which are typically referred to as the variance components of (1), while $\boldsymbol{\beta}$ is usually referred to as its fixed effect.

Let $\hat{\boldsymbol{\beta}}$ and $\hat{\mathbf{u}}$ denote estimates of the fixed and random effects in (1). The EBLUP of the area i mean of the y_j under (1) is then

$$\hat{\bar{y}}_i^{EBLUP} = N_i^{-1}\left\{n_i \bar{y}_{si} + \left(N_i - n_i\right)\left(\bar{\mathbf{x}}_{ri}^T \hat{\boldsymbol{\beta}} + \bar{\mathbf{z}}_{ri}^T \hat{\mathbf{u}}\right)\right\}, \tag{2}$$

where $\hat{\mathbf{u}} = \left(\hat{\mathbf{u}}_1^T, \ldots, \hat{\mathbf{u}}_m^T\right)^T$ denotes the vector of the estimated area specific random effects and we use indices of s and r to denote sample and non-sample quantities respectively. Thus, $\bar{y}_{si}$ is the average of the n_i sample values of y_j from area i and $\bar{\mathbf{x}}_{ri}$ and $\bar{\mathbf{z}}_{ri}$ denoting the vectors of average values of $\mathbf{x}_j$ and $\mathbf{z}_j$ respectively for the $N_i - n_i$ non-sampled units in the same area.

From a Robust Projective viewpoint, (2) can be made insensitive to sample outliers by replacing $\hat{\boldsymbol{\beta}}$ and $\hat{\mathbf{u}}$ by outlier robust alternatives. To motivate this approach, we initially assume the variance components $\boldsymbol{\theta}$ are known, so the covariance matrices $\boldsymbol{\Sigma}_u$ and $\boldsymbol{\Sigma}_e$ in (1) are known. Put

$\mathbf{V}_s = \mathbf{\Sigma}_{es} + \mathbf{Z}_s \mathbf{\Sigma}_u \mathbf{Z}_s^T$ where $\mathbf{\Sigma}_{es}$ denotes the sample component of $\mathbf{\Sigma}_e$. Then the BLUE of the fixed effects vector $\boldsymbol{\beta}$ is

$$\tilde{\boldsymbol{\beta}} = \left\{ \mathbf{X}_s^T \mathbf{V}_s^{-1} \mathbf{X}_s \right\}^{-1} \mathbf{X}_s^T \mathbf{V}_s^{-1} \mathbf{y}_s , \tag{3}$$

while the BLUP of the random effects vector $\mathbf{u}$ is

$$\tilde{\mathbf{u}} = \Sigma_u \mathbf{Z}_s^T \mathbf{V}_s^{-1} \left(\mathbf{y}_s - \mathbf{X}_s \tilde{\boldsymbol{\beta}} \right). \tag{4}$$

It is easy to see that (3) and (4) are solutions to

$$\mathbf{X}_s^T \mathbf{V}_s^{-1} \left(\mathbf{y}_s - \mathbf{X}_s \boldsymbol{\beta} \right) = \mathbf{0} \tag{5}$$

and

$$\mathbf{\Sigma}_u \mathbf{Z}_s^T \mathbf{V}_s^{-1} \left(\mathbf{y}_s - \mathbf{X}_s \boldsymbol{\beta} \right) - \mathbf{u} = \mathbf{0} . \tag{6}$$

A straightforward way to make the solutions to (5) and (6) robust to sample outliers is therefore to replace them by

$$\mathbf{X}_s^T \mathbf{V}_s^{-1/2} \psi \left(\mathbf{V}_s^{-1/2} \{ \mathbf{y}_s - \mathbf{X}_s \boldsymbol{\beta} \} \right) = \mathbf{0} \tag{7}$$

and

$$\mathbf{\Sigma}_u \mathbf{Z}_s^T \mathbf{V}_s^{-1/2} \psi \left(\mathbf{V}_s^{-1/2} \{ \mathbf{y}_s - \mathbf{X}_s \beta \} \right) - \mathbf{\Sigma}_u^{1/2} \psi \left(\mathbf{\Sigma}_u^{-1/2} \mathbf{u} \right) = \mathbf{0} . \tag{8}$$

Here ψ is a bounded influence function and $\psi(\mathbf{a})$ denotes the vector defined by applying ψ to every component of $\mathbf{a}$. Unfortunately, since $\mathbf{V}_s$ is not a diagonal matrix, the solution to (8) can be numerically unstable. An alternative approach was therefore suggested by Fellner (1986), who noted that any solution to (5) and (6) was also a solution to

$$\mathbf{X}_s^T \mathbf{\Sigma}_{es}^{-1} \left(\mathbf{y}_s - \mathbf{X}_s \boldsymbol{\beta} - \mathbf{Z}_s \mathbf{u} \right) = \mathbf{0} \quad \text{and} \quad \mathbf{Z}_s^T \mathbf{\Sigma}_{es}^{-1} \left(\mathbf{y}_s - \mathbf{X}_s \boldsymbol{\beta} - \mathbf{Z}_s \mathbf{u} \right) - \mathbf{\Sigma}_u^{-1} \mathbf{u} = \mathbf{0} .$$

Fellner (1986) suggested that these alternative estimating equations (and hence their solutions) be made outlier robust by replacing them by

$$\mathbf{X}_s^T \mathbf{\Sigma}_{es}^{-1/2} \psi \left(\mathbf{\Sigma}_{es}^{-1/2} \{ \mathbf{y}_s - \mathbf{X}_s \boldsymbol{\beta} - \mathbf{Z}_s \mathbf{u} \} \right) = \mathbf{0} \tag{9}$$

and

$$\mathbf{Z}_s^T \mathbf{\Sigma}_{es}^{-1/2} \psi \left(\mathbf{\Sigma}_{es}^{-1/2} \{ \mathbf{y}_s - \mathbf{X}_s \boldsymbol{\beta} - \mathbf{Z}_s \mathbf{u} \} \right) - \mathbf{\Sigma}_u^{-1/2} \psi \left(\mathbf{\Sigma}_u^{-1/2} \mathbf{u} \right) = \mathbf{0} . \tag{10}$$

Since (9) and (10) assume the variance components $\boldsymbol{\theta}$ are known, their usefulness is somewhat limited unless outlier robust estimators of these parameters can also be defined. Richardson and Welsh (1995) propose two outlier robust variations to the maximum likelihood estimating equations for $\boldsymbol{\theta}$. One of these (ML Proposal II) leads to an estimating equation for the variance component θ_k of $\boldsymbol{\theta}$ of the form

$$\psi\left\{(\mathbf{y}_s-\mathbf{X}_s\boldsymbol{\beta})^T\mathbf{V}_s^{-1/2}\right\}\mathbf{V}_s^{-1/2}(\partial\mathbf{V}_s/\partial\theta_k)\mathbf{V}_s^{-1/2}\psi\left\{\mathbf{V}_s^{-1/2}(\mathbf{y}_s-\mathbf{X}_s\boldsymbol{\beta})\right\}=tr\left\{\mathbf{D}_n^{\psi}(\partial\mathbf{V}_s/\partial\theta_k)\right\}, \tag{11}$$

where $\partial\mathbf{V}_s/\partial\theta_k$ denotes the first order partial derivative of $\mathbf{V}_s$ with respect to the variance component θ_k and, for $Z\sim N(0,1)$, $\mathbf{D}_n^{\psi}=E\left\{\psi^2(Z)\right\}\mathbf{V}_s^{-1}$.

Sinha and Rao (2009) describe an approach to outlier robust estimation of $\boldsymbol{\beta}$ and $\mathbf{u}$ in (1) that builds on these results, substituting approximate solutions to both (7) and (11) into the Fellner estimating equation (10) to obtain an outlier robust estimate of the area effect $\mathbf{u}$. In particular, their approach replaces (7) by

$$\mathbf{X}_s^T\mathbf{V}_s^{-1}\mathbf{U}_s^{1/2}\psi\left(\mathbf{U}_s^{-1/2}\{\mathbf{y}_s-\mathbf{X}_s\boldsymbol{\beta}\}\right)=\mathbf{0}, \tag{12}$$

where $\mathbf{U}_s=diag\left(\mathbf{V}_s\right)$, and replaces (11) by

$$\psi\left\{(\mathbf{y}_s-\mathbf{X}_s\boldsymbol{\beta})^T\mathbf{U}_s^{-1/2}\right\}\mathbf{U}_s^{1/2}\mathbf{V}_s^{-1}(\partial\mathbf{V}_s/\partial\theta_k)\mathbf{V}_s^{-1}\mathbf{U}_s^{1/2}\psi\left\{\mathbf{U}_s^{-1/2}(\mathbf{y}_s-\mathbf{X}_s\boldsymbol{\beta})\right\}=tr\left\{\mathbf{D}_n^{\psi}(\partial\mathbf{V}_s/\partial\theta_k)\right\}. \tag{13}$$

Since the solutions to (12) and (13) depend on the influence function ψ, we denote them by a superscript of ψ below. The Sinha and Rao (2009) Robust Projective alternative to (2) is then

$$\hat{\bar{y}}_i^{SR}=\bar{\mathbf{x}}_i^T\hat{\boldsymbol{\beta}}^{\psi}+\bar{\mathbf{z}}_i^T\hat{\mathbf{u}}^{\psi}. \tag{14}$$

Note that (14) estimates the area *i* mean under (1). A minor modification restricts this to the mean of the non-sampled units in area *i*, in which case (14) becomes

$$\hat{\bar{y}}_i^{REBLUP}=N_i^{-1}\left\{n_i\bar{y}_{si}+\left(N_i-n_i\right)\left(\bar{\mathbf{x}}_{ri}^T\hat{\boldsymbol{\beta}}^{\psi}+\bar{\mathbf{z}}_{ri}^T\hat{\mathbf{u}}^{\psi}\right)\right\}. \tag{15}$$

Hereafter we call this estimator the Robust EBLUP (REBLUP). An alternative methodology for outlier robust SAE is the M-quantile regression-based method described by Chambers and Tzavidis (2006). This is based on a linear model for the M-quantile regression of $\mathbf{y}$ on $\mathbf{X}$, i.e.

$$m_q(\mathbf{X})=\mathbf{X}\boldsymbol{\beta}_q, \tag{16}$$

where $m_q(\mathbf{X})$ denotes the M-quantile of order q of the conditional distribution of $\mathbf{y}$ given $\mathbf{X}$. An estimate $\hat{\boldsymbol{\beta}}_q$ of $\boldsymbol{\beta}_q$ can be calculated for any value of q in the interval (0,1), and for each unit in sample we define its unique M-quantile coefficient under this fitted model as the value q_j such that $y_j = \mathbf{x}_j^T \hat{\boldsymbol{\beta}}_{q_j}$, with the sample average of these coefficients in area *i* denoted by $\overline{q}_i$. The M-quantile estimate of the mean of y_j in area *i* is then

$$\hat{\overline{y}}_i^{MQ} = N_i^{-1}\left\{ n_i \overline{y}_{si} + \left(N_i - n_i \right) \overline{\mathbf{x}}_{ri}^T \hat{\boldsymbol{\beta}}_{\overline{q}_i} \right\}. \tag{17}$$

Note that the regression M-quantile model (16) depends on the influence function ψ underpinning the M-quantile. When this function is bounded, sample outliers have limited impact on $\hat{\boldsymbol{\beta}}_q$. That is, (17) corresponds to assuming that all non-sample units in area *i* follow the working model (16) with $q = \overline{q}_i$, in the sense that one can write $y_j = \mathbf{x}_j^T \boldsymbol{\beta}_{\overline{q}_i} + \text{ noise}$ for all such units.

3. Robust Predictive Estimation for Small Areas

A problem with the Robust Projective approach is that it assumes all non-sampled units follow the working model, or, in what essentially amounts to the same thing, that any deviations from this model are noise and so cancel out 'on average'. Thus, under the linear mixed model (1) one can see that provided the individual errors of the non-sampled units are symmetrically distributed about zero, the REBLUP (15) by Sinha and Rao (2009) will perform well since it is based on the implicit assumption that the average of these errors over the non-sampled units in area *i* converges to zero. The M-quantile estimator MQ (17) is no different since it assumes that the errors $y_j - \mathbf{x}_j^T \boldsymbol{\beta}_{\overline{q}_i}$ from the area *i*-specific M-quantile regression model are 'noise' and hence also cancel out on average. Note that this does not mean that these non-sample units are not outliers. It is just that their behaviour is such that our best prediction of their corresponding average value is zero.

Welsh and Ronchetti (1998) consider the issue of outlier robust prediction within the context of population level survey estimation. Starting with a working linear model linking the population values of y_j and $\mathbf{x}_j$, and sample data containing representative outliers with respect to this model, they extend the approach of Chambers (1986) to robust prediction of the empirical distribution function of the population values of y_j. Their argument immediately applies to robust prediction of the empirical distribution function of the area *i* values of y_j, and leads to a predictor of the form

$$\hat{F}_i^{\psi\phi}(t) = N_i^{-1}\left[\sum_{j\in s_i} I(y_j \le t) + n_i^{-1}\sum_{j\in s_i}\sum_{k\in r_i} I\left(\mathbf{x}_k^T\hat{\boldsymbol{\beta}}^{\psi} + \omega_{ij}^{\psi}\phi\left\{\left(y_j - \mathbf{x}_j^T\hat{\boldsymbol{\beta}}^{\psi}\right)\big/\omega_{ij}^{\psi}\right\} \le t\right)\right]. \tag{18}$$

Here $\hat{\boldsymbol{\beta}}^{\psi}$ denotes an M-estimator of the regression parameter in the linear working model based on a bounded influence function ψ, ω_{ij}^{ψ} is a robust estimator of the scale of the residual $y_j - \mathbf{x}_j^T\hat{\boldsymbol{\beta}}^{\psi}$ in area i and ϕ denotes a bounded influence function that satisfies $|\phi| \ge |\psi|$. Tzavidis *et al*. (2010) note that the robust estimator of the area i mean of the y_j defined by (18) is just the expected value functional defined by it, which is

$$\hat{\bar{y}}_i^{\psi\phi} = \int t d\hat{F}_i^{\psi\phi}(t) = N_i^{-1}\left[n_i\bar{y}_{si} + \left(N_i - n_i\right)\left(\bar{\mathbf{x}}_{ri}^T\hat{\boldsymbol{\beta}}^{\psi} + n_i^{-1}\sum_{j\in s_i}\omega_{ij}^{\psi}\phi\left\{\left(y_j - \mathbf{x}_j^T\hat{\boldsymbol{\beta}}^{\psi}\right)\big/\omega_{ij}^{\psi}\right\}\right)\right]. \tag{19}$$

These authors therefore suggest an extension to the M-quantile estimator (17) by replacing $\hat{\boldsymbol{\beta}}^{\psi}$ in (19) by $\hat{\boldsymbol{\beta}}_{\bar{q}_i}$, which leads to a 'bias-corrected' version of (17), hereafter MQ-BC, given by

$$\hat{\bar{y}}_i^{MQ-BC} = N_i^{-1}\left[n_i\bar{y}_{si} + \left(N_i - n_i\right)\left(\bar{\mathbf{x}}_{ri}^T\hat{\boldsymbol{\beta}}_{\bar{q}_i} + n_i^{-1}\sum_{j\in s_i}\omega_{ij}^{MQ}\phi\left\{\left(y_j - \mathbf{x}_j^T\hat{\boldsymbol{\beta}}_{\bar{q}_i}\right)\big/\omega_{ij}^{MQ}\right\}\right)\right] \tag{20}$$

and ω_{ij}^{MQ} is a robust estimator of the scale of the residual $y_j - \mathbf{x}_j^T\hat{\boldsymbol{\beta}}_{\bar{q}_i}$ in area i.

The use of the two influence functions ψ and ϕ in (20) is worthy of comment. The first, ψ, underpins $\hat{\boldsymbol{\beta}}_q$, and hence $\hat{\boldsymbol{\beta}}_{\bar{q}_i}$. Its purpose is to ensure that sample outliers have little or no influence on the fit of the working M-quantile model. As a consequence it is bounded and down-weights these outliers. The second, ϕ, is still bounded but 'less restrictive' than ψ (since $|\phi| \ge |\psi|$) and its purpose is to define an adjustment for the bias caused by the fact that the first two terms on the right hand side of (20) treat sample outliers as self-representing. A similar argument can be used to modify the REBLUP (15). In particular, a Robust Predictive version of this estimator, hereafter REBLUP-BC, mimics the bias correction idea used in (20) and leads to

$$\hat{\bar{y}}_i^{REBLUP-BC} = \hat{\bar{y}}_i^{REBLUP} + \left(1 - n_iN_i^{-1}\right)n_i^{-1}\sum_{j\in s_i}\omega_{ij}^{\psi}\phi\left\{\left(y_j - \mathbf{x}_j^T\hat{\boldsymbol{\beta}}^{\psi} - \mathbf{z}_j^T\hat{\mathbf{u}}^{\psi}\right)\big/\omega_{ij}^{\psi}\right\}, \tag{21}$$

where the ω_{ij}^{ψ} are now robust estimates of the scale of the area i residuals $y_j - \mathbf{x}_j^T\hat{\boldsymbol{\beta}}^{\psi} - \mathbf{z}_j^T\hat{\mathbf{u}}^{\psi}$.

4. MSE Estimation for Robust Predictors

In this Section we propose two different analytic methods of MSE estimation for robust predictors of small area means under the Robust Projective and Robust Predictive approaches. Both are developed on the assumption that the working model for inference conditions on the realised values of the area effects, and so the proposed MSE estimators are conditional estimators. In Section 4.1 we apply the ideas set out by Chambers *et al*. (2009) to define a pseudo-linearization estimator of the conditional MSE of the REBLUP (15). Similar conditional MSE estimators for the REBLUP-BC (21), MQ (17) and MQ-BC (20) predictors follow directly. In Section 4.2 we use first order approximations to the variances of solutions of estimating equations to develop conditional MSE estimators for the REBLUP (15) and the REBLUP-BC (21). Analogous MSE estimators for the MQ (17) and the MQ-BC (20) predictors based on this approach are described in the Appendix.

4.1 Pseudo-linearization approach to MSE estimation for small area predictors

Sinha and Rao (2009) proposed a parametric bootstrap-based estimator for the MSE of REBLUP. Here we describe an analytical estimator of the conditional MSE of the REBLUP that is less computationally demanding. The proposed estimator is based on the pseudo-linearization approach to MSE estimation described by Chambers *et al*. (2009), which can be used for predictors that can be expressed as weighted sums of the sample values. Since the REBLUP can be expressed in a pseudo-linear form, i.e. as a weighted sum of the sample values of *y*, this approach is immediately applicable. To start, we note that under model (1), and assuming that the variance components are known, the Robust BLUP or RBLUP of $\bar{y}_i$ can be expressed as

$$\hat{\bar{y}}_i^{RBLUP} = \sum_{j \in s} w_{ij}^{RBLUP} y_j = \left(\mathbf{w}_{is}^{RBLUP}\right)^T \mathbf{y}_s , \tag{22}$$

where

$$\left(\mathbf{w}_{is}^{RBLUP}\right)^T = N_i^{-1}\left\{\mathbf{1}_s^T + \left(N_i - n_i\right)\left[\bar{\mathbf{x}}_{ir}^T \mathbf{A}_s + \bar{\mathbf{z}}_{ir}^T \mathbf{B}_s\left(\mathbf{I}_s - \mathbf{X}_s \mathbf{A}_s\right)\right]\right\}.$$

Here

- $\mathbf{A}_s = \left(\mathbf{X}_s^T \mathbf{V}_s^{-1} \mathbf{U}_s^{1/2} \mathbf{W}_{1s} \mathbf{U}_s^{-1/2} \mathbf{X}_s\right)^{-1} \mathbf{X}_s^T \mathbf{V}_s^{-1} \mathbf{U}_s^{1/2} \mathbf{W}_{1s} \mathbf{U}_s^{-1/2}$, with $\mathbf{W}_{1s}$ a $n \times n$ diagonal matrix of weights with

j-th component $w_{1j} = \psi\left(U_j^{-1/2}\left\{y_j - \mathbf{x}_j^T\tilde{\boldsymbol{\beta}}^{\psi}\right\}\right) \Big/ U_j^{-1/2}\left\{y_j - \mathbf{x}_j^T\tilde{\boldsymbol{\beta}}^{\psi}\right\}$;

- $\mathbf{B}_s = \left(\mathbf{Z}_s^T\boldsymbol{\Sigma}_{es}^{-1/2}\mathbf{W}_{2s}\boldsymbol{\Sigma}_{es}^{-1/2}\mathbf{Z}_s + \boldsymbol{\Sigma}_u^{-1/2}\mathbf{W}_{3s}\boldsymbol{\Sigma}_u^{-1/2}\right)^{-1}\left(\mathbf{Z}_s^T\boldsymbol{\Sigma}_{es}^{-1/2}\mathbf{W}_{2s}\boldsymbol{\Sigma}_{es}^{-1/2}\right)$, with $\mathbf{W}_{2s}$ a $n \times n$ diagonal matrix of weights with j-th component $w_{2j} = \psi\left(\left(\sigma_e^{\psi}\right)^{-1}\left\{y_j - \mathbf{x}_j^T\tilde{\boldsymbol{\beta}}^{\psi} - \mathbf{z}_j\tilde{\mathbf{u}}^{\psi}\right\}\right) \Big/ \left(\sigma_e^{\psi}\right)^{-1}\left\{y_j - \mathbf{x}_j^T\tilde{\boldsymbol{\beta}}^{\psi} - \mathbf{z}_j\tilde{\mathbf{u}}^{\psi}\right\}$, and $\mathbf{W}_{3s}$ is a $m \times m$ diagonal matrix of weights with i-th component $w_{3i} = \psi\left(\left(\sigma_u^{\psi}\right)^{-1}\tilde{u}_i^{\psi}\right) \Big/ \left(\sigma_u^{\psi}\right)^{-1}\tilde{u}_i^{\psi}$;
- $\tilde{\boldsymbol{\beta}}^{\psi}$ and $\tilde{u}_i^{\psi}$ are the solutions to (12) and (13) when variance components are known.

In addition, $\mathbf{1}_s$ is the n-vector with j-th component equal to one whenever the corresponding sample unit is in area i and is zero otherwise. The REBLUP (15) can be expressed in exactly the same way, except that all quantities in the weight vector $\mathbf{w}_{is}^{RBLUP}$ that depend on (unknown) variance components now need a 'hat', in which case we denote it by $\mathbf{w}_{is}^{REBLUP}$. Given this pseudo-linear representation for the REBLUP, a simple first order approximation to its MSE is developed assuming the conditional version of the model (1), i.e. the random effects are considered to be fixed, but unknown, quantities. Let $I(j \in i)$ denote the indicator for whether unit j is in area i. The estimator of the conditional MSE of the REBLUP is then written as

$$\widehat{MSE}\left(\hat{\bar{y}}_i^{REBLUP}\right) = \hat{V}\left(\hat{\bar{y}}_i^{REBLUP}\right) + \left\{\hat{B}\left(\hat{\bar{y}}_i^{REBLUP}\right)\right\}^2, \tag{23}$$

where

$$\hat{V}(\hat{\bar{y}}_i^{REBLUP}) = N_i^{-2}\sum\nolimits_{j \in s}\left\{a_{ij}^2 + (N_i - n_i)n^{-1}\right\}\lambda_j^{-1}(y_j - \hat{\mu}_j)^2$$

is the estimate of the conditional prediction variance of (23) with $a_{ij} = N_i w_{ij}^{REBLUP} - I(j \in i)$ and

$$\hat{B}(\hat{\bar{y}}_i^{REBLUP}) = \sum\nolimits_{j \in s} w_{ij}^{REBLUP}\hat{\mu}_j - N_i^{-1}\sum\nolimits_{j \in (r_i \cup s_i)}\hat{\mu}_j$$

is the estimate of its conditional prediction bias. In order to implement (23) we need to define $\hat{\mu}_j$ and $\hat{\lambda}_j$. Here $\hat{\mu}_j = \sum\nolimits_{k \in s}\phi_{kj}y_k$ is an unbiased linear estimator of the conditional expected value $\mu_j = E\left(y_j \mid \mathbf{x}_j, \mathbf{u}^{\psi}\right)$ and $\lambda_j = \left\{1 - 2\phi_{jj} + \sum\nolimits_{k \in s}\phi_{kj}^2\right\}$ is a scaling constant. Because of the well-known shrinkage effect associated with BLUPs, replacing $\hat{\mu}_j$ by the EBLUP of μ_j under (1) can lead to biased estimation of the conditional prediction variance. Chambers *et al*. (2009) therefore recommend that $\hat{\mu}_j$ be computed as the 'unshrunken' version of the EBLUP for μ_j. See also Salvati *et al*. (2010). Note that the MSE estimator (23) ignores the extra variability associated with estimation of the variance components, and hence is a first order approximation to the actual conditional MSE of the REBLUP.

The MSE estimator for the REBLUP-BC (21) is obtained using the same pseudo-linearization approach as outlined above. The only difference is that the weights w_{ij}^{REBLUP} used in (23) are now replaced by corresponding REBLUP-BC weights. Furthermore, since the REBLUP-BC is an approximately unbiased estimator of the small area mean, the squared bias term in (23) is omitted.

It has been empirically demonstrated that this method of MSE has good repeated sampling properties for realistic small area applications - see Chandra and Chambers (2009), Chambers and Tzavidis (2006), Chandra *et al*. (2007) and Tzavidis *et al*. (2010). Although empirical results (see Chambers *et al.*, 2009) show that this method of MSE estimation performs reasonably well in terms of bias, this improved bias performance comes at the cost of increased variability. In particular, when the area-specific sample sizes are very small, the use of this type of MSE estimator can lead to MSE estimates with high variance. As a result, analysts must be cautious when applying this MSE estimator with very small area-specific sample sizes.

4.2 Linearization based MSE estimation for small area predictors

In what follows we build on the linearization ideas set out in Street *et al*. (1988) to propose a new estimator of the MSE of a small area estimator that is defined by the solution of a set of robust estimating equations. We then illustrate this approach by applying it to estimation of the conditional MSE of the REBLUP (15) and the REBLUP-BC (21). The corresponding MSE estimator for the EBLUP predictor (2) can be obtained as a special case of the MSE estimator of the REBLUP (15). In the economy of space, the development omits some technical details, but these are available from the authors upon request. Note that when used with an estimator based on a mixed model, the proposed MSE estimator provides a second order approximation to the conditional MSE since it includes a term for the contribution to the variability resulting from the estimation of the variance components.

MSE estimation for REBLUP and REBLUP-BC

Under model (1) the conditional prediction variance of the Robust BLUP (RBLUP) of $\bar{y}_i$ can be expressed as

$$
\begin{aligned}
Var_{\mathbf{u}}\left(\hat{\bar{y}}_i^{RBLUP} - \bar{y}_i\right) &= Var_{\mathbf{u}}\left\{N_i^{-1}\sum_{j\in r_i}\left(\mathbf{x}_j^T\tilde{\boldsymbol{\beta}}^{\psi} + \mathbf{z}_j^T\tilde{\mathbf{u}}^{\psi}\right) - N_i^{-1}\sum_{j\in r_i} y_j\right\} \\
&= \left(1-n_iN_i^{-1}\right)^2 \bar{\mathbf{x}}_{ri}^T Var_{\mathbf{u}}\left(\tilde{\boldsymbol{\beta}}^{\psi}\right)\bar{\mathbf{x}}_{ri} + \left(1-n_iN_i^{-1}\right)^2 \bar{\mathbf{z}}_{ri}^T Var_{\mathbf{u}}\left(\tilde{\mathbf{u}}^{\psi}\right)\bar{\mathbf{z}}_{ri} + \left(1-n_iN_i^{-1}\right)^2 Var_{\mathbf{u}}\left(\bar{e}_{ri}\right),
\end{aligned} \tag{24}
$$

where we assume independence between $\tilde{\boldsymbol{\beta}}^{\psi}$ and $\tilde{\mathbf{u}}^{\psi}$. Here a superscript of $\mathbf{u}$ is used to denote moments that are conditioned on the realised values of the area effects. As a result, in order to be able to calculate the prediction variance of RBLUP, we need to estimate $Var_{\mathbf{u}}(\tilde{\boldsymbol{\beta}}^{\psi})$ and $Var_{\mathbf{u}}(\tilde{\mathbf{u}}^{\psi})$, where $\tilde{\mathbf{u}}^{\psi} = \left(\tilde{\mathbf{u}}_1^{\psi T},\ldots,\tilde{\mathbf{u}}_m^{\psi T}\right)^T$. For doing this, put $\boldsymbol{\delta} = \left(\boldsymbol{\beta}^{\psi T},\mathbf{u}^{\psi T}\right)^T$, so $\tilde{\boldsymbol{\delta}} = \left(\tilde{\boldsymbol{\beta}}^{\psi T},\tilde{\mathbf{u}}^{\psi T}\right)^T$ with corresponding 'true' value $\boldsymbol{\delta}_0 = \left(\boldsymbol{\beta}_0^{\psi T},\mathbf{u}_0^{\psi T}\right)^T$. Then, from equations (10) and (12), $\mathbf{H}(\tilde{\boldsymbol{\delta}}) = \mathbf{0}$ where

$$
\mathbf{H}(\boldsymbol{\delta}) = \begin{pmatrix} \mathbf{H}_{\beta^{\psi}}(\boldsymbol{\delta}) \\ \mathbf{H}_{u^{\psi}}(\boldsymbol{\delta}) \end{pmatrix} = \begin{pmatrix} \mathbf{X}_s^T\mathbf{V}_s^{-1}\mathbf{U}_s^{1/2}\psi\left(\mathbf{U}_s^{-1/2}\left\{\mathbf{y}_s - \mathbf{X}_s\boldsymbol{\beta}^{\psi}\right\}\right) = \mathbf{0} \\ \mathbf{Z}_s^T\boldsymbol{\Sigma}_{es}^{-1/2}\psi\left(\boldsymbol{\Sigma}_{es}^{-1/2}\left\{\mathbf{y}_s - \mathbf{X}_s\boldsymbol{\beta}^{\psi} - \mathbf{Z}_s\mathbf{u}^{\psi}\right\}\right) - \boldsymbol{\Sigma}_u^{-1/2}\psi\left(\boldsymbol{\Sigma}_u^{-1/2}\mathbf{u}^{\psi}\right) = \mathbf{0} \end{pmatrix}.
$$

We then use results by Welsh and Richardson (1997) and Sinha and Rao (2009) on the asymptotic variance of solutions to an estimating equation to obtain a first order approximation to $Var_{\mathbf{u}}(\tilde{\boldsymbol{\delta}})$ and by extension to the conditional prediction variance of the RBLUP. To do this, we note that

$$
Var_{\mathbf{u}}(\tilde{\boldsymbol{\delta}}) \approx \left\{E_{\mathbf{u}}\left(\partial_{\boldsymbol{\delta}}\mathbf{H}_0\right)\right\}^{-1} Var_{\mathbf{u}}\left\{\mathbf{H}(\boldsymbol{\delta}_0)\right\}\left[\left\{E_{\mathbf{u}}\left(\partial_{\boldsymbol{\delta}}\mathbf{H}_0\right)\right\}^{-1}\right]^T
$$

where

$$
\begin{aligned}
Var_{\mathbf{u}}\left(\mathbf{H}_{\beta^{\psi}}(\boldsymbol{\delta}_0)\right) &= Var_{\mathbf{u}}\left[\psi\left\{U_j^{-1/2}\left(y_j - \mathbf{x}_j^T\boldsymbol{\beta}_0^{\psi}\right)\right\}\right]\mathbf{X}_s^T\mathbf{V}_s^{-1}\mathbf{U}_s\mathbf{V}_s^{-1}\mathbf{X}_s \\
Var_{\mathbf{u}}\left(\mathbf{H}_{u^{\psi}}(\boldsymbol{\delta}_0)\right) &= E_{\mathbf{u}}\left[\psi^2\left\{\left(\sigma_e^{\psi}\right)^{-1}\left(y_j - \mathbf{x}_j^T\boldsymbol{\beta}_0^{\psi} - \mathbf{z}_j^T\mathbf{u}_0^{\psi}\right)\right\}\right]\mathbf{Z}_s^T\boldsymbol{\Sigma}_{es}^{-1}\mathbf{Z}_s,
\end{aligned} \tag{25}
$$

and

$$
E_{\mathbf{u}}\left(\partial_{\boldsymbol{\delta}}\mathbf{H}_0\right) = E_{\mathbf{u}}\begin{bmatrix} \partial_{\beta^{\psi}}\mathbf{H}_{\beta^{\psi}}(\boldsymbol{\delta}_0) & \partial_{\beta^{\psi}}\mathbf{H}_{u^{\psi}}(\boldsymbol{\theta}_0) \\ \partial_{u^{\psi}}\mathbf{H}_{\beta^{\psi}}(\boldsymbol{\delta}_0) & \partial_{u^{\psi}}\mathbf{H}_{u^{\psi}}(\boldsymbol{\theta}_0) \end{bmatrix} = \begin{bmatrix} E_{\mathbf{u}}\left\{\partial_{\beta^{\psi}}\mathbf{H}_{\beta^{\psi}}(\boldsymbol{\delta}_0)\right\} & E_{\mathbf{u}}\left\{\partial_{\beta^{\psi}}\mathbf{H}_{u^{\psi}}(\boldsymbol{\theta}_0)\right\} \\ \mathbf{0} & E_{\mathbf{u}}\left\{\partial_{u^{\psi}}\mathbf{H}_{u^{\psi}}(\boldsymbol{\theta}_0)\right\} \end{bmatrix}
$$

with

$$
\begin{aligned}
E_{\mathbf{u}}\left\{\partial_{\beta^{\psi}}\mathbf{H}_{\beta^{\psi}}(\boldsymbol{\delta}_0)\right\} &= -\mathbf{X}_s^T\mathbf{V}_s^{-1}\mathbf{U}_s^{1/2}E_{\mathbf{u}}\left[\psi'\left\{\mathbf{U}_s^{-1/2}\left(\mathbf{y}_s - \mathbf{X}_s\boldsymbol{\beta}_0^{\psi}\right)\right\}\right]\mathbf{U}_s^{-1/2}\mathbf{X}_s \\
E_{\mathbf{u}}\left\{\partial_{u^{\psi}}\mathbf{H}_{u^{\psi}}(\boldsymbol{\theta}_0)\right\} &= -\mathbf{Z}_s^T\boldsymbol{\Sigma}_{es}^{-1/2}E_{\mathbf{u}}\left[\psi'\left\{\boldsymbol{\Sigma}_{es}^{-1/2}\left(\mathbf{y}_s - \mathbf{X}_s\boldsymbol{\beta}_0^{\psi} - \mathbf{Z}_s\mathbf{u}_0^{\psi}\right)\right\}\right]\boldsymbol{\Sigma}_{es}^{-1/2}\mathbf{Z}_s - \boldsymbol{\Sigma}_u^{-1/2}\left[\psi'\left\{\boldsymbol{\Sigma}_u^{-1/2}\mathbf{u}_0^{\psi}\right\}\right]\boldsymbol{\Sigma}_u^{-1/2}.
\end{aligned} \tag{26}
$$

Since

$$\left\{E_{\mathbf{u}}\left(\partial_{\boldsymbol{\delta}}\mathbf{H}_0\right)\right\}^{-1} = \begin{bmatrix} \left[E_{\mathbf{u}}\left\{\partial_{\beta^\psi}\mathbf{H}_{\beta^\psi}(\boldsymbol{\delta}_0)\right\}\right]^{-1} & -\left[E_{\mathbf{u}}\left\{\partial_{\beta^\psi}\mathbf{H}_{\beta^\psi}(\boldsymbol{\delta}_0)\right\}\right]^{-1} E_{\mathbf{u}}\left\{\partial_{\beta^\psi}\mathbf{H}_{u^\psi}(\boldsymbol{\theta}_0)\right\}\left[E_{\mathbf{u}}\left\{\partial_{u^\psi}\mathbf{H}_{u^\psi}(\boldsymbol{\theta}_0)\right\}\right]^{-1} \\ \mathbf{0} & \left[E_{\mathbf{u}}\left\{\partial_{u^\psi}\mathbf{H}_{u^\psi}(\boldsymbol{\theta}_0)\right\}\right]^{-1} \end{bmatrix},$$

the previous expressions lead to the sandwich-type estimators:

$$\begin{aligned} \hat{V}\left(\tilde{\boldsymbol{\beta}}^\psi\right) &= \left\{\hat{E}\left(\partial_{\beta^\psi}\mathbf{H}_{0\beta^\psi}\right)\right\}^{-1}\hat{V}\left\{\mathbf{H}_{0\beta^\psi}\right\}\left[\left\{\hat{E}\left(\partial_{\beta^\psi}\mathbf{H}_{0\beta^\psi}\right)\right\}^{-1}\right]^T \\ \hat{V}\left(\tilde{\mathbf{u}}^\psi\right) &= \left\{\hat{E}\left(\partial_{u^\psi}\mathbf{H}_{0u^\psi}\right)\right\}^{-1}\hat{V}\left\{\mathbf{H}_{0u^\psi}\right\}\left[\left\{\hat{E}\left(\partial_{u^\psi}\mathbf{H}_{0u^\psi}\right)\right\}^{-1}\right]^T, \end{aligned} \tag{27}$$

where

- $\hat{E}\left\{\partial_{\beta^\psi}\mathbf{H}_{0\beta^\psi}\right\} = -\mathbf{X}_s^T\mathbf{V}_s^{-1}\mathbf{U}_s^{1/2}\mathbf{R}\mathbf{U}_s^{-1/2}\mathbf{X}_s$;
- $\hat{E}\left\{\partial_{u^\psi}\mathbf{H}_{0u^\psi}\right\} = -\mathbf{Z}_s^T\boldsymbol{\Sigma}_{es}^{-1/2}\mathbf{T}\boldsymbol{\Sigma}_{es}^{-1/2}\mathbf{Z}_s - \boldsymbol{\Sigma}_u^{-1/2}\mathbf{Q}\boldsymbol{\Sigma}_u^{-1/2}$;
- $\hat{V}\left\{\mathbf{H}_{0\beta^\psi}\right\} = (n-p)^{-1}\lambda_1\sum_{j=1}^{n}\psi^2\left(r_j\right)\mathbf{X}_s^T\mathbf{V}_s^{-1}\mathbf{U}_s\mathbf{V}_s^{-1}\mathbf{X}_s$; and
- $\hat{V}\left\{\mathbf{H}_{0u^\psi}\right\} = (n-p)^{-1}\lambda_2\sum_{j=1}^{n}\psi^2(t_j)\mathbf{Z}_s^T\boldsymbol{\Sigma}_{es}^{-1}\mathbf{Z}_s$.

Note that the above estimators assume use of a Huber Proposal 2 influence function with tuning constant c, $\mathbf{R}$ is a $n\times n$ diagonal matrix with j-th diagonal element equal to 1 if $-c < r_j < c$, 0 otherwise, with $r_j = U_j^{-1/2}\left(y_j - \mathbf{x}_j^T\tilde{\boldsymbol{\beta}}^\psi\right)$. $\mathbf{T}$ is a diagonal matrix of dimension $n\times n$ with j-th diagonal element equal to 1 if $-c < t_j < c$, 0 otherwise, with $t_j = \left(\sigma_e^\psi\right)^{-1}\left(y_j - \mathbf{x}_j^T\tilde{\boldsymbol{\beta}}^\psi - \mathbf{z}_j^T\mathbf{u}^\psi\right)$ and $\mathbf{Q}$ is a $m\times m$ diagonal matrix with i-th diagonal element equal to 1 if $-c < q_i < c$, 0 otherwise, with $q_i = \left(\sigma_u^{2\psi}\right)^{-1/2}u_i^\psi$. The values

$$\lambda_1 = \left\{1 + pn^{-1}Var_{\mathbf{u}}\left(\psi'\left(r_j\right)\right)\left[E_{\mathbf{u}}\left(\psi'\left(r_j\right)\right)\right]^{-2}\right\} \text{ and } \lambda_2 = \left\{1 + pn^{-1}Var_{\mathbf{u}}\left(\psi'(\tilde{t}_i)\right)\left[E_{\mathbf{u}}\left(\psi'(\tilde{t}_i)\right)\right]^{-2}\right\}$$

are bias correction terms (Huber, 1981). From (24), an estimator of the conditional prediction variance of RBLUP can then be written as:

$$\hat{V}\left(\hat{\bar{y}}_i^{RBLUP} - \bar{y}_i\right) = h_{1i}(\tilde{\boldsymbol{\delta}}) + h_{2i}(\tilde{\boldsymbol{\delta}}) + h_{3i}(\tilde{\boldsymbol{\delta}}), \tag{28}$$

where

- $h_{1i}(\tilde{\boldsymbol{\delta}}) = \left(1 - n_iN_i^{-1}\right)^2\bar{\mathbf{z}}_{ri}^T\hat{V}\left(\tilde{\mathbf{u}}^\psi\right)\bar{\mathbf{z}}_{ri}$ is due to the estimation of random effects;
- $h_{2i}(\tilde{\boldsymbol{\delta}}) = \left(1 - n_iN_i^{-1}\right)^2\bar{\mathbf{x}}_{ri}^T\hat{V}\left(\tilde{\boldsymbol{\beta}}^\psi\right)\bar{\mathbf{x}}_{ri}$ is due to estimation of $\boldsymbol{\beta}$ by $\tilde{\boldsymbol{\beta}}^\psi$; and

- $h_{3i}(\tilde{\boldsymbol{\delta}}) = \left(1 - n_i N_i^{-1}\right)^2 \hat{V}\left(\bar{e}_{ri}\right)$ can be calculated just using the data from area i, so $\hat{V}(\bar{e}_{ri}) = (N_i - n_i)^{-1}(n_i - 1)^{-1}\sum_{j\in s_i}\left(y_j - \mathbf{x}_j^T\tilde{\boldsymbol{\beta}}^{\psi} - \mathbf{z}_j^T\tilde{\mathbf{u}}^{\psi}\right)^2$, or by using data from the entire sample, in which case $\hat{V}(\bar{e}_{ri}) = (N_i - n_i)^{-1}(n_i - 1)^{-1}\sum_h\sum_{j\in s_h}\left(y_j - \mathbf{x}_j^T\tilde{\boldsymbol{\beta}}^{\psi} - \mathbf{z}_j^T\tilde{\mathbf{u}}^{\psi}\right)$.

Finally, we add an estimator of the squared conditional bias to (28), leading to an estimator of the MSE of the RBLUP of the form:

$$\widehat{MSE}\left(\hat{\bar{y}}_i^{RBLUP}\right) = h_{1i}(\tilde{\boldsymbol{\delta}}) + h_{2i}(\tilde{\boldsymbol{\delta}}) + h_{3i}(\tilde{\boldsymbol{\delta}}) + \left\{\hat{B}\left(\hat{\bar{y}}_i^{RBLUP}\right)\right\}^2, \tag{29}$$

where $\hat{B}\left(\hat{\bar{y}}_i^{RBLUP}\right)$ is the estimator of the conditional bias defined following (23). The corresponding estimator of the MSE of the REBLUP (15) is defined by adding an extra term to (29) to account for the increased variability due to the estimation of the variance components:

$$\widehat{MSE}\left(\hat{\bar{y}}_i^{REBLUP}\right) = \widehat{MSE}\left(\hat{\bar{y}}_i^{RBLUP}\right) + \hat{E}\left[\left(\hat{\bar{y}}_i^{REBLUP} - \hat{\bar{y}}_i^{RBLUP}\right)^2\right]. \tag{30}$$

A Taylor approximation to the last term in (30) can be obtained following the approach by Prasad and Rao (1990). To start, we note that

$$\hat{\bar{y}}_i^{REBLUP} - \hat{\bar{y}}_i^{RBLUP} \approx \left(N_i^{-1}\sum_{j\in r_i}\mathbf{z}_j^T\right)\sum_{k=1}^{2}\left(\partial_{\theta_k}\mathbf{B}_s\right)\left(\mathbf{y}_s - \mathbf{X}_s\boldsymbol{\beta}^{\psi}\right)\left\{\hat{\theta}_k - \theta_k\right\},$$

where $\mathbf{B}_s$ has been defined before in this paper, and $\boldsymbol{\theta} = \left(\sigma_u^{2\psi}, \sigma_e^{2\psi}\right)$ is the vector of the variance components. Assuming that the derivative of $\left(\tilde{\boldsymbol{\beta}}^{\psi} - \boldsymbol{\beta}^{\psi}\right)$ with respect to $\boldsymbol{\theta}$ is of lower order, the second term on the right hand side of (30) is then approximated by

$$h_{4i}(\tilde{\boldsymbol{\delta}}) = \left(N_i^{-1}\sum_{j\in r_i}\mathbf{z}_j^T\right)\Upsilon Var(\hat{\sigma}_u^{\psi 2}, \hat{\sigma}_e^{\psi 2})\left(N_i^{-1}\sum_{j\in r_i}\mathbf{z}_j^T\right)^T + o\left(m^{-1}\right) \tag{31}$$

where

$$\Upsilon = \sum_{k=1}^{2}\left\{\left(\partial_{\sigma_u^{\psi 2},\sigma_e^{\psi 2}}\mathbf{B}_s\right)\left[\sum_j\sum_l\left\{\left(\mathbf{z}_j^T\tilde{\mathbf{u}}^{\psi}\right)\left(\mathbf{z}_l^T\tilde{\mathbf{u}}^{\psi}\right) + \sigma_e^{\psi 2}\mathbf{I}(j=l)\right\}\right]\left(\partial_{\sigma_u^{\psi 2},\sigma_e^{\psi 2}}\mathbf{B}_s\right)\right\}.$$

Note that $Var(\hat{\sigma}_u^{\psi 2}, \hat{\sigma}_e^{\psi 2})$ in expression (31) is obtained using the results for the asymptotic distribution of $\left(\sigma_u^{2\psi}, \sigma_e^{2\psi}\right)$ given in Sinha and Rao (2009). Consequently, (30) can be approximated by:

$$\widehat{MSE}\left(\hat{\bar{y}}_i^{REBLUP}\right)=h_{1i}(\tilde{\boldsymbol{\delta}})+h_{2i}(\tilde{\boldsymbol{\delta}})+h_{3i}(\tilde{\boldsymbol{\delta}})+h_{4i}(\tilde{\boldsymbol{\delta}})+\left\{\hat{B}\left(\hat{\bar{y}}_i^{RBLUP}\right)\right\}^2. \quad (32)$$

An estimator of the conditional MSE of the REBLUP is obtained by replacing $\tilde{\boldsymbol{\delta}}=\left(\tilde{\boldsymbol{\beta}}^{\psi T},\tilde{\mathbf{u}}^{\psi T}\right)^T$ by $\hat{\boldsymbol{\delta}}=\left(\hat{\boldsymbol{\beta}}^{\psi T},\hat{\mathbf{u}}^{\psi T}\right)^T$ in (32) and leads to:

$$mse\left(\hat{\bar{y}}_i^{REBLUP}\right)=h_{1i}(\hat{\boldsymbol{\delta}})+h_{2i}(\hat{\boldsymbol{\delta}})+h_{3i}(\hat{\boldsymbol{\delta}})+h_{4i}(\hat{\boldsymbol{\delta}})+\left\{\hat{B}\left(\hat{\bar{y}}_i^{REBLUP}\right)\right\}^2. \quad (33)$$

Note that $E\left[h_{2i}(\hat{\boldsymbol{\delta}})\right]=h_{2i}(\boldsymbol{\delta})+o\left(m^{-1}\right)$, $E\left[h_{3i}(\hat{\boldsymbol{\delta}})\right]=h_{3i}(\boldsymbol{\delta})+o\left(m^{-1}\right)$ and $E\left[h_{4i}(\hat{\boldsymbol{\delta}})\right]=h_{4i}(\boldsymbol{\delta})+o\left(m^{-1}\right)$. However, we can improve upon $h_{1i}(\hat{\boldsymbol{\delta}})$ as an estimator of $h_{1i}(\boldsymbol{\delta})$ because its bias is generally of the same order as $h_{2i}(\hat{\boldsymbol{\delta}})$, $h_{3i}(\hat{\boldsymbol{\delta}})$ and $h_{4i}(\hat{\boldsymbol{\delta}})$. We use a Taylor series expansion of $h_{1i}(\hat{\boldsymbol{\delta}})$ around $\boldsymbol{\theta}=\left(\sigma_u^{\psi 2},\sigma_e^{\psi 2}\right)$ to evaluate the bias of $h_{1i}(\hat{\boldsymbol{\delta}})$:

$$h_{1i}(\hat{\boldsymbol{\delta}})=h_{1i}(\boldsymbol{\delta})+\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right)^T\nabla h_{1i}\left(\tilde{\boldsymbol{\delta}}\right)+\frac{1}{2}\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right)^T\nabla^2 h_{1i}\left(\boldsymbol{\delta}\right)\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right)=h_{1i}(\boldsymbol{\delta})+\Delta_1+\Delta_2.$$

If $\hat{\boldsymbol{\theta}}$ is unbiased for $\boldsymbol{\theta}$ then $E\left[\Delta_1\right]=0$. In general, if $\hat{\boldsymbol{\theta}}$ is biased, $E\left[\Delta_1\right]$ is of lower order than $E\left[\Delta_2\right]$, and so

$$\begin{aligned}E\left[h_{1i}(\hat{\boldsymbol{\delta}})\right]&\approx h_{1i}(\boldsymbol{\delta})+\frac{1}{2}\left(1-n_iN_i^{-1}\right)^2 tr\left\{\bar{\mathbf{z}}_{ri}^T\nabla^2 h_{1i}(\boldsymbol{\delta})\bar{\mathbf{z}}_i Var(\hat{\sigma}_u^{2\psi},\hat{\sigma}_e^{2\psi})\right\}+o\left(m^{-1}\right)\\&=h_{1i}(\boldsymbol{\delta})+h_{5i}(\boldsymbol{\delta})+o\left(m^{-1}\right).\end{aligned}$$

The proposed estimator of the conditional MSE of the REBLUP is therefore

$$mse\left(\hat{\bar{y}}_i^{REBLUP}\right)=h_{1i}(\hat{\boldsymbol{\delta}})+h_{2i}(\hat{\boldsymbol{\delta}})+h_{3i}(\hat{\boldsymbol{\delta}})+h_{4i}(\hat{\boldsymbol{\delta}})+h_{5i}(\hat{\boldsymbol{\delta}})+\left\{\hat{B}\left(\hat{\bar{y}}_i^{REBLUP}\right)\right\}^2. \quad (34)$$

Note that an estimate of the conditional MSE of the EBLUP is easily calculated by setting the tuning constant for the influence function in (34) so that no outlier modification occurs, e.g. setting $c > 100$.

We take a similar approach to defining an estimator of the conditional MSE of the REBLUP-BC. To start, we develop an approximation to the conditional prediction variance of this predictor when the variance components are known, i.e. for the RBLUP-BC. In this case the prediction error is

$$\hat{\bar{y}}_i^{RBLUP-BC}-\bar{y}_i=\left(1-n_iN_i^{-1}\right)\left\{\left(\bar{\mathbf{x}}_{ri}^T\tilde{\boldsymbol{\beta}}^{\psi}+\bar{\mathbf{z}}_{ri}^T\tilde{\mathbf{u}}^{\psi}\right)+n_i^{-1}\sum_{j\in s_i}\omega_{ij}^{\psi}\phi\left(\frac{y_j-\mathbf{x}_j^T\tilde{\boldsymbol{\beta}}^{\psi}-\mathbf{z}_j^T\tilde{\mathbf{u}}^{\psi}}{\omega_{ij}^{\psi}}\right)-\bar{y}_{ri}\right\}.$$

The second (BC) term inside the braces on the right hand side of this expression can be expanded using a

Taylor series approximation. When the tuning constant used in ϕ is large, so $\phi' \approx 1$, this approximation becomes

$$n_i^{-1}\sum_{j\in s_i}\omega_{ij}^{\psi}\phi\left(\frac{y_j-\mathbf{x}_j^T\tilde{\boldsymbol{\beta}}^{\psi}-\mathbf{z}_j^T\tilde{\mathbf{u}}^{\psi}}{\omega_{ij}^{\psi}}\right)\approx n_i^{-1}\sum_{j\in s_i}\omega_{ij}^{\psi}\phi\left(\frac{y_j-\mathbf{x}_j^T\boldsymbol{\beta}^{\psi}-\mathbf{z}_j^T\mathbf{u}^{\psi}}{\omega_{ij}^{\psi}}\right)-\left(\tilde{\boldsymbol{\beta}}^{\psi}-\boldsymbol{\beta}^{\psi}\right)^T\bar{\mathbf{x}}_{si}-\left(\tilde{\mathbf{u}}^{\psi}-\mathbf{u}^{\psi}\right)^T\bar{\mathbf{z}}_{si}.$$

Substituting in the preceding expression for the prediction error of the RBLUP-BC leads to

$$\hat{\bar{y}}_i^{RBLUP-BC}-\bar{y}_i\approx\left(1-n_iN_i^{-1}\right)\left\{\left(\bar{\mathbf{x}}_{si}^T\boldsymbol{\beta}^{\psi}+\bar{\mathbf{z}}_{si}^T\mathbf{u}^{\psi}\right)+T_{1i}+T_{2i}-\bar{y}_{ri}\right\}, \tag{35}$$

where $T_{1i}=n_i^{-1}\sum_{j\in s_i}\omega_{ij}^{\psi}\phi\left(\frac{y_j-\mathbf{x}_j^T\boldsymbol{\beta}^{\psi}-\mathbf{z}_j^T\mathbf{u}^{\psi}}{\omega_{ij}^{\psi}}\right)$ and $T_{2i}=\left(\bar{\mathbf{x}}_{ri}-\bar{\mathbf{x}}_{si}\right)^T\tilde{\boldsymbol{\beta}}^{\psi}+\left(\bar{\mathbf{z}}_{ri}-\bar{\mathbf{z}}_{si}\right)^T\tilde{\mathbf{u}}^{\psi}$. Since the covariance between T_{1i} and T_{2i} should be of a lower order of magnitude than either of their variances, we can write down an estimator of the conditional variance of the RBLUP-BC of the form

$$\widehat{MSE}\left(\hat{\bar{y}}_i^{RBLUP-BC}\right)=h_{1i}^{BC}(\tilde{\boldsymbol{\delta}})+h_{2i}^{BC}(\tilde{\boldsymbol{\delta}})+h_{3i}^{BC}(\tilde{\boldsymbol{\delta}})+h_{4i}^{BC}(\tilde{\boldsymbol{\delta}}), \tag{36}$$

where

- $h_{1i}^{BC}(\tilde{\boldsymbol{\delta}})=\left(1-n_iN_i^{-1}\right)^2\left\{\left(\bar{\mathbf{z}}_{ri}-\bar{\mathbf{z}}_{si}\right)^T\hat{V}\left(\tilde{\mathbf{u}}^{\psi}\right)\left(\bar{\mathbf{z}}_{ri}-\bar{\mathbf{z}}_{si}\right)\right\}$;
- $h_{2i}^{BC}(\tilde{\boldsymbol{\delta}})=\left(\bar{\mathbf{x}}_{ri}-\bar{\mathbf{x}}_{si}\right)^T\hat{V}\left(\tilde{\boldsymbol{\beta}}^{\psi}\right)\left(\bar{\mathbf{x}}_{ri}-\bar{\mathbf{x}}_{si}\right)$;
- $h_{3i}^{BC}(\tilde{\boldsymbol{\delta}})=\hat{V}\left(\bar{e}_{ri}\right)$; and
- $h_{4i}^{BC}(\tilde{\boldsymbol{\delta}})=n_i^{-1}(n_i-p)^{-1}\sum_{j\in s_i}\left\{\omega_{ij}^{\psi}\phi\left(\frac{y_j-\mathbf{x}_j^T\tilde{\boldsymbol{\beta}}^{\psi}-\mathbf{z}_j^T\tilde{\mathbf{u}}^{\psi}}{\omega_{ij}^{\psi}}\right)\right\}^2$.

The estimator of the conditional MSE of the REBLUP-BC is then obtained by adding a term to (36) to account for the additional uncertainty due to estimation of the variance components. The same approach as already used for the REBLUP can be used to obtain the approximation

$$\begin{aligned}h_{5i}^{BC}(\tilde{\boldsymbol{\delta}})&=E\left[\left(\hat{\bar{y}}_i^{REBLUP-BC}-\hat{\bar{y}}_i^{RBLUP-BC}\right)^2\right]\\&\approx\left(1-n_iN_i^{-1}\right)^2\mathbf{D}_i^TVar\left\{\sum_{k=1}^{2}\left(\partial_{\theta_k}\mathbf{B}_s\right)\left(\mathbf{y}_s-\mathbf{X}_s\boldsymbol{\beta}^{\psi}\right)\left\{\hat{\theta}_k-\theta_k\right\}\right\}\mathbf{D}_i+o\left(m^{-1}\right),\end{aligned} \tag{37}$$

where $\mathbf{D}_i=\bar{\mathbf{z}}_{ri}-n_i^{-1}\sum_{j\in s_i}\phi'\left(\frac{y_j-\mathbf{x}_j^T\boldsymbol{\beta}^{\psi}-\mathbf{z}_j^T\mathbf{B}_s\left(\mathbf{y}_s-\mathbf{X}_s\boldsymbol{\beta}^{\psi}\right)}{\omega_{ij}^{\psi}}\right)\mathbf{z}_j$. Note that $\mathbf{D}_i=\mathbf{0}$ when ϕ is the identity function, e.g. as in the version of BC described in Chambers *et al.* (1993), and the model only contains random intercepts. The resulting estimator of the conditional MSE of the REBLUP-BC is then

$$mse\left(\hat{\bar{y}}_i^{REBLUP-BC}\right) = h_{1i}^{BC}(\hat{\boldsymbol{\delta}}) + h_{2i}^{BC}(\hat{\boldsymbol{\delta}}) + h_{3i}^{BC}(\hat{\boldsymbol{\delta}}) + h_{4i}^{BC}(\hat{\boldsymbol{\delta}}) + h_{5i}^{BC}(\hat{\boldsymbol{\delta}}) + h_{6i}^{BC}(\hat{\boldsymbol{\delta}}). \quad (38)$$

Note that like $h_{1i}(\hat{\boldsymbol{\delta}})$ in expression (34) $h_{1i}^{BC}(\hat{\boldsymbol{\delta}})$ is also biased and $h_{6i}^{BC}(\hat{\boldsymbol{\delta}})$ is its bias correction term. This term is computed by using a Taylor expansion similar to $h_{5i}(\hat{\boldsymbol{\delta}})$ in expression (34). As with the estimator of the conditional MSE of the REBLUP-BC based on the pseudo-linearization approach, no squared conditional bias estimator is used with (38). Estimators of the conditional MSE for the MQ and MQ-BC predictors can be obtained similarly (see the Appendix).

5. Model-Based Simulations

We provide model-based simulation results illustrating the performances of the different outlier robust small area predictors and of the corresponding MSE estimators described in Sections 3 and 4. Population data are generated for $m = 40$ small areas, with samples selected by simple random sampling without replacement within each area. Population and sample sizes are the same for all areas, and are fixed at either $N_i = 100, n_i = 5$ or $N_i = 300, n_i = 15$. Values for *x* are generated as independently and identically distributed from a lognormal distribution with a mean of 1.0 and a standard deviation of 0.5 on the log scale. Values for *Y* are generated as $y_{ij} = 100 + 5x_{ij} + u_i + \varepsilon_{ij}$, where the random area and individual effects are independently generated according to four scenarios:

- [0,0] – No outliers: $u \sim N(0,3)$ and $\varepsilon \sim N(0,6)$.
- [e,0] – Individual outliers only: $u \sim N(0,3)$ and $\varepsilon \sim \delta N(0,6) + (1-\delta)N(20,150)$, where δ is an independently generated Bernoulli random variable with $\Pr(\delta = 1) = 0.97$, i.e. the individual effects are independent draws from a mixture of two normal distributions, with 97% on average drawn from a ‘well-behaved’ $N(0,6)$ distribution and 3% on average drawn from an outlier $N(20,150)$ distribution.
- [0,u] – Area outliers only: $u \sim N(0,3)$ for areas 1-36, $u \sim N(9,20)$ for areas 37-40 and $\varepsilon \sim N(0,6)$, i.e. random effects for areas 1–36 are drawn from a ‘well behaved’ $N(0,3)$ distribution, with those for

areas 37–40 drawn from an outlier $N(9,20)$ distribution. Individual effects are not outlier-contaminated.

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

Each scenario is independently simulated 500 times. For each simulation the population values are generated according to the underlying scenario, a sample is selected in each area and the sample data are then used to compute estimates of each of the actual area means for *y*.

Five different estimators are used for this purpose - the standard EBLUP, see (2), which serves as a reference; the projective M-quantile estimator MQ, see (17); the robust bias-corrected predictive MQ estimator MQ-BC, see (20); the robust projective REBLUP estimator of Sinha and Rao (2009), see (15); and its robust bias-corrected version REBLUP-BC, see (21). In all cases the 'projective' influence function ψ is a Huber Proposal 2 type with tuning constant $c=1.345$. In contrast, the 'predictive', less restrictive, influence function ϕ used in MQ-BC and REBLUP-BC is also a Huber Proposal 2 type, but with a larger tuning constant, $c=3$.

The performance of these estimators across the different areas and simulations is assessed by computing the median values of their area specific relative bias and relative root mean squared error, where the relative bias of an estimator $\hat{\bar{y}}_i$ for the actual mean $\bar{y}_i$ of area *i* is the average across simulations of the errors $\hat{\bar{y}}_i-\bar{y}_i$ divided by the corresponding average value of $\bar{y}_i$, and its relative root mean squared error is the square root of the average across simulations of the squares of these errors, again divided by the average value of $\bar{y}_i$. Table 1 presents these median values for the different simulation scenarios and different estimators.

The relative bias results set out in Table 1 confirm our expectations regarding the behaviour of the projective estimators (EBLUP, REBLUP and MQ) and the bias-corrected predictive estimators (REBLUP-BC and MQ-BC). The former are more biased than the latter (see scenarios with area and individual outliers) as a consequence of their implicit assumption that although outlier variances may be inflated relative to non-outliers, outlier effects still have zero expectation. This increase in bias is most pronounced when there are outliers in the area effects, which is not unexpected since that is when area means are most affected by the presence of outliers in the population data. Turning to the median RRMSE results, we see that claims in the literature (e.g. Chambers and Tzavidis, 2006) about the superior outlier robustness of MQ compared with the EBLUP certainly hold true – provided the outliers are in individual effects. If there are outliers in area effects, then MQ appears to offer no extra protection compared to the EBLUP, and in fact performs worse, mainly due to its sharply increasing bias in this situation. Similarly, when we compare the EBLUP and the REBLUP we see that if outliers are associated with individual effects, then the REBLUP offers better RRMSE performance than the EBLUP. However, the gap between these two estimators narrows considerably when outliers are associated with area effects. In contrast, the two bias-corrected predictive estimators seem relatively robust in terms of RRMSE performance. Nevertheless, due to the increased variability as a consequence of their bias corrections, both BC estimators are not as efficient as the projective estimators when outliers are associated with individual effects, but both also do not fail when there are outliers in the area effects. Finally, the REBLUP-BC estimator appears to be performing better than the MQ-BC estimator for those scenarios where the use of predictive estimators offers gains.

We now examine the performance of the different MSE estimators. Here, we are mainly interested in the performance of alternative MSE estimators for estimating the MSE of the robust predictive estimators. However, we do also comment on the performance of the different MSE estimators when used for estimating the MSE of projective estimators under a range of scenarios.

MSE estimation for the REBLUP and REBLUP-BC is implemented via the pseudo-linearization MSE estimator (23) (hereafter CCT) and via the linearization-based MSE estimators (34) and (38) (hereafter CCST). For the MQ and the MQ-BC the MSE estimators (A5) and (A7), which correspond to the CCST (see the Appendix for details), and the CCT are used (see Chambers et al. 2009 for details). For the REBLUP we further used the bootstrap procedure proposed by Sinha and Rao (2009), which is implemented by generating 100 bootstrap samples in each Monte Carlo run. Finally, the MSE of the EBLUP is estimated by using the Prasad-Rao (PR) estimator, but in addition we also evaluate the performance of the CCT and the CCST, obtained as a special case of (34), for estimating the MSE of this estimator. The results of the MSE estimators for each scenario and for each estimator are shown in Table 2 where we report the median values of their area specific relative bias and the relative root mean squared error.

We start first by evaluating the performance of the MSE estimators we proposed in this paper (CCT and CCST) for estimating the MSE of the robust predictive estimators REBLUP-BC and MQ-BC. We note while for the REBLUP-BC the CCST has a somewhat better performance, for the MQ-BC the CCT performs better in terms of bias. Although the RRMSEs of both MSE estimators have similar orders of magnitude, it appears that the CCST is less stable than the CCT when used to estimate the MSE of the REBLUP-BC but the reverse is true for the estimation of the MSE of the MQ-BC. Perhaps an improvement in the MSE estimation of the REBLUP-BC can be offered by using the parametric bootstrap MSE proposed by Sinha and Rao (2009). This estimator has an advantage both in terms of relative bias and RRMSE. However, for the case of the MQ-BC our only options are currently the CCT and CCST estimators.

Turning now to MSE estimation for the robust projective estimators we note the following. For the REBLUP estimator the parametric bootstrap MSE provides a very good approximation to the true MSE. Having said this, the CCST also provides a good alternative in this case. For example, in the case of outliers both at area level and at individual level the CCST records the lowest relative bias and the lowest RRMSE. For the MQ estimator the CCST estimator performs better than the CCT estimator. However, for the scenario with outliers at both the area and the unit level we observe that both CCT and CCST estimator record very high relative biases. Nevertheless, in the case of this scenario we should use a robust projective estimator in which case both the bootstrap or the CCST MSE estimators perform well. For the case of the EBLUP estimator, the PR estimator performs, as expected, well in the [0,0] scenario and also records small relative bias for some of the scenarios with outliers. The PR MSE estimator is also more stable than the CCT and the CCST for almost all scenarios. The CCT estimator has an impressively small relative bias for all scenarios. However, as pointed out by Chambers *et al*. (2009), the bias robustness of this MSE estimator comes at the price of high variability especially in the case of very small area sample sizes. Finally, the CCST performs worse in terms of relative bias but is more stable than the CCT.

6. Design-Based Simulation

Design-based simulations complement model-based simulations for SAE since they allow us to evaluate the performance of SAE methods in the context of a real population and realistic sampling methods where we do not know the precise source of contamination. From a finite population perspective we believe that this type of simulation constitutes a more practical and appropriate representation of the SAE problem. Furthermore, it provides a good illustration of why a focus on conditional MSE is likely to be closer to the MSE of interest for analysts using small area methods.

The population underpinning the design-based simulation is based on a data set obtained under the Environmental Monitoring and Assessment Program (EMAP) of the U.S. Environmental Protection Agency. The background to this data set is that between 1991 and 1995 EMAP conducted a survey of lakes in the North-Eastern states of the U.S. The data collected in this survey consists of 551 measurements from a sample of 334 of the 21,026 lakes located in this area. The lakes making up this population are grouped into 113 8-digit Hydrologic Unit Codes (HUCs), of which 64 contained less than 5 observations and 27 did not have any observations. In our simulation, we defined HUCs as the small areas of interest, with lakes grouped within HUCs. The variable of interest is Acid Neutralizing Capacity (ANC), an indicator of the acidification risk of water bodies. A total of 1000 independent random samples of lake locations are then taken from the population of 21,026 lake locations by randomly selecting locations in the 86 HUCs that contained EMAP sampled lakes, with sample sizes in these HUCs set to the greater of five and the original EMAP sample size. A two-level (level 1 is the lake and level 2 is the HUC) mixed model has been fitted to the synthetic population data. The Shapiro-Wilk normality test, which rejects the null hypothesis that the residuals follow a normal distribution (p-values: level 1 = 0.03555, level 2 = 1.715e-14), indicates that the Gaussian assumptions of the mixed model are not met. Using a model that relaxes these assumptions, such as an M-quantile model with a bounded influence function, therefore seems reasonable for these data. Details on the exact data generation mechanism and the characteristics of the population can be found in Salvati *et al.* (2011).

Table 3 shows the median relative bias and the median relative root MSE of the different predictors (EBLUP, REBLUP, MQ, REBLUP-BC, MQ-BC) and Table 4 reports the median relative bias and the median relative root MSE of the corresponding estimators of the MSE of these predictors calculated using the same sample as the one used to compute the small area point estimators. The MQ-BC and REBLUP-BC predictors work well both in terms of bias and RRMSE, while the EBLUP and MQ have the highest RRMSE, with the MQ also recording the largest negative bias. The REBLUP shows a good performance in terms of RRMSE but records a large negative bias. These results suggest that the proposed Robust predictive estimators offer in this case the most balanced performance both in terms of bias and MSE.

We now examine the performance of the different methods of MSE estimation. The behaviour of the 'true' (empirical) root MSE and of its estimators for each area and for each approach is depicted in Figures 1, 2 and 3. Examination of these results can be useful for understanding the reasons as to why the MSE estimators perform differently.

Figure 1 shows the results for the EBLUP predictor. We start by noting that the PR estimator does not capture the between area differences in the design-based RMSE of the EBLUP, while the CCT MSE estimator for the EBLUP tracks the irregular profile of the area-specific empirical MSE very well. The CCST also works reasonably well but produces somewhat over-smoothed estimates of the area-specific empirical MSEs. These results confirm the poor design-based properties of the PR estimator (Longford, 2007). Figure 2 presents the results for the REBLUP and REBLUP-BC predictors. For REBLUP (top plot) it is evident that CCT tends to underestimate the 'true' area-specific MSE, mainly because its squared bias component underestimates the actual squared bias of this predictor. The bootstrap MSE estimator produces over-smoothed estimates of the area-specific 'true' MSEs. The CCST estimator tracks the area-specific empirical MSE but shows underestimation in a few areas. For the REBLUP-BC, the CCST MSE estimator (bottom plot) tracks the irregular profile of the area-specific empirical MSE very well, while the bootstrap MSE estimator for the REBLUP-BC generates over-smoothed estimates of the area-specific empirical MSE. Finally, Figure 3 illustrates the results for MQ (top plot) and MQ-BC (bottom plot) predictors. The MSE estimators have a similar behaviour. They track the irregular profile of the area-specific 'true' MSE very well for MQ-BC but for MQ both the CCT and CCST underestimate the 'true' area-specific MSE.

By combining the results in Table 4 with the results shown in Figures 1, 2 and 3 we conclude that the CCST and the bootstrap MSE estimators offer two good alternatives for estimating the MSE of the REBLUP-BC whereas the CCST and CCT estimators estimate reasonably well the MSE of the MQ-BC estimator. The CCST estimator is also a very good alternative to the bootstrap estimator when the target is to estimate the MSE of the REBLUP. These conclusions are in line with the conclusions we reached by analysing the results from the model-based simulation.

Finally, we note that one could combine the linear mixed and M-quantile model-based estimators with the MSE estimation method described in Section 4 to generate ‘normal theory’ confidence intervals for the small area means. Coverage results based on such intervals have been produced and are available from the authors. However, this use of the estimated MSE to construct confidence intervals, though widespread, has been criticised. Chatterjee *et al*. (2008) discuss the use of bootstrap methods for constructing confidence intervals for small area parameters and argue that there is no guarantee that the asymptotic behaviour underpinning normal theory confidence intervals extends to small area estimation.

7. Final Remarks

In this paper we explore the extension of the Robust Predictive approach to SAE and we propose two analytical mean squared error (MSE) estimators for outlier robust predictors of small area means. The first is a bias-robust MSE estimator that is based on the 'pseudo-linearization' approach discussed by Chambers *et al*. (2009). The second method is a linearization-based MSE estimation that is based on first order approximations to the variances of solutions of estimating equations.

The empirical results we report in Sections 5 and 6 show that the bias-corrected predictive estimators (REBLUP-BC and MQ-BC) are less biased and can be more efficient than the projective estimators (EBLUP, REBLUP and MQ) in the presence of area and individual outliers. What is also evident from these results is that the bias correction of the predictive estimators comes at the cost of higher variability. As a result we expect that the use of the predictive estimators will pay dividends only when the use of model diagnostics suggest that there are significant departures from the assumed working small area model. One approach for controlling the bias-variance trade off when using the robust predictive approaches is by selecting optimal tuning constants c and ϕ to be used for computing these estimators. This can be potentially achieved by using a cross-validation criterion. Defining this cross validation criterion is an avenue for future research.

The pseudo-linearization and linearization-based MSE estimators we described in Section 4 and in particular the CCST, offer a good approach for estimating the MSE of the robust predictive estimators. Together with the parametric bootstrap MSE estimator, pseudo-linearization and linearization-based MSE estimators present a collection of MSE estimators that can be used for estimating the MSE of alternative robust small area predictors. Finally, the CCST is developed under a conditional version of the linear mixed model. It should be possible though to develop an unconditional version of the CCST MSE estimator that averages over the distribution of the random area effects under a linear mixed model, and so reduces to the Prasad-Rao MSE estimator in the case of the EBLUP. This presents an additional avenue for further research.

Acknowledgements

Work partially supported by the SAMPLE project (SSH-CT-2007-217565; FP7-SSH-2007-1) and by ARC Linkage Grant LP0776810. The authors are grateful to Space-Time Aquatic Resources Modeling and Analysis Program (STARMAP) for data availability. The views expressed here are solely those of the authors.

Appendix: Linearization based MSE estimation for MQ and MQ-BC

For fixed q, the prediction variance of the MQ predictor (17) based on the M-quantile approach is

$$Var(\hat{\bar{y}}_i^{MQ} - \bar{y}_i) = \left(1 - n_i N_i^{-1}\right)^2 \left\{ \bar{\mathbf{x}}_{ri}^T Var(\hat{\boldsymbol{\beta}}_{\bar{q}_i}) \bar{\mathbf{x}}_{ri} \right\} + \left(1 - n_i N_i^{-1}\right)^2 Var(\bar{e}_{ri}) . \quad \text{(A1)}$$

It follows that we need to estimate $Var(\hat{\boldsymbol{\beta}}_{\bar{q}_i})$ in order to be able to calculate an estimate of the prediction variance of this predictor. The starting point, as usual, is the first order approximation based on the estimating equations for $\hat{\boldsymbol{\beta}}_{\bar{q}_i}$. Putting $q = \bar{q}$,

$$Var_0\left(\hat{\boldsymbol{\beta}}_q\right) \approx \left\{ E_0\left(\partial_{\beta_q} \mathbf{H}_0\right) \right\}^{-1} Var_0\left\{ \mathbf{H}\left(\boldsymbol{\beta}_{0q}\right) \right\} \left[\left\{ E_0\left(\partial_{\beta_q} \mathbf{H}_0\right) \right\}^{-1} \right]^T , \quad \text{(A2)}$$

with $\mathbf{H}(\boldsymbol{\beta}_{0q}) = \sum_{j=1}^{n} \mathbf{x}_j \psi_q(r_j) = \mathbf{X}_s^T \psi_q(r_{0q})$, where ψ_q is a bounded influence function depending on q, $\psi_q(r_{0q})$ is the n-vector with elements $\psi_q(r_{j0q}) = \psi_q\left\{ \omega_{j0q}^{-1}\left(y_j - \mathbf{x}_j^T \boldsymbol{\beta}_{0q} \right) \right\}$ and ω_{j0q} is a robust estimator of the scale of the residual $y_j - \mathbf{x}_j^T \boldsymbol{\beta}_{0q}$. The $Var_0\left\{ \mathbf{H}(\beta_{0q}) \right\}$ component of (A2) can then be written as

$$Var_0\left\{ \mathbf{H}(\beta_{0q}) \right\} = \mathbf{X}_s^T \left\{ E_0\left\{ \psi_q(r_{0q}) \psi_q^T(r_{0q}) \right\} \right\} \mathbf{X}_s ,$$

because the y values are conditionally uncorrelated and $E_0\left\{ \psi_q(r_{0q}) \right\} = 0,\ \forall\, q$. Assuming a Huber-type influence function, we obtain

$$E_0\left(\partial_{\beta_q}\mathbf{H}_{0q}\right)=\mathbf{X}_s^T E_0\left[2\frac{d}{d\boldsymbol{\beta}_q}\psi_q(r_{0q})\Big|_{\boldsymbol{\beta}_q=\boldsymbol{\beta}_{0q}}\right]=-2\mathbf{X}_s^T\mathbf{C}\mathbf{X}_s\,,$$

where $\mathbf{C}$ is a $n\times n$ diagonal matrix with *j*-th diagonal component

$$\omega_{j0q}^{-1}E_{0q}\left\{qI\left(0<r_{j0q}\le c\right)+(1-q)I\left(-c<r_{j0q}\le 0\right)\right\}.$$

These expressions then lead to the estimators:

$$\widehat{Var}(\hat{\boldsymbol{\beta}}_q)=(n-p)^{-1}\left\{\sum_{j\in s}\psi_q^2\left(\hat{r}_{jq}\right)\right\}\left[n^{-1}\sum_{j\in s}\psi_q'\left(\hat{r}_{jq}\right)\right]^{-2}\left(\mathbf{X}_s^T\mathbf{X}_s\right)^{-1}, \qquad \text{(A3)}$$

where $\hat{r}_{jq}=\hat{\omega}_{jq}^{-1}\left(y_j-\mathbf{x}_j^T\hat{\boldsymbol{\beta}}_q\right)$. That is, the Street *et al*. (1988) estimator when $q=0.5$.

The estimator of the prediction variance of the MQ predictor when $q=\overline{q}$ can be written as:

$$\hat{V}\left(\hat{\bar{y}}_i^{MQ}\right)=\left(1-n_iN_i^{-1}\right)^2\left\{\overline{\mathbf{x}}_{ri}^T\widehat{Var}\left(\hat{\boldsymbol{\beta}}_{\overline{q}_i}\right)\overline{\mathbf{x}}_{ri}\right\}+\left(1-n_iN_i^{-1}\right)^2\hat{V}\left(\overline{e}_{ri}\right) \qquad \text{(A4)}$$

with $\hat{V}(\overline{e}_{ri})=(N_i-n_i)^{-1}(n-1)^{-1}\sum_h\sum_{j\in s_h}\left(y_j-\mathbf{x}_j^T\hat{\boldsymbol{\beta}}_{\overline{q}_h}\right)^2$. Moreover, since we are taking a conditional approach, we have to add an estimator of the squared bias based on:

$$\hat{B}\left(\hat{\bar{y}}_i^{MQ}\right)=N_i^{-1}\left\{\sum_k\sum_{j\in s_k}w_j\mathbf{x}_j^T\hat{\boldsymbol{\beta}}_{\overline{q}_k}-\sum_{j\in i}\mathbf{x}_j^T\hat{\boldsymbol{\beta}}_{\overline{q}_i}\right\},\text{ where } w_j=\begin{cases}\frac{N_i}{n_i}+b_j & \text{if } j\in i\\ b_j & \text{otherwise}\end{cases}$$

and $\mathbf{b}^T=\left(\sum_{j\in r_i}\mathbf{x}_j^T-\frac{N_i-n_i}{n_i}\sum_{j\in s_i}\mathbf{x}_j^T\right)(\mathbf{X}_s^T\mathbf{W}(\overline{q}_i)\mathbf{X}_s)^{-1}\mathbf{X}_s^T\mathbf{W}(\overline{q}_i)$ is a $1\times n$ vector. The final expression for the estimator of the MSE of the MQ predictor is therefore:

$$mse\left(\hat{\bar{y}}_i^{MQ}\right)=\left(1-n_iN_i^{-1}\right)^2\left\{\overline{\mathbf{x}}_{ri}^T\hat{V}\left(\hat{\boldsymbol{\beta}}_{\overline{q}_i}\right)\overline{\mathbf{x}}_{ri}\right\}+\left(1-n_iN_i^{-1}\right)^2\hat{V}\left(\overline{e}_{ri}\right)+\left\{N_i^{-1}\left\{\sum_k\sum_{j\in s_k}w_j\mathbf{x}_j^T\hat{\boldsymbol{\beta}}_{\overline{q}_k}-\sum_{j\in i}\mathbf{x}_j^T\hat{\boldsymbol{\beta}}_{\overline{q}_i}\right\}\right\}^2 .\text{(A5)}$$

Note that expression (A5) is a first order approximation to the asymptotic prediction variance of the MQ predictor, and so it could underestimate its MSE. The estimation or prediction error for MQ-BC is

$$\hat{\bar{y}}_i^{MQ-BC}-\bar{y}_i=N_i^{-1}\sum_{j\in r_i}\mathbf{x}_j^T\hat{\boldsymbol{\beta}}_{\overline{q}_i}+\frac{N_i-n_i}{N_in_i}\sum_{j\in s_i}\omega_{ij}^{\psi}\phi\left(\frac{y_j-\mathbf{x}_j^T\hat{\boldsymbol{\beta}}_{\overline{q}_i}}{\omega_{ij}^{\psi}}\right)-N_i^{-1}\sum_{j\in r_i}y_j.$$

We can write

$$\frac{1}{n_i}\sum_{j\in s_i}\omega_{ij}^{\psi}\phi\left(\frac{y_j-\mathbf{x}_j^T\hat{\boldsymbol{\beta}}_{\bar{q}_i}}{\omega_{ij}^{\psi}}\right)\approx\frac{1}{n_i}\sum_{j\in s_i}\omega_{ij}^{\psi}\phi\left(\frac{y_j-\mathbf{x}_j^T\boldsymbol{\beta}_{\bar{q}_i}}{\omega_{ij}^{\psi}}\right)+\left(\hat{\boldsymbol{\beta}}_{\bar{q}_i}-\boldsymbol{\beta}_{\bar{q}_i}\right)^T\partial_{\beta_{\bar{q}_i}}\left\{\frac{1}{n_i}\sum_{j\in s_i}\omega_{ij}^{\psi}\phi\left(\frac{y_i-\mathbf{x}_i^T\boldsymbol{\beta}_{\bar{q}_i}}{\omega_{ij}^{\psi}}\right)\right\}$$

and if the tuning constant used in the BC term is large, $\phi'\approx 1$ and covariance between the first and second terms on the right hand side should be of a lower order of magnitude than either of their variances, so

$$Var\left(\hat{\bar{y}}_i^{MQ-BC}-\bar{y}_i\right)=\left(1-n_iN_i^{-1}\right)^2\left[\{\bar{\mathbf{x}}_{ri}-\bar{\mathbf{x}}_{si}\}^T Var\left(\hat{\boldsymbol{\beta}}_{\bar{q}_i}\right)\{\bar{\mathbf{x}}_{ri}-\bar{\mathbf{x}}_{si}\}+Var\left(\bar{e}_{ri}\right)+\frac{1}{n_i^2}\sum_{j\in s_i}E\left\{\omega_{ij}^{\psi}\phi\left(\frac{y_j-\mathbf{x}_j^T\boldsymbol{\beta}_{\bar{q}_i}}{\omega_{ij}^{\psi}}\right)\right\}^2\right].\quad\text{(A6)}$$

The corresponding estimator of the MSE of MQ BC, when $q=\hat{\bar{q}}$, is therefore:

$$mse(\hat{\bar{y}}_i^{MQ-BC})=\left(1-n_iN_i^{-1}\right)^2\left[\{\bar{\mathbf{x}}_{ri}-\bar{\mathbf{x}}_{si}\}^T\hat{V}\left(\hat{\boldsymbol{\beta}}_{\bar{q}_i}\right)\{\bar{\mathbf{x}}_{ri}-\bar{\mathbf{x}}_{si}\}+\hat{V}\left(\bar{e}_{ri}\right)+\frac{1}{n_i^2}\sum_{j\in s_i}\left\{\omega_{ij}^{\psi}\phi\left(\frac{y_j-\mathbf{x}_j^T\hat{\boldsymbol{\beta}}_{\bar{q}_i}}{\omega_{ij}^{\psi}}\right)\right\}^2\right].\quad\text{(A7)}$$

References

Chambers, R.L. (1986) Outlier robust finite population estimation. *J. Am. Statist. Ass.*, **81**, 1063-1069.

Chambers, R.L., Dorfman, A.H. and Wehrly, T.E. (1993) Bias robust estimation in finite populations using nonparametric calibration. *J. Am. Statist. Ass.*, **88**, 268-277.

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

Chambers, R., Chandra, H. and Tzavidis, N. (2009) On bias-robust mean squared error estimation for linear predictors for domains. Working Papers, 09-08, Centre for Statistical and Survey Methodology, The University of Wollongong, Australia. (Available from http://cssm.uow.edu.au/publications.)

Chandra H., Salvati N. and Chambers R. (2007) Small area estimation for spatially correlated populations- a comparison of direct and indirect model-based methods. *Statist. Trans.*, **8**, 887-906.

Chandra, H. and Chambers, R. (2009) Multipurpose weighting for small area estimation. *J. Off. Stat.*, **25 (3)**, 379-395.

Chatterjee, S., Lahiri, P. and Huilin, L. (2008) Parametric bootstrap approximation to the distribution of EBLUP and related prediction intervals in linear mixed models. *Ann. Statist.*, **36**, 1221-1245.

Fellner, W. H. (1986) Robust estimation of variance components. Technometrics, 28, 51-60.

Huber, P. (1981) *Robust Statistics*. J. Wiley, New York.

Longford, N.T. (2007) On standard errors of model-based small-area estimators. *Survey Method.*, **33**, 69-79.

Prasad, N.G.N. and Rao, J.N.K. (1990) The estimation of the mean squared error of small area estimators. *J. Am. Statist. Ass.*, **85**, 163-171.

Royall, R.M. and Cumberland, W.G. (1978) Variance estimation in finite population sampling. *J. Am. Statist. Ass.*, **73**, 351 - 358.

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

Richardson, A.M. and Welsh, A.H. (1995) Robust restricted maximum likelihood in mixed linear models. *Biometrics,* **51**, 1429-1439.

Salvati, N., Tzavidis, N., Pratesi, M. and Chambers, R. (2011) Small area estimation Via M-quantile geographically weighted regression. *TEST,* DOI: 10.1007/s11749-010-0231-1.

Salvati, N., Chandra, H., Ranalli, M.G. and Chambers, R. (2010). Small Area Estimation Using a Nonparametric Model Based Direct Estimator. *Comput. Statist. Data Anal.*, **54**, 2159-2171.

Sinha, S.K. and Rao, J.N.K. (2009) Robust small area estimation. *Canad. J. Stat.*, **37**, 381-399.

Street, J. O., Carroll, R. J., and Ruppert, D. (1988) A note on computing robust regression estimates via iterative reweighted least squares. *Am. Statist*, **42**, 152–154.

Tzavidis, N., Marchetti, S. and Chambers, R. (2010) Robust prediction of small area means and distributions. *Aust. & NZ J. Stat.*, **52**, 167-186.

Welsh, A.H. and Richardson, A.M. (1997) Approaches to the robust estimation of mixed models. G. S. Maddala and C. R. Rao, eds., *Handbook of Statistics,* **15**, Chapter 13.

Welsh, A.H. and Ronchetti, E. (1998) Bias-calibrated estimation from sample surveys containing outliers. *J. R. Statist. Soc.* B, **60**, 413-428.

Table 1. Model-based simulation results: performances of predictors of small area means.

Scenario/Areas	No outliers	Individual outliers	Area outliers		Both types	
	[0,0]	[e,0]	[0,u]/1-36	[0,u]/37-40	[e,u]/1-36	[e,u]/37-40
Estimator	*Median values of Relative Bias (expressed as a percentage)*					
EBLUP	0.019	-0.019	0.097	-0.536	0.166	-1.592
REBLUP	0.027	-0.391	0.108	-0.468	-0.296	-0.998
MQ	0.020	-0.428	0.088	-0.942	-0.323	-0.988
REBLUP-BC	0.022	-0.286	0.026	0.020	-0.276	-0.318
MQ-BC	0.022	-0.276	0.030	-0.068	-0.262	-0.297
	Median values of Relative Root MSE (expressed as a percentage)					
EBLUP	0.805	1.215	0.854	0.966	1.369	2.389
REBLUP	0.822	1.008	0.842	1.019	0.985	1.436
MQ	0.824	1.030	0.833	1.464	1.008	1.570
REBLUP-BC	0.913	1.232	0.918	0.859	1.240	1.270
MQ-BC	0.913	1.238	0.915	0.931	1.256	1.486

Table 2. Performance of Root MSE estimators in model-based simulation experiments.

	Scenario/Areas	No outliers	Individual outliers	Area outliers		Both types	
		[0,0]	[e,0]	[0,u]/1-36	[0,u]/37-40	[e,u]/1-36	[e,u]/37-40
Estimator	MSE Estimator	*Median values of Relative Bias (expressed as a percentage)*					
EBLUP	Prasad-Rao	-0.34	1.74	3.82	-17.31	11.32	-40.86
	CCT	3.61	31.24	1.55	2.15	5.95	-3.05
	CCST	5.64	33.95	4.78	77.26	8.52	8.28
REBLUP	CCT	-17.71	-15.76	-20.24	-34.79	-19.51	-36.63
	CCST	-2.01	-8.46	-5.31	-3.58	-7.91	-22.51
	Bootstrap	-1.19	-4.42	7.38	-19.42	11.37	-31.44
MQ	CCT	-2.98	-16.29	-12.56	6.69	-24.02	177.42
	CCST	0.11	-8.21	-7.77	8.95	-14.10	163.38
REBLUP-BC	CCT	-10.56	-12.46	-11.88	-10.54	-12.57	-18.37
	CCST	12.98	7.79	12.19	13.63	7.90	4.67
	Bootstrap	-0.21	-6.76	-0.52	-1.25	-4.90	-12.96
MQ-BC	CCT	-6.35	3.48	-7.19	3.92	1.87	5.96
	CCST	-7.18	-11.38	-7.42	3.21	-11.42	-9.20
		Median values of Relative Root MSE (expressed as a percentage)					
EBLUP	Prasad-Rao	6.24	18.57	7.20	17.90	22.28	43.19
	CCT	31.51	76.20	31.25	28.37	61.57	51.30
	CCST	26.65	66.72	15.20	88.30	29.28	39.97
REBLUP	CCT	29.52	30.82	28.67	28.58	29.00	38.70
	CCST	27.86	28.47	20.89	22.87	20.25	29.24
	Bootstrap	10.27	34.92	10.67	14.62	16.61	33.04
MQ	CCT	61.94	61.50	59.88	43.76	59.67	205.30
	CCST	54.77	49.14	50.63	40.58	45.34	189.92
REBLUP-BC	CCT	33.64	45.20	33.21	33.56	45.48	47.18
	CCST	38.14	51.03	37.65	37.63	50.34	53.71
	Bootstrap	10.12	15.27	10.20	10.60	14.53	18.35
MQ-BC	CCT	36.68	65.37	36.19	38.33	65.70	64.26
	CCST	33.93	44.81	33.55	35.30	44.65	50.55

Table 3. Median values of the relative bias (RB) and relative root mean squared error (RRMSE) of point estimators in the design-based simulation. All values are expressed as percentages and are averaged over the regions of interest.

Estimator	RB(%)	RRMSE(%)
EBLUP	10.79	35.18
REBLUP	-13.08	30.59
MQ	-22.98	35.07
REBLUP-BC	-4.13	31.94
MQ-BC	-6.17	31.57

Table 4. Performance of Root MSE estimators in design-based simulation: median values of the percentage relative bias and relative root MSE.

Estimator\MSE estimator	Prasad-Rao	CCT	CCST	Bootstrap
Median values of Relative Bias (expressed as a percentage)				
EBLUP	6.37	1.79	5.85	
REBLUP		-23.06	3.59	32.12
MQ		-31.59	-24.48	
REBLUP-BC		-14.58	3.51	0.48
MQ-BC		-6.40	-11.01	
Median values of Relative Root MSE (expressed as a percentage)				
EBLUP	30.61	30.67	28.16	
REBLUP		45.79	43.72	61.95
MQ		62.19	55.88	
REBLUP-BC		39.78	43.13	39.81
MQ-BC		45.53	38.38	

Figure 1. Area specific values of true RMSE (solid line with □) and average estimated RMSE (dashed line) obtained in the design-based simulation. Values for the PR estimator are indicated by △, those for the CCT estimator are indicated by ○, and those for the CCST estimator are indicated by +. Plots show results for the EBLUP predictor.

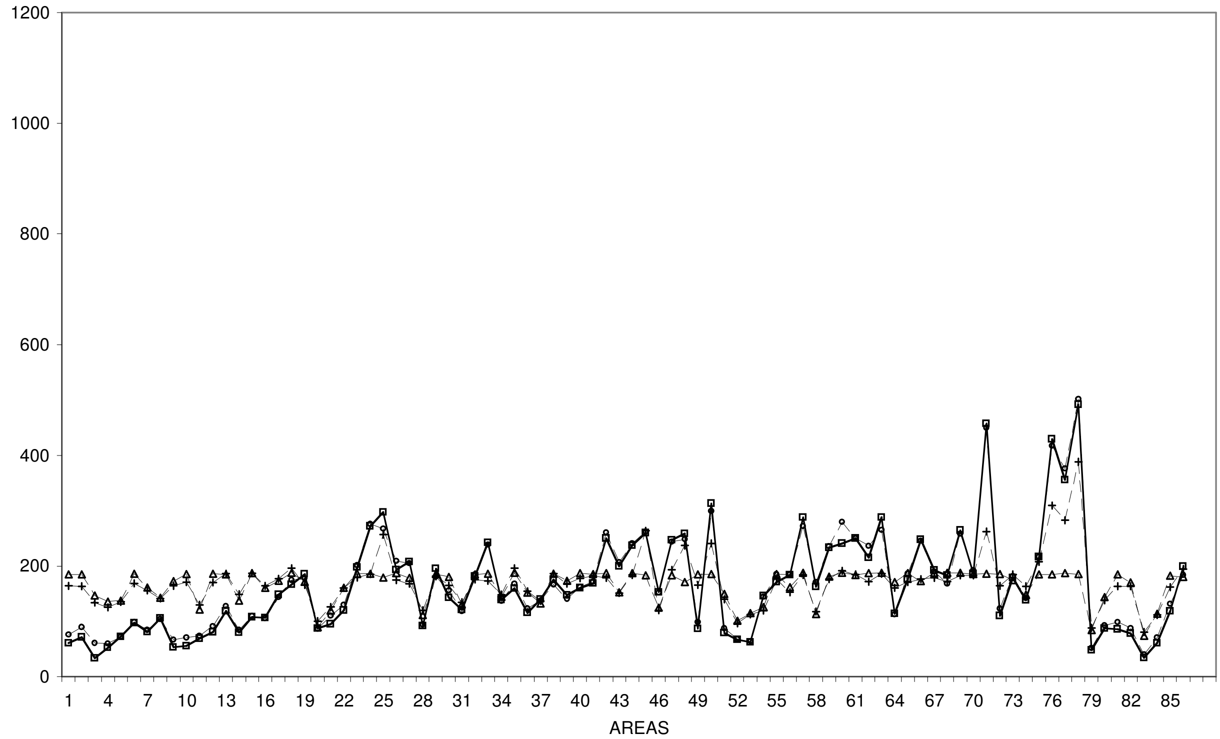
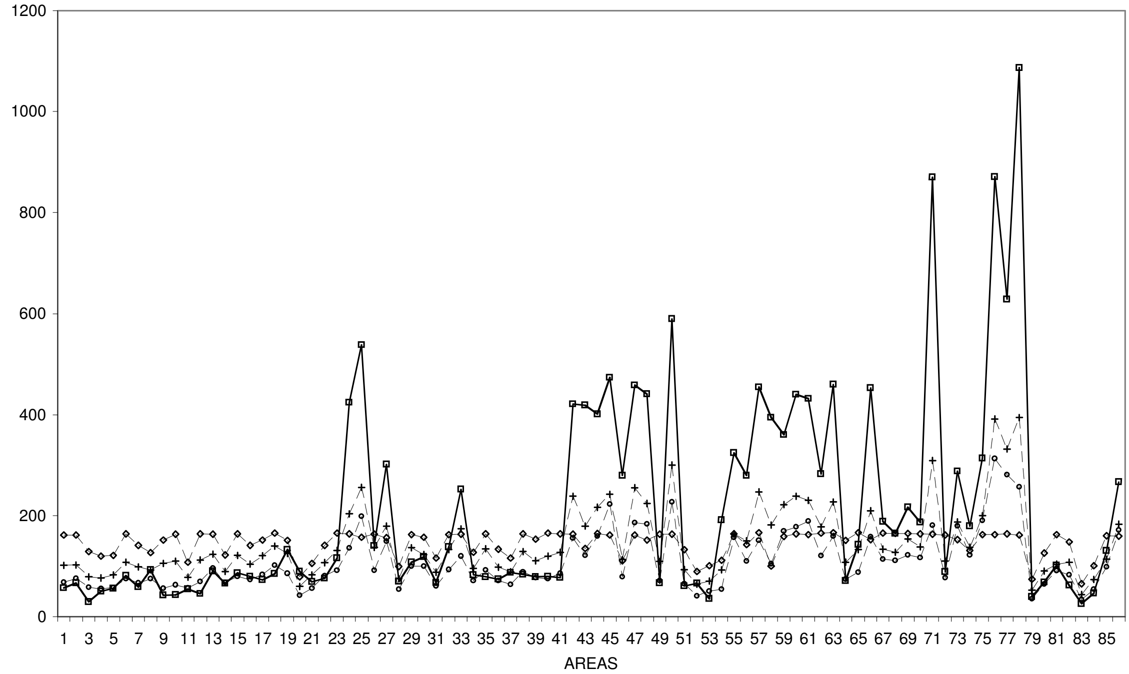


Figure 2. Area specific values of true RMSE (solid line with □) and average estimated RMSE (dashed line) obtained in the design-based simulation. Values for the CCT estimator are indicated by ○, those for the CCST estimator are indicated by +, while those for the MSE bootstrap estimator are indicated by ◇. Plots show results REBLUP (top) and REBLUP-BC (bottom) predictors.



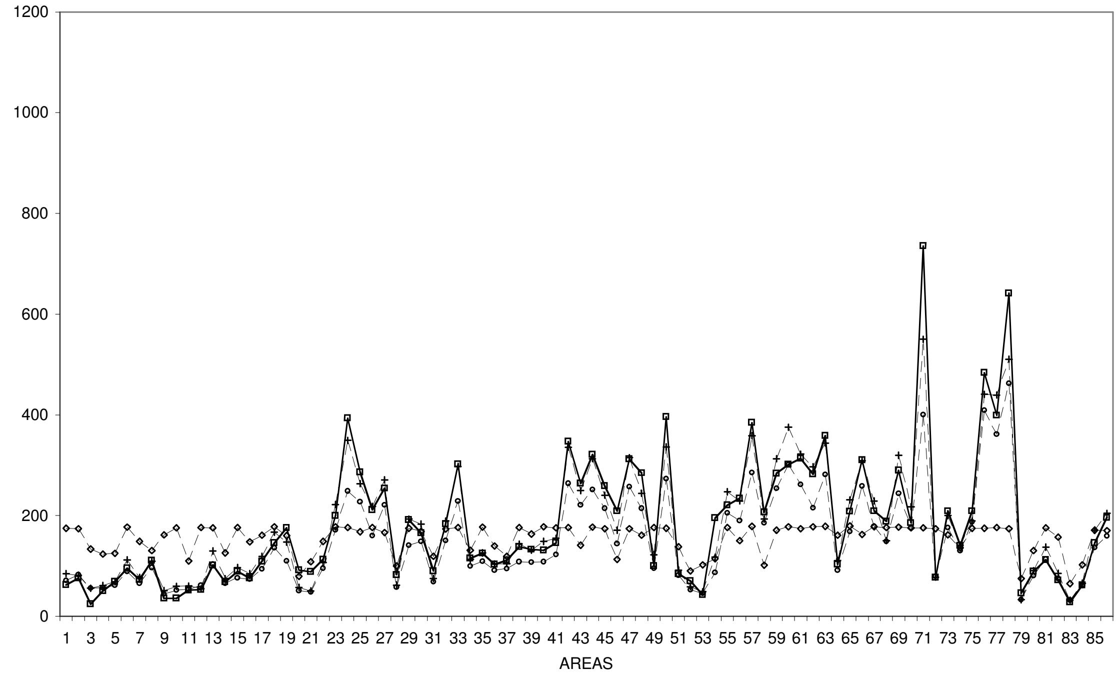


Figure 3. Area specific values of true RMSE (solid line with □) and average estimated RMSE (dashed line) obtained in the design-based simulation. Values for the CCT estimator are indicated by ○, while those for the CCST estimator are indicated by +. Plots show results MQ (top) and MQ-BC (bottom) predictors.